Intercalibration of survey trawl gear using paired hauls

Draft prepared by Uffe Høgsbro Thygesen*, Kasper Kristensen, Teunis Jansen, Jan E. Beyer. Additional authors to be included.

This version compiled November 29, 2014.

Abstract

We present a statistical method for intercalibration of surveys, i.e. determining the relative selectivity of two gear types or two vessels. The relative selectivity is estimated for each size class. The method relies on data from paired trawl hauls performed with the two gear types. The method models the size spectrum of the underlying population at each station, size-structured clustering of fish at small temporal and spatial scales, as well as the relative selectivity of the two gear types in each length class. The model is based on Poisson distributed catches conditional on log-Gaussian variables that describe the expected cathches, which allows for overdispersion and correlation between catch counts in neighboring size classes. We apply the method to catches of hake (M. Paradoxus and M. Capensis) in paired trawl hauls. In one case we compare two vessels, RV Dr. Fridtjof Nansen and FV Blue Sea, using the same gear. In a second case we compare the RV Africana, with “old” and “new” gear, and the RV Dr. Fridtjof Nansen, with Gisund gear. The results demonstrate that it is feasible to estimate the relative selectivity in each size class, but also that confidence limits are quite wide. From the results we are able to intercalibrate indirectly the “new” and “old” Africana gear, although no paired hauls used these two gears, but in doing so the confidence regions widen.

1 Introduction

Fishery-independent surveys are of pivotal importance for stock assessments and basic biological research. The vessels, riggings and gears applied in these

*DTU Aqua, Jægersborg Allé 1, DK-2920 Charlottenlund, Denmark, uht@aqua.dtu.dk

surveys often develop or shift over time leading to changes in selectivity and catch efficiency (Lewy, Nielsen, and Hovgård, 2004; Miller and Trenkel, 2013; Thorson and Ward, 2014). Time series and spatial distribution data can therefore not be used without accounting for the vessel-gear-specific selectivity. This is done by performing dedicated intercalibration experiments with two or more vessel-gear combinations. The relative difference in catch rates are measured by performing pairwise simultaneous deployments of the gears in the same area. This design minimizes the time-space variation of the fished population between gear deployments. Finally, species and size-specific intercalibration factors can be estimated from the sample data.

Multiple calibration procedures has been proposed and applied over time. (Lewy, Nielsen, and Hovgård, 2004) focused on the disturbance effect that a first haul has on the fished local fished population, and the implications for the second haul. The length-dependence on the relative selectivity was described using a polynomial in length, the coefficients of which was estimated in a GLM framework, including overdispersion relative to Poisson counts. Alternatives to such fixed polynomials include orthogonal polynomials or GAMs (Miller and Trenkel, 2013). A typical problem of these data is the large number of zero catches; therefore (Thorson and Ward, 2014) considered delta-GLMM's, where the probability of zero catch is explicitly modeled.

A common phenomenon for size structures in catches is that not only is the numbers in each length group overdispersed, there is also strong tendency to positive correlations between nearby size classes (Kristensen et al., 2014). The most obvious explanation for this is that individuals tend to aggregate with individuals of similar size. This applies also to non-schooling species; although the phenomenon can be explained by direct interactions between fish, it can also be caused by features in the local habitat that attract or repel individuals of certain size.

This paper presents a feasibility study of a novel method which addresses explicitly such small-scale size-structured clustering. The model is based on unobserved random functions of size, which describe the local abundance at each station, the small-scale clustering in each haul, as well as the relative selectivity, where the latter is the primary objective of inference. We develop and describe the method, and demonstrate it by applying it in two cases: First, a vessel intercalibration experiment, where the objective is to investigate differences between two vessels which use the same gear. Second, a gear-vessel intercalibration experiment where the objective is to investigate differences between two vessels which use a total of three different gears.

