DTU Library

# Introducing selfisher: open source software for statistical analyses of fishing gear selectivity

**Brooks, Mollie Elizabeth; Melli, Valentina; Savina, Esther Anne Charlotte Marie; Santos, Juan; Millar, Russell B.; O'Neill, Finbarr G; Veiga-Malta, Tiago; Krag, Ludvig Ahm; Feekings, Jordan P.**

*Published in:*
Canadian Journal of Fisheries and Aquatic Sciences

*Link to article, DOI:*
10.1139/cjfas-2021-0099

*Publication date:*
2022

*Document Version*
Peer reviewed version



*Citation (APA):*
Brooks, M. E., Melli, V., Savina, E. A. C. M., Santos, J., Millar, R. B., O'Neill, F. G., Veiga-Malta, T., Krag, L. A., & Feekings, J. P. (2022). Introducing selfisher: open source software for statistical analyses of fishing gear selectivity. *Canadian Journal of Fisheries and Aquatic Sciences*, *79*(8), 1189-1197. https://doi.org/10.1139/cjfas-2021-0099

# Introducing `selfisher`: open source software for statistical analyses of fishing gear selectivity

Authors and affiliations:

Mollie E Brooks[a,*], Valentina Melli[a], Esther Savina[a], Juan Santos[b], Russell Millar[c], Finbarr Gerard O'Neill[a], Tiago Veiga-Malta[a], Ludvig Ahm Krag[a], Jordan Paul Feekings[a]

[a] National Institute of Aquatic Resources, Technical University of Denmark, Lyngby, 2800 Denmark

[b] Thünen Institute of Baltic Sea Fisheries, Rostock, Germany

[c] Department of Statistics, The University of Auckland, Private Bag 92019, Auckland, New Zealand

* Corresponding author

Running Headline: R package `selfisher` for gear selectivity

# Abstract

There is a need to improve fishing methods to select for certain sizes and species while excluding others. Experiments are conducted to quantify selectivity of fishing gears and how variables such as gear design (e.g. mesh size, mesh shape), environmental parameters (e.g. light, turbidity, substrate) or biological parameters (e.g. fish condition) alter selectivity; the resulting data need to be analyzed using specialized statistical methods in many cases. Here, we present a new tool for analyzing this type of data: an R package named `selfisher`. It allows estimating multiple fixed effects (e.g. fish length, total catch weight, environmental variables) and random effects (e.g. haul). A bootstrapping procedure is also provided. We demonstrate its use via four case studies including (A) covered codend analyses of four gears, (B) a paired gear study with numerous covariates, (C) a catch comparison study of unpaired hauls of gillnets and (D) a catch comparison study of paired hauls using polynomials and splines. This software will make it easier to model selectivity, teach statistical methods, and make analyses more repeatable.

Keywords: geometric similarity, catch comparison, covered codend, gillnet, mesh size, paired gear, trawl

# Introduction

The wasteful practice of returning organisms back to the sea, often dead or dying, is commonly known as discarding and is responsible for approximately 11% of total commercial catches or 9.1 million tons (Pérez Roda et al., 2019). Consequently, substantial effort is spent on trying to reduce these unwanted catches by developing fishing gears that select for certain species and sizes of individuals while allowing others to avoid capture (Kennelly and Broadhurst, 2021). Experiments are conducted to measure the selectivity of fishing gear and statistical models are used to characterize the selectivity patterns. The selectivity of fishing gear is commonly described by a retention curve, i.e. the probability of being retained in the net, which is usually a function of individual length or size and may vary between hauls (Wileman et al., 1996). Between-haul variation may be random (due to unmeasured variables), or it may depend on observed covariates such as total catch weight (Fryer, 1991; Suuronen & Millar, 1992; Erickson et al., 1996; O'Neill & Kynoch, 1996), environmental variables that affect animal behavior including light, temperature, and bottom substrate (He, 1993; Walsh & Hickey, 1993; Ryer & Barnett, 2006, Somerton et al., 2013), or the body condition of individuals (Özbilgin et al., 2007; Ferro et al., 2008).

