Modelling of age-at-size data

This document explains how we model the length-stratified empirical dataset collected across multiple survey dates. The method convolves a single-cohort simulation with an annual spawning distribution and an age-to-ring mapping function to predict the age distribution at length on any given survey day.

Cohort evolution

We describe the population density $u(l,t)$ for $l > l_0$ (where l_0 is the egg size) with the PDE

$$\frac{\partial u}{\partial t} = -\frac{\partial J}{\partial l} - \mu u$$

where the flux J is

$$J(l) = k(L_\infty^0 - l)u - \alpha l \frac{\partial u}{\partial l}$$

and the mortality μ is μ_0/l. The positive parameters k, L_∞^0, d and μ_0 need to be estimated from data. The superscript 0 on L_∞ indicates that this is the asymptotic length in the absence of diffusion. The average growth rate is given by

$$g(l) = k(L_\infty^0 - l)$$

with $L_\infty = L_\infty^0 + \alpha/k$ being the asymptotic length when diffusion is included.

This PDE allows us to calculate the temporal evolution of a single cohort of fish born at the same instant. To do this numerically we discretize the size into length bins $l_1, l_2, \dots, l_{n_l}$ and at time zero place all individuals into the smallest size bin. From the output of this simulation we get the impulse response matrix $\mathbf{G}$, where the element

$$G(l,a) = n(l,a)dl$$

is the number of fish of size bin l (with width dl) that have survived to age a from the initial pulse of recruits.

Spawning season

We now need to incorporate the information about how the spawning intensity varies through the year. We describe this with a function $S(\tau)$, where $\tau \in [0,1]$ represents the fractional time of year.

The expected number of fish of size l and true age a in the population on a given survey date, t_{survey}, is found by weighting the impulse response by the spawning intensity S. Let $\tau(t)$ be the function that returns the fractional time of year for any date t. The population structure, $\mathbf{N}_{pop}$, is given by:

$$N_{pop}(l,a|t_{survey}) = G(l,a) \cdot S(\tau(t_{survey} - a))$$

This convolution integrates the fate of all cohorts born throughout the year into a single snapshot of the population.

One way to model the relative intensity of spawning throughout the year is by a von Mises probability density function:

$$S(\tau) = \frac{e^{\kappa \cos(2\pi(\tau-\mu))}}{2\pi I_0(\kappa)}$$

where μ is the mean time of year for spawning, κ is the concentration parameter, and $I_0(\kappa)$ is the modified Bessel function of the first kind of order zero, which serves as a normalization constant.

Annuli from true age

The number K of observed otolith rings is a deterministic function of a fish's true age a and the date of the survey, t_{survey}. This mapping, $\mathcal{K}(a, t_{survey})$, is defined as:

$$\mathcal{K}(a, t_{survey}) = \sum_{y=Y_{birth}+1}^{Y_{survey}-1} \mathbb{I}(t_{ring,y} - t_{birth} \geq a_{min})$$

where:

- $t_{birth} = t_{survey} - a$ is the birth date.
- Y_{birth} and Y_{survey} are the years of birth and survey, respectively.
- $t_{ring,y}$ is the date of ring formation in year y.
- a_{min} is the minimum age required for the first ring to form.
- $\mathbb{I}(\cdot)$ is the indicator function, which is 1 if the condition is true and 0 otherwise.

The expected number of fish of size l that would be observed with K rings, $\mathbf{N}_{model}$, is obtained by aggregating the true population structure according to the age-to-ring mapping function:

$$N_{model}(l, K|t_{survey}) = \int_0^{a_{max}} N_{pop}(l, a|t_{survey}) \cdot \mathbb{I}(\mathcal{K}(a, t_{survey}) = K) \, da$$

In the discrete-time implementation, this integral becomes a sum over all age steps a_i where $\mathcal{K}(a_i, t_{survey}) = K$.

Predicted age distribution

For a given survey date t_{survey}, the probability of a fish in size class l having K rings is:

$$P(K|l, t_{survey}) = \frac{N_{model}(l, K|t_{survey})}{\sum_{K'} N_{model}(l, K'|t_{survey})}$$

Likelihood of observed data

Let the empirical data for a survey on date $t_{survey,j}$ consist of a set of counts $C(K|l, j)$. Let $n_{l,j}$ is the number of fish sampled from size class l. We model these observations as a draw from a multinomial distribution for each size class and each survey:

$$(C(K = 0|l, j), C(K = 2|l, j), ..., C(K = n_K|l, j) \sim \text{Multinomial}(n_{l,j}, \mathbf{p}_{l,j})$$

where $\mathbf{p}_{l,j}$ is the vector of probabilities $[P(K = 0|l, t_{survey,j}), P(K = 1|l, t_{survey,j}), ...]$.

The total negative log likelihood for the observations is then

$$NLL = \sum_j \sum_l \sum_k C_{K=k|l,j} \log(P(K = k|l, t_{survey,j}).$$