# Non-standard modeling

Olav Nikolai Risdal Breivik

# Non-standard modeling

- RTMB is a general tool
- The likelihood needs to be differentiable w.r.t. model parameters
- Posterior distribution of latent effects needs to be *approximately* Gaussian
- Validate your model:
  - always do a simulation study
  - always do a jitter analysis
  - always investigate residuals
- Do not let other packages limit you!

# Example: State space assessment model

The state space assessment model SAM is widely used by the International Council for the Exploration of the Sea

- Used to provide quotas for approximately 40-50 fish species world wide
- Combine many sources of data
- Population dynamics
- Includes latent effects
- It is a non-standard model
  - packages like **`mgcv`** or **`INLA`** cant be used

# State space assessment model

SAM assumes:

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

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.

# Spline

- Can use `mgcv` to set up P-splines
- See `spline/pSpline.R` for an example

```
gam_setup = gam(Richness ~ s(ROCK, bs = "cs") ,...)
S = gam_setup$smooth[[1]]$S[[1]]
X = gam_setup$X[,-1]  # Design matrix, without intercept

...
#Inside objective function:
nll = nll- (0.5*m*log_lambda - 0.5*lambda*t(beta)% * % S  % * %beta);
spline = X % * %beta;

...
obj = RTMB::MakeADFun(f, par,random="beta")
```

- Now you can include a P-spline in a non-standard model!