In some cases, catch data collected in selectivity studies could be analyzed with standard binomial generalized linear models, for which there are plenty of software options available. Binomial models are appropriate because, in many gear selectivity experiments, individuals end up in one of two compartments (e.g. codend vs cover; gear 1 vs gear 2; or test gear vs control gear; Wileman et al., 1996), i.e. there are two possible outcomes as in coin flips. However, a substantial amount of the analyses in this field are specialized and require specialized software. For example, obtaining confidence intervals on predictions typically requires accounting for extra-binomial variability (overdispersion) between and within hauls

(Millar, 1993; Millar et al., 2004). Paired gear studies (where the test gear is tested against a control one with retention probability equal to one for the given species and lengths of interest) is another example that does not conform to a typical logistic regression model because the probability model is more complex as we will show below. This paper presents newly developed open source software that is specifically designed for modelling fishing gear selectivity, something that was previously limited. The package, `selfisher`, is implemented in the R statistical computing environment which is commonly used for many modern fisheries analyses (R core team 2020). The package is written with an interface that will be familiar to many users of regression methods in R. By making the software free and openly available, we aim to improve repeatability of analyses and enable teaching these analytical methods in classrooms for the next generation of fisheries scientists.

In this paper, first, we describe the implementation of the `selfisher` R package and provide a general description of how models are estimated by the package, including dealing with subsampled catches, a common occurrence in gear selectivity studies (e.g. Larsen et al., 2018; Melli et al., 2019; Veiga-Malta et al., 2020). We also describe ways that the package can be used to address the common issue of overdispersion (Millar et al., 2004). We describe three general categories of statistical models, divided based on the mathematical probabilities underlying the estimation in our method, while omitting details about experimental designs and code to run the models. Then, we describe the bootstrapping procedure used to account for variation within and between hauls in selectivity (Millar, 1993), potentially resampling from distinct pools of hauls based on gear type or tactic as in unpaired hauls (Herrmann et al., 2017; Savina et al., 2017). Then, we briefly describe four case studies used to illustrate the package's capabilities, while providing more thorough descriptions with R code in the supplementary material. The case studies are (A) a covered codend study of haddock (*Melanogrammus aeglefinus*) with four different codends (O'Neill et al., 2016), (B) a paired gear study in a brown shrimp (*Crangon crangon*) trawl fishery where one trawl is nonselective (Santos et al., 2018), (C) a catch comparison study of unpaired hauls of gillnets avoiding an unwanted crab (*Cancer pagurus*) (Savina et al., 2017), and (D) a catch comparison study of paired hauls in a Norway lobster (*Nephrops norvegicus*) trawl fishery (Melli et al., 2018). Finally, we discuss current limitations of the `selfisher` package and potential future advancements.

## Implementation of `selfisher`

We designed the `selfisher` package to be flexible and robust for fitting and assessing a variety of gear selectivity models that can be represented with a binomial distribution. The code for `selfisher` was developed by modifying the R package `glmmTMB` (Brooks et al., 2017) because `glmmTMB` already had the capabilities needed for fitting and analyzing binomial mixed effect models. Previously, `glmmTMB` was developed by adapting the popular user interface from `lme4` (Bates et al., 2015) and increasing the model flexibility and fitting robustness by doing estimation with `TMB` (i.e. Template Model Builder, Kristensen et al., 2016). Prior to the development of `glmmTMB`, `TMB` was developed based on the algorithm of AD Model Builder (`ADMB`), which performs maximum likelihood estimation (MLE) in a fast and robust way (Fournier et al., 2012; Miller, 2013). The algorithm is fast and robust because it has information on the gradients of the likelihood surface via automatic differentiation. Additionally, `TMB` improves robustness by providing binomial and beta-binomial likelihood functions that are numerically stable even when probabilities are near zero or one. Thus, through inheritance, `selfisher` has a flexible user interface with `lme4`-style syntax and robust `TMB` code underlying the model estimation which is done using the same MLE algorithm as `ADMB`.