2 Methods

2.1 Statistical model

The intercalibration model is a statistical model which explains the size composition of the catch in survey trawl hauls. The observed quantities are count data, N_{ijk}, which represents number of individuals caught in station $i = 1, \ldots, n_s$, using gear $j = 1, 2$, and in length group $k = 1, \ldots, n_l$.

We assume that these catches are Poisson distributed, conditional on the swept area A_{ij} and three sets of random variables, which all depend on the size class: The local background size spectrum Φ_{ik}, which is specific to the station, haul-specific fluctuations R_{ijk} in the size spectrum, and the relative selectivity S_{jk} which is specific to the gear. First, Φ_{ik} represents the size composition of the fish at station i, as would be observed with a hypothetical gear with "typical" size selectivity. Specifically, $\exp(\Phi_{ik})$ is the expected number of fish caught in size group k at station i with a hypothetical gear which lies in between the two gears $j = 1$ and $j = 2$.

Next, R_{ijk} is akin to the "nugget effect" in spatial statistics, and represents small-scale clustering of fish. This is particular to both stations and gears, since the two gears are used at slightly different locations and times, and therefore these clusters have moved or regrouped between hauls at the same station.

Finally, S_{jk} is the main object of interest, and represents the selectivity of gear j in size group k. Since we do not know the actual size 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 Φ, 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, Φ, R and S, are given prior distributions: The size spectrum at each station, i.e. Φ_{ik}, is considered a random walk over size groups:

$$\Phi_{ik} - \Phi_{i(k-1)} \sim N(0, \sigma_\Phi^2) \text{ for } k > 1, \quad \Phi_{i1} \sim N(0, \sigma_1^2)$$

Here, the variance σ_1^2 is fixed at a "large" value 10, while σ_Φ^2 is estimated. This model enforces some continuity in the size spectrum, but makes no additional assumption about its shape. 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 (Kristensen et al., 2014), where emphasis is exactly on this spatiotemporal structure.

