# Dissect SAM: Catch observations

Anders Nielsen and Olav Nikolai Breivik

# Dissect SAM: Catch observations

SAM assumes

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

were

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

Observe:

$$\log C_{a,y} = \log \left( \frac{F_{a,y}}{F_{a,y} + M_{a,y}} (1 - e^{-F_{a,y}-M_{a,y}}) N_{a,y} \right) + \epsilon^{c}_{a,y}$$

$$\log I^{(s)}_y = \log(Q^{(s)}_a e^{-(F_{a,y}+M_{a,y})day^{(s)}/365} N_{a,y}) + \epsilon^{s}_{a,y}$$

Assumes $\boldsymbol{\eta}_y$, $\boldsymbol{\xi}_y$ and $\boldsymbol{\epsilon}^C_y$ and $\boldsymbol{\epsilon}^s_y$ all Gaussian distributed.

# Observations

- Observations provides information about the system trough the observation equations:

$$\log C_{a,y} = \log\left(\frac{F_{a,y}}{F_{a,y}+M_{a,y}}(1-e^{-F_{a,y}-M_{a,y}})N_{a,y}\right)+\epsilon^{c}_{a,y}$$

$$\log I^{(s)}_{y} = \log(Q^{(s)}_{a}e^{-(F_{a,y}+M_{a,y})day^{(s)}/365}N_{a,y})+\epsilon^{s}_{a,y}$$

- When we have set up the system for **F** and **N**, we indirectly gain insight about those quantities trough observations and the observation equations.

# Catch equation

We observe catch:

$$\log C_{a,y} = \log\left(\frac{F_{a,y}}{F_{a,y}+M_{a,y}}(1-e^{-F_{a,y}-M_{a,y}})N_{a,y}\right) + \epsilon^{c}_{a,y}$$

where $\epsilon^c_y \sim N(0, \mathbf{\Sigma}_C)$.

Note

- $\frac{F_{a,y}}{F_{a,y}+M_{a,y}}$ is the proportion died in fishery
- $(1-e^{-F_{a,y}-M_{a,y}})N_{a,y}$ is the total amount of fish died

# Exercise: Catch observations

- We will now investigate the implementation of the catch equation.
- Data about **F** and **N** are provided in `Cobs.RData`.
  - It may sound artificial that we know these quantities, but now we will now only focus on the catch equation.
- Exercise a) Implement the model

$$\log C_{a,y} \sim N(\log \hat{C}_{a,y}, \sigma^2)$$

where

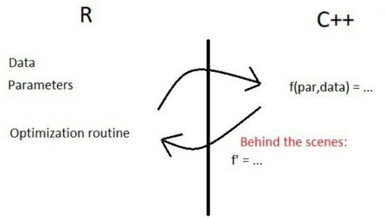
$$\begin{aligned}\log \hat{C}_{a,y} &= \log\left(\frac{F_{a,y}}{F_{a,y}+M_{a,y}}(1-e^{-F_{a,y}-M_{a,y}})N_{a,y}\right)\\ &= \log F_{a,y} - \log(F_{a,y}+M_{a,y}) + \log(1-e^{-F_{a,y}-M_{a,y}}) + \log N_{a,y}\end{aligned}$$

- Exercise b) Include separate variance parameter per age.

# Exercise: Catch observations

TMB output while fitting the model:

```
> fit <- nlminb(obj$par, obj$fn, obj$gr)
outer mgc:  502.1106
outer mgc:  375.0361
outer mgc:  300.1951
outer mgc:  16.76699
outer mgc:  8.446832
outer mgc:  0.1371361
outer mgc:  0.00109791
outer mgc:  1.442425e-07
>
```

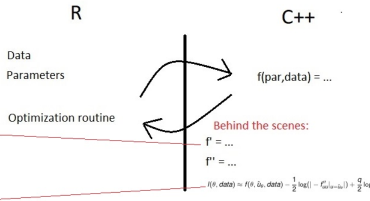



- Explain the output

# Note

In the full model, optimal **F** and **N** are found in the inner optimization.

```
outer mgc:  2.125321
  4:     158.72116: -0.405600  1.00234
iter: 1  value: 167.6788 mgc: 0.9840473 ustep: 1
iter: 2  value: 167.6786 mgc: 0.01624102 ustep: 1
iter: 3  value: 167.6786 mgc: 1.878016e-05 ustep: 1
iter: 4  mgc: 3.701173e-11
iter: 1  value: 153.1634 mgc: 0.05411768 ustep: 1
iter: 2  value: 153.1634 mgc: 0.0001761833 ustep: 1
iter: 3  mgc: 1.925322e-09
iter: 1  value: 153.1634 mgc: 0.05411768 ustep: 1
iter: 2  value: 153.1634 mgc: 0.0001761833 ustep: 1
iter: 3  mgc: 1.925322e-09
outer mgc:  0.4393079
  5:     158.64804: -0.348985  1.00601
iter: 1  value: 150.4947 mgc: 0.2225958 ustep: 1
iter: 2  value: 150.4947 mgc: 0.0004934857 ustep: 1
iter: 3  mgc: 9.436551e-09
iter: 1  value: 152.9936 mgc: 0.006990456 ustep: 1
iter: 2  value: 152.9936 mgc: 6.445189e-07 ustep: 1
iter: 3  mgc: 1.110223e-14
iter: 1  value: 152.9936 mgc: 0.006990456 ustep: 1
iter: 2  value: 152.9936 mgc: 6.445189e-07 ustep: 1
iter: 3  mgc: 1.110223e-14
outer mgc:  0.01172901
```