All models in `selfisher` involve comparing the catches from two compartments (e.g. test vs control gear, gear 1 vs gear 2, or codend vs cover), which gives rise to data that can be analysed as a binomial response, subject to the use of appropriate methods to allow for overdispersion as described below. Due to technicalities of the underlying code, in `selfisher` syntax, the binomial response must be specified as a proportion (i.e. proportion of individuals of a given length in one compartment of one haul with respect to the total) and a total (i.e. total number of individuals of a given length in either compartment in one haul), as we will demonstrate in case studies; this is in contrast to other binomial regression methods that take a two-column response variable. There are three main categories of experimental designs in which each individual has two possible outcomes, i.e. studies that produce binomial data and can be analyzed with `selfisher`. The categories are (1) a selective net inside an outer nonselective small-mesh cover net (covered codend), (2) a selective net compared to a nonselective net (paired gear), and (3) a comparison between two selective gears (catch comparison). These can all be modeled using `selfisher` as we describe in sections below, but first we describe some generalities.

# Retention models

All three categories of analyses involve estimating a retention model; covered codend and paired gears experiments allow one to estimate the absolute retention, i.e. retention out of the population encountered by the gear, while in a catch comparison experiment the estimated retention is relative to that of a baseline gear. Regardless, the mathematical formulation is general. We use $r(l)$ to refer to the retention model throughout this text, but as we describe in case studies below, it may depend on covariates other than length, $l$. See Table 1 for a description of all notations. As in binomial generalized linear models (GLMs), retention models use a link function to keep the retention probability in the range from zero to one. The most common link is the "`logit`" (i.e. logistic), but other options include, "`probit`" (i.e. normal probability ogive), "`cloglog`" (i.e. negative extreme value), "`loglog`" (i.e. extreme value/Gompertz), or "`Richards`". The software default is the logit link. To fit retention model shapes that are more diverse than the built-in link functions, it is possible to use a logit link with more complex models such as polynomials (Holst & Revill, 2009) or smooth functions (Skalski & Perez-Comas,1993; Munro & Somerton, 2001; Fryer et al., 2003; Somerton et al., 2013) as demonstrated in case studies C and D below.

## Selectivity statistics $l_{50}$ and SR

Two estimated summary statistics of interest are the length with 50% probability of retention ($l_{50}$) and the selection range (SR, i.e. the width of the range of length classes with 25% to 75% retention probabilities). Note that they only apply to models where retention probability monotonically increases with length, such as in covered codend and paired gear studies. In simple models with only length or size as a predictor of retention, then $l_{50}$ and SR can be extracted from a model using the function `L50SR()`. In more complex models, such as the covered codend and paired gear case studies below, which involve additional covariates besides length, there are multiple ways to extract $l_{50}$ and SR estimates. The covered codend case study demonstrates how to extract them algorithmically by finding the lengths that correspond to retention probabilities 0.25, 0.50, and 0.75 for each given value of covariates in the model. The paired gear example solves for $l_{50}$ and SR mathematically using a model's estimated coefficients. For either method, confidence intervals can be obtained by bootstrapping.

## Subsampled catches

Often, in cases of abundant catches, it may not be feasible to measure the length of every individual that is caught and, in those cases, only a fraction of the catch may be measured. This leads to additional statistical complexity in the analyses, but we have made `selfisher` capable of handling subsampling in any model. Here we denote the approximate fraction of individuals in compartment *i*, of haul *h*, and length class *l* that were sampled as $s_{i,h,l}$. Although subsampling doesn't always depend on *l*, we have written the package to be flexible enough to handle cases where each observed count (e.g. $n_{i,h,l}$) has a different subsampling fraction. It is sufficient to include the ratio of subsampling fractions (i.e. the subsampling ratio) in a model, rather than each compartment's fraction individually, $q_{h,l} = s_{i,h,l}/s_{j,h,l}$, assuming here that *i* is the compartment being considered as "success" in the binomial context and *j* is the alternative compartment. If raising factors were recorded in the data instead of subsampling fractions, then care should be taken when calculating the subsampling ratio to account for the fact that a raising factor is the inverse ($1/s_{i,h,l}$) of a subsampling fraction. Subsampling fractions are between zero and one, while raising factors are greater than or equal to one.