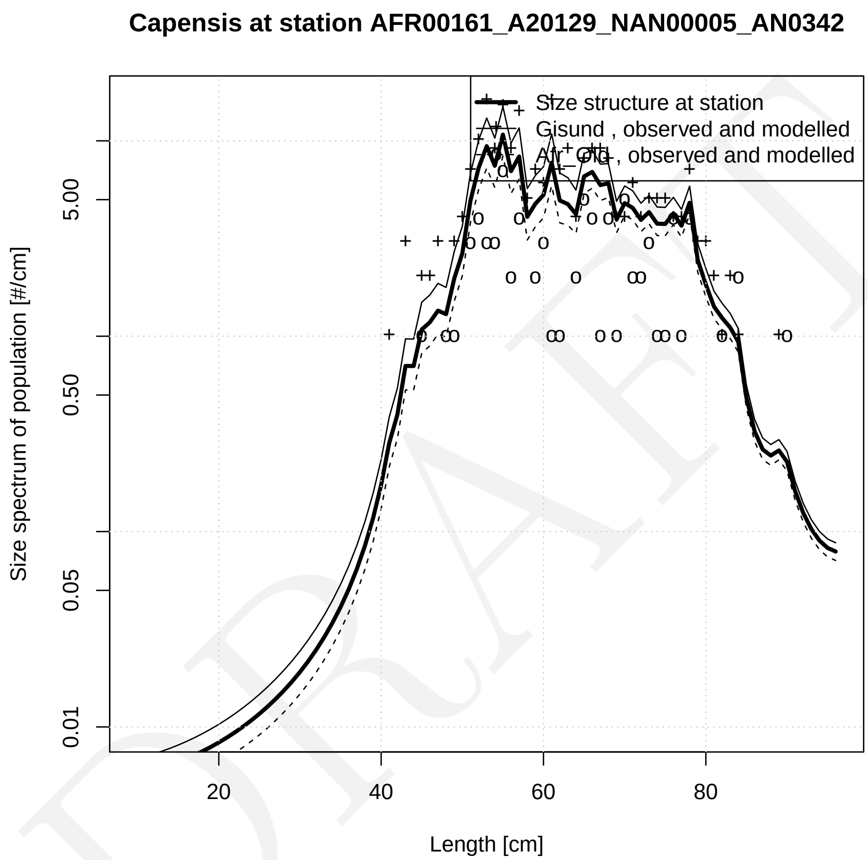


Figure 1: Example of data and model components. Estimated size spectrum Φ at one particular station (thick solid lines). Different nugget effects R apply to the two hauls and results in different size structures encountered by the two hauls (thin solid and dashed lines). Relative gear selectivity S modifies the expected catch in each size group and for each haul (not shown). Observed counts N in each size group and in each haul are shown with "o" and "+", respectively. Note log scale on the count axis; zero catches are not shown.

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 and correlation coefficient which is estimated.

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

$$S_{jk} - S_{j(k-1)} \sim N(0, \sigma_S^2)$$

We assume infinite variance on the first size group, S_{j1}.

2.2 Implementation

The statistical model in the previous defines the joint distribution of the count data, N, and the unobserved random variables Φ, R, S, for given parameters σ_S, σ_Φ, and the two parameters (scale and range) defining the nugget effect. The unobserved Φ, 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 Φ, R, and in particular S are reported.

The computations are performed in `R` version 3.1.2 (R Core Team, 2014); we use the `TMB` package (Kristensen, 2013) for evaluating the likelihood function and its derivatives.

2.3 Data

We perform two studies: First, an intercalibration study between two vessels using the same gear, and second a gear-vessel intercalibration study between two vessels using different gear.

Vessel intercalibration case

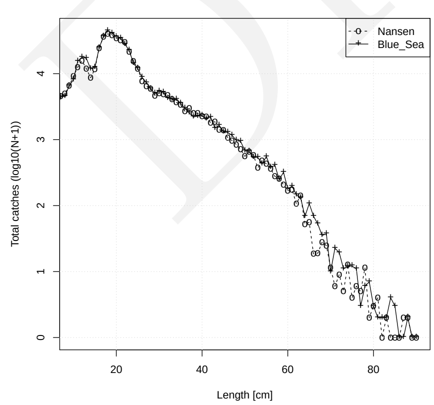
Following independence of Namibia in 1990, abundance of Namibia's hake stocks was monitored by trawl surveys conducted by the Norwegian fisheries research vessel Dr Fridtjof Nansen. From 2000 the Ministry of Fisheries and Marine Resources in Namibia (MFMR) conducted the surveys using the commercial trawler Blue Sea. In 1998 and 1999, before the shift, extensive intercalibration experiments were performed by completing the entire annual survey in parallel with both vessels. Both vessels used the same Gisund fishing gear and rigging.

Catch data collected from these surveys were extracted from the NAN-SIS database in November 2014 (Strømme, 1992). The analysis was based on 341 of the 365 pairs of trawl hauls. 24 pairs were excluded because the trawl duration were less than 15 minutes and/or the difference in trawl duration exceeded 10 minutes.

Catch in numbers per length group and the hauling distance were available for each haul. Figure 2 shows all catches, summed over all stations, for the two species.

Gear intercalibration case

The hake stocks in Namibia and South Africa has been surveyed with three different trawl gears, "Gisund", "Africana New" and "Africana Old" . A series of intercalibration experiments has made catch data available for estimation of gear catchability factors, however, estimation of species and size specific intercalibration factors has proven to be problematic (Rademeyer and Butterworth, 2013; Cotter, 2012; Brandão, Rademeyer, and Butterworth, 2004). 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. Figure 3 shows all catches, summed over all stations, for the two species.