# Catch observation configurations in SAM

- Use `keyVarObs` to modify the catch observation variance ($\Sigma_{a,a}$)
- First line is typically the catch
- Example with same variance parameter for all ages:

```
$keyVarObs
   0   0   0   0   0   0   0   0   0   0   0
  -1   1   1   1   1   1   1   1   1   1  -1
  -1   2   2   2   2   2   2   2   2   2  -1
   3   3   3   3   3   3   3   3   3   3  -1
```

# Catch observation configurations in SAM

- We can construct a link between predicted observations and associated variance in SAM

Let $\mu_{a,y}$ be the predicted observation for age $a$ at year $y$ on natural scale, and let $v_{a,y}$ be the corresponding variance. With this option we impose the assumption that

$$v_{a,y} = \alpha \mu_{a,y}^{\beta},$$

and estimate $\alpha$ and $\beta$ internally in SAM.

- `keyVarObs` couples the $\alpha$ parameters.
- `predVarObsLink` couples the $\beta$ parameters.

Exercise: Include this structure in the previous exercise

- Hint: Variance of log observation is given by:

$$\sigma_{a,y}^2 = \log\left(\alpha_a \mu_{a,y}^{\beta_a - 2} + 1\right).$$

# External observation variances in SAM

- We may know an external estimate of the uncertainty of $\log \hat{C}_{a,y}$
- We can include such data as a relative weighting factor for each observation
- See the vignette `obsCovarOptions` for details

# Exercise: Include external covariance structures

A script for fitting SAM to North East Arctic cod is provided in the folder `externalCovarianceEx`.

- Externally estimated log catch covariance matrices are provided in `covCatch.RData`
- A SAM script without using external variances are provided in `script.R`
- Exercise a: Read the vignette `obsCovarOptions` and include the external covariance matrices in the assessment.
- Exercise b: Based on AIC, do we obtain an improvement?

# Total catches

- We inform the system with observations trough observations equations.
- Need to express observations as realisations of a probability distribution with parameters provided by parameters in the model (e.g. **F** and **N**).
- Say we only observe total catch, and not per age.
- We can include that observation trough:

$$\log C_y = \log\left(\sum_a \frac{F_{a,y}}{F_{a,y}+M_{a,y}}(1-e^{-F_{a,y}-M_{a,y}})N_{a,y}cw_{a,y}\right)+\epsilon_y^C$$

where $\epsilon_y^C \sim N(0,\sigma_y^2)$.

  - Here $cw_{a,y}$ is mean catch weight at age $a$ in year $y$.

# Total catches

- Can use the delta method for constructing variance of $\epsilon_y^{Catch}$:

$$\sigma_y^2 = h_y^t \Sigma_C h_y$$

where

$$h_y = \left(\frac{\hat{C}_{1,y} cw_{1,y}}{\sum_a \hat{C}_{a,y} cw_{a,y}}, \cdots, \frac{\hat{C}_{A,y} cw_{A,y}}{\sum_a \hat{C}_{a,y} cw_{a,y}}\right)$$

and

$$\hat{C}_{A,y} = \frac{F_{a,y}}{F_{a,y} + M_{a,y}}(1 - e^{-F_{a,y} - M_{a,y}})N_{a,y}.$$

- We can now include total catch in years when catch at age is not available.
- Procedure in SAM: Include total catch as a survey (of age -1), set `keyBiomassTreat =3` and set catch per age to -1 in the corresponding years
- Similarly we can include total landings (`keyBiomassTreat = 4`)

# Biomass indices

- Observations of total spawning stock biomass
- SSB observations can be represented internally as:

$$\log \mathrm{SSB}_y = \log\left(\sum_a Q_a^{(s)} e^{-(F_{a,y}+M_{a,y})day^{(s)}/365} N_{a,y} m_{a,y} w_{a,y}\right) + \epsilon_y^{SSB}$$

  where $m_{a,y}$ is proportion mature and $w_{a,y}$ is mean stock weight.
- Procedure in SAM: Including as survey of age -1 and set `keyBiomassTreat = 0`.
- Similarity we can include:
  - Catch index (`keyBiomassTreat = 1`)
  - FSB index (`keyBiomassTreat = 2`)
  - TSB index (`keyBiomassTreat = 5`)