distribution of the stock, we cannot estimate the absolute selectivity, but only the relative selectivity between the two gears. This corresponds to enforcing $S_{1k} = -S_{2k}$

Given these random variables $\Phi$, $R$, $S$, we assume that count data is Poisson distributed:

$$N_{ijk}|\Phi, R, S \sim \text{Poisson}(A_{ij} \cdot \exp(\Phi_{ik} + S_{jk} + R_{ijk}))$$

The swept area $A_{ij}$ is an input to the model. The unobserved random variables, $\Phi$, $R$ and $S$, are given prior distributions: The size spectrum at each station, i.e. $\Phi_{ik}$, is considered a random walk over size groups:

$$\Delta\Phi_{ik} \sim N(0, \sigma_\Phi^2) \text{ for } k = 1, \ldots, n_l - n_\Phi \quad .$$

Here, $\Delta$ is the difference operator. This enforces continuity in the size spectrum; the most probable spectrum is flat. To ensure that the spectrum is a well defined stochastic process, we complement this with initial conditions

$$\Phi_{ik} \sim N(0, \sigma_1^2) \text{ for } k = 1 \quad .$$

Here, the variance $\sigma_1^2$ is fixed at a "large" value 10. In contrast, the parameter $\sigma_\Phi^2$ is estimated. We assume independence between stations, i.e. we do not attempt to model any large-scale spatiotemporal structure of the population. We note that this is the main difference between this model and the GeoPop model, where emphasis is exactly on this spatiotemporal structure.

The residual or "nugget effect" $R_{ijk}$ models size-structured clustering of the fish at small spacial and temporal scales. Thus, this effect is independent between hauls, even those taken at same station $i$ but with different gear $j$. For a given haul, i.e. for given station $i$ and gear $j$, the nugget effect is a mean 0 first order autoregressive process of size, with a variance $\sigma_N^2$ and correlation coefficient $\phi$ which is estimated.

The relative selectivity $S_{jk}$, which we aim to estimate, is modeled as a random walk in size:

$$\Delta S_{jk} \sim N(0, \sigma_S^2) \text{ for } k = 1, \ldots, n_l - n_S$$

We assume infinite variance on the first size group, $S_{j1}$, i.e. only the increments in the selectivity process enter into the likelihood function.

| | Africana Old | | Africana New | |
|---|---|---|---|---|
| Parameter | Estimate | Std. Error | Estimate | Std. Error |
| $\log \sigma_\Phi$ | 0.077 | 0.02 | 0.17 | 0.02 |
| $\phi$ | 0.927 | 0.01 | 0.92 | 0.01 |
| $\log \sigma_N$ | -0.039 | 0.05 | -0.08 | 0.05 |
| $\log \sigma_S$ | -3.145 | 0.25 | -2.76 | 0.24 |

Table 1: Parameter estimates

**SI 1.2: Implementation**

The statistical model in the previous defines the joint distribution of the count data, $N$, and the unobserved random variables $\Phi$, $R$, $S$, for given parameters $\sigma_S$, $\sigma_\Phi$, and the two parameters (scale and range) defining the nugget effect. The unobserved $\Phi$, $R$ and $S$ are integrated out using the Laplace approximation, to yield the likelihood function as a function of the four parameters. The likelihood function is maximized to yield estimates of the four parameters, after which the posterior means of the $\Phi$, $R$, and in particular $S$ are reported.

The computations are performed in `R` version 3.1.2 (R Core Team, 2015); we use the Template Model Builder (TMB) package (Kristensen et al., 2016) for evaluating the likelihood function and its derivatives.

**SI 1.3: Data**

The data base consisted of a total of 236 pairs of trawl hauls performed by RVs Africana and Dr. Fridtjof Nansen. The Gisund gear was used onboard Fridtjof Nansen, while RVs Africana deployed two gear types: "Africana Old" (108 hauls) and "Africana New" (128 hauls). Catch in numbers per length group and the swept area (hauling distance multiplied by wing spread) were available for each haul.

**SI 1.4: Results**

The obtained intercalibration curves are seen in figure 1. Notice that Gisund overall is more effective than both the Old and the New Africana, in particular in the small size classes. The difference between size classes is statistically significant ($p < 10^{-4}$). For size classes larger than 30 cm, say, the intercalibration curves show little variation with size and although this has not been tested, it is plausible that these variations are not statistically significant. Estimated parameters are seen in table 1.

# References

Kristensen K, Nielsen A, Berg C, Skaug H (2016) TMB: Automatic differentiation and Laplace approximation. *J. Stat. Softw* 70.

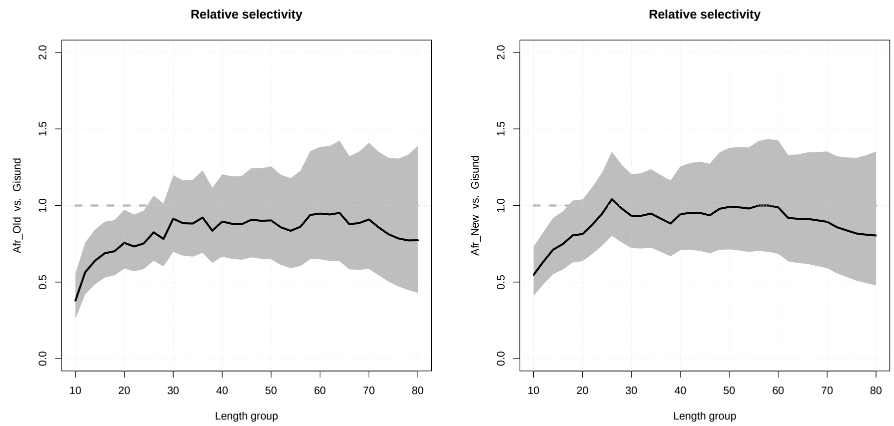


Figure 1: Relative selectivity (gear calibration factor), comparing catches of M. paradoxus with Gisund gear and the “Old” and “New” gear on the R/V Africana. Large values indicate that Africana is more effective. Solid curve: Estimated relative selectivity (posterior mode). Grey region: Marginal 95 % confidence intervals.

R Core Team (2015) *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing, Vienna, Austria.