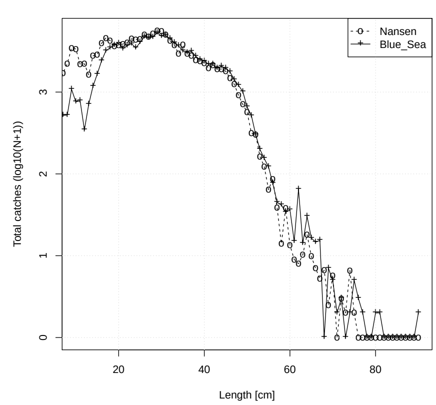


Figure 2: Total catches for the vessel intercalibration study, by size, summed over all hauls. Left panel: M. Capensis. Right panel: M. Paradoxus.

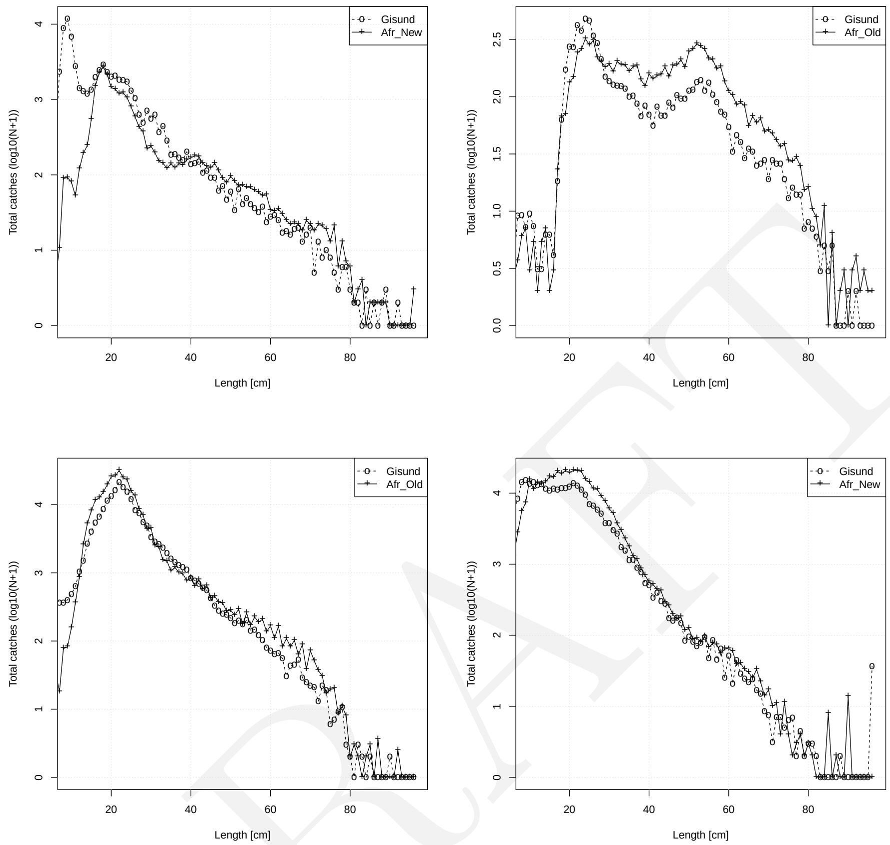


Figure 3: Total catches for the gear intercalibration study, by size, summed over all hauls. Top row: M. Capensis. Bottom row: M. Paradoxus.

3 Results

3.1 Vessel intercalibration

The comparison between RVs Dr. Fridtjof Nansen and Blue Sea is seen in figure 4. Since the gear used on the two vessels are the same, a reasonable hypothesis could have been that there is no size structure to these calibration factors. However, this hypothesis is rejected for both species ($p < 0.01$ for Capensis and $p < 0.001$ for Paradoxus). The overall patterns of relative selectivity are similar for both species, namely that the FV Blue Sea is more efficient at catching larger hakes than the RV Dr. Fridtjof Nansen. The size dependency is more pronounced for paradoxus, where the FV Blue Sea is less efficient in the small size classes. The selection of small capensis is similar for the two vessels.

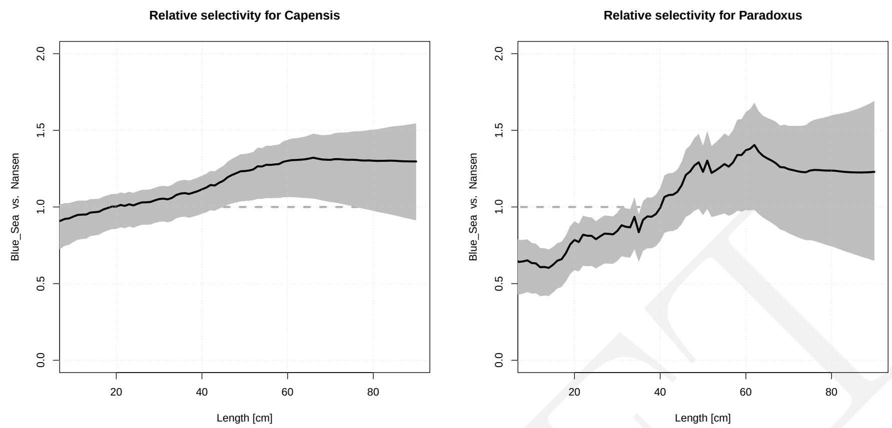


Figure 4: Relative selectivity (vessel calibration factor), comparing Gisund on RV Dr. Fridtjof Nansen and FV Blue Sea. Large values indicate that the FV Blue Sea has higher selectivity. Solid curve: Estimated relative selectivity (posterior mode). Grey region: Marginal 95 % confidence intervals. Left panel: M. Capensis. Right panel: M. Paradoxus.

3.2 Gear intercalibration for M. Capensis

For M. Capensis the intercalibration curves are estimated to be fairly smooth (figure 5). Both the Africana “Old” and “New” have selectivities below Gisund in the small size classes. The Africana “Old” has practically the same selectivity as Gisund for fish larger than 40 cm, while the Africana “New” has a selectivity which approaches that of Gisund, but still remain less. Confidence regions remain tolerable; there is a slight tendency that they widen for both small and large fish due to small catch numbers in these extreme size classes.

Since both Africana “Old” and “New” can be compared to Gisund, we can indirectly estimate the intercalibration factor between Africana “Old” and “New” (figure 7, left panel). Confidence regions are wider, due to the indirectness of this estimation, and although the New gear seems to select less, this result is only borderline significant.

3.3 Gear intercalibration for M. Paradoxus

For M. Paradoxus the intercalibration curves are estimated to be considerably more fluctuating (figure 6). Both Africana “Old” and “New” have considerably

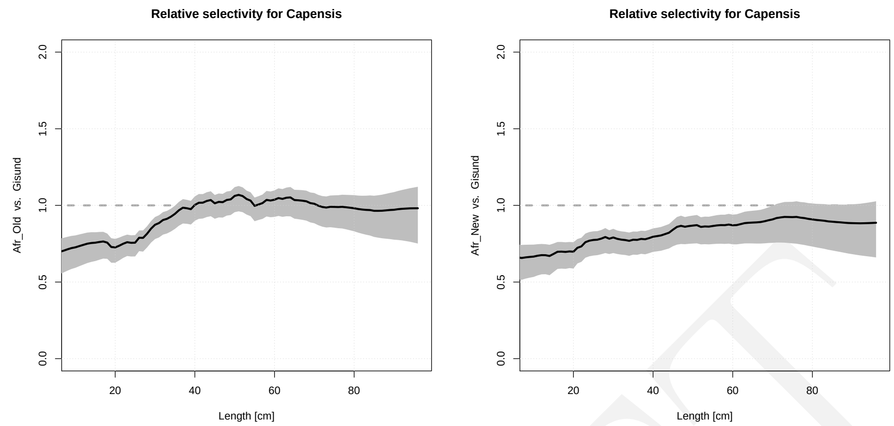


Figure 5: Relative selectivity (gear calibration factor) for M. Capensis of Africana vs Gisund. Solid curve: Estimated relative selectivity (posterior mode). Grey region: Marginal 95 % confidence intervals. Left panel: Africana "Old" gear. Right panel: Africana "New" gear.

smaller selectivity in the small size groups, due to large catches of these size classes by RV Fridtjof Nansen. Except in the smallest size classes, there is little difference in selectivity between the Africana gears and the Gisund. The confidence regions are fairly wide for large fish. In terms of absolute uncertainty, the confidence regions are quite narrow for small fish; the relative uncertainty is more moderate (not shown).

Again, we can indirectly obtain intercalibration factors between the two Africana gears (figure 7, right panel). There is some indication that the "New" gear has higher selectivity than the "Old" gear, but this result is also marginally insignificant. For very small fish (10-15 cm), the "Old" gear seems to catch much less. This result should be treated with caution since it depends on the larger catches of RV Fridtjof Nansen in this size range.

4 Discussion

It is interesting to notice that the vessel effect apparently can not be taken as a constant, but varies with the size of the fish. This also highlights that the results of the vessel-gear intercalibration study may be a combination of a gear effect and a vessel effect.

Intercalibration factors can be modeled as constants which apply to all size classes, as size-dependent functions using parametric forms, or unparametri-

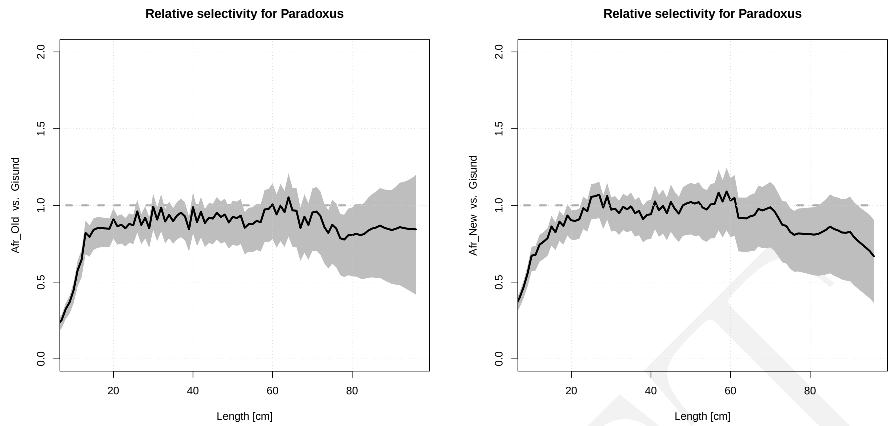


Figure 6: Relative selectivity (gear calibration factor) for M. Paradoxus of Africana vs Gisund. Solid curve: Estimated relative selectivity (posterior mode). Grey region: Marginal 95 % confidence intervals. Left panel: Africana "Old" gear. Right panel: Africana "New" gear.

cally. Here, we have taken the last approach, mostly because it is difficult to hypothesize a reasonable functional form prior to seeing the data. If we postulate a specific functional form, then it is likely that parameters in this form can be estimated with seemingly high accuracy, but it is difficult to assess the sensitivity of the results to miss-specification of the functional form. As a result, we would be prone to overestimate our confidence in the obtained intercalibration curves.

Arguably, the nonparametric curves we have obtained in this preliminary study could be just a first step. Parametric forms could be built on top of the current model. From an implementation point of view, this is a minor extension, but the issues with model validation needs to be addressed.

5 Conclusion

We have demonstrated the feasibility of estimating size-specific intercalibration factors from paired trawl hauls, using the "GeoPop" approach where data is assumed Poisson distributed while overdispersion and covariance structure is explained by unobserved random fields, which represent stock size composition, small scale size structured clustering, and gear selectivity.

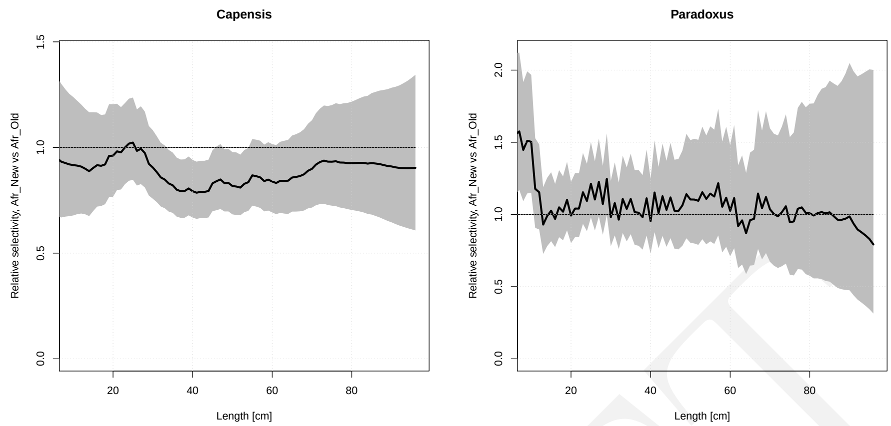


Figure 7: Relative selectivity of Africana New vs. Africana Old. Left panel: M. Capensis. Right panel: M. Paradoxus. Solid curve: Estimated relative selectivity (posterior mode). Grey region: Marginal 95 % confidence intervals.

References

Brandão, A, RA Rademeyer, and DS Butterworth (2004). First attempt to obtain a multiplicative bias calibration factor between the africana with the old and the new gear. *Unpublished report, Marine and Coastal Management, South Africa WG/11/04/D: H* 26.

Cotter, John (2012). An inter-calibration study for surveys of hake, merluccius capensis and m. paradoxus, carried out by rs africana (south africa), rs blue sea 1 (namibia), and rv dr. fridtjof nansen (norway-fao). Technical report, Benguela Current Commission.

Fryer, RJ (1991). A model of between-haul variation in selectivity. *ICES Journal of Marine Science: Journal du Conseil* 48(3): 281–290.

Kristensen, Kasper (2013). *TMB: General random effect model builder tool inspired by ADMB.* R package version 1.0.

Kristensen, Kasper, Uffe Høgsbro Thygesen, Ken Haste Andersen, and Jan E. Beyer (2014). Estimating spatio-temporal dynamics of size-structured populations. *Canadian Journal of Fisheries and Aquatic Sciences* 71: 326–336.

Lewy, Peter, J Rasmus Nielsen, and Holger Hovgård (2004). Survey gear calibration independent of spatial fish distribution. *Canadian Journal of Fisheries and Aquatic Sciences* 61(4): 636–647.

Millar, Russell B (1992). Estimating the size-selectivity of fishing gear by conditioning on the total catch. *Journal of the American Statistical Association* 87(420): 962–968.

Miller, Timothy J and Verena Trenkel (2013). A comparison of hierarchical models for relative catch efficiency based on paired-gear data for us northwest atlantic fish stocks. *Canadian Journal of Fisheries and Aquatic Sciences* 70(9): 1306–1316.

R Core Team (2014). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Rademeyer, Rebecca A and Doug S Butterworth (2013). An update of the catchability calibration factor between the africana with the old and the new gear, with an attempt to estimate its length-dependence. p. HAKE/P1. MARAM IWS.

Strømme, T (1992). *Software for Fishery Survey Data Logging and Analysis: User's Manual*, Vol. 4. Food & Agriculture Org.

Thorson, James T and Eric J Ward (2014). Accounting for vessel effects when standardizing catch rates from cooperative surveys. *Fisheries Research* 155: 168–176.