The most general way to account for subsampling in a `selfisher` model is to specify the subsampling ratio $q_{h,l}$ using the argument `qratio` in a call to the `selfisher` function. In covered codend models with logit links (Millar, 1994) or catch comparison models with logit links (Holst & Revill, 2009) an `offset` could be used equivalently, but using `qratio` will be more broadly applicable because it can be used with any type of link including in paired gear models in addition to the other types.

# Three categories of experimental designs

## Covered codend

One way of characterizing the selectivity of towed gear is to capture the individuals that escape the net using a small-mesh cover, commonly known as the covered codend method.

 



Then, the statistical model has strong information on retention because the total number of individuals encountered in each length class is directly observed. In this category, we compare the number of individuals sampled in the cover in haul *h* with length *l* ( $n_{1,h,l}$) and codend ($n_{2,h,l}$) by modeling the proportion $\frac{n_{1,h,l}}{n_{1,h,l}+n_{2,h,l}}$ as the probability of being retained and sampled in the codend divided by the probability of being retained and sampled in the codend or escaping to the cover and being sampled (Millar, 1994):

$\phi_{covered,h,l} = \frac{r(l)s_{1,h,l}}{r(l)s_{1,h,l}+(1-r(l))s_{2,h,l}}$ or more simply in terms of the subsampling ratio:

$\phi_{covered,h,l} = \frac{r(l)q_{h,l}}{r(l)q_{h,l}+(1-r(l))}$.

See Table 1 for definitions of all symbols.

## Paired gear (where one gear is nonselective)

Another way to characterize the selectivity of a fishing gear is to acquire knowledge on the population available to be caught in each haul. Paired gear studies accomplish this by deploying a control gear or codend (besides the one whose selectivity is being measured) that has full retention for the species and lengths of interest, i.e. all individuals of the given species entering that gear are retained. We assume that there is a probability, $p_h$, of entering the test gear in haul *h*, given that the individual goes into either the test (*t*) or control (*c*) gear. The proportion of individuals observed in the test gear in haul *h* with length class *l* compared to the total number of individuals observed $\frac{n_{t,h,l}}{n_{t,h,l}+n_{c,h,l}}$ is modeled as the probability of entering the test net, being retained, and being sampled, divided by the probability of entering either net, being retained and being counted:

$$\phi_{paired,h,l} = \frac{p_h s_{t,h,l} r(l)}{p_h s_{t,h,l} r(l) + (1-p_h)s_{c,h,l}}$$

It is convenient to write the model more simply in terms of the subsampling ratio:

$$\phi_{paired,h,l} = \frac{p_h q_{h,l} r(l)}{p_h q_{h,l} r(l) + 1 - p_h}.$$

In paired gear studies, the ideal relative fishing power is 50% (i.e. *p*=0.5). If this is known *a priori* then it is possible to fix *p* at 0.5 by specifying `pformula=~0` in the `selfisher` model. If another value of *p* is known *a priori* due to differences in effort, such as swept areas, then that can be specified using the `qratio` argument. For example, if the study has (haul-specific) subsampling fractions *st* and *sc* as well as (haul-specific) swept areas of *at* and *ac* for the test and control gear respectively, then one could use argument `qratio = at / ac * st / sc` together with `pformula=~0` (e.g. Somerton et al., 2013).

# Catch comparison (where both gears are selective)

Catch comparison studies compare the catches in two gears, both of which are selective. Consequently, there is no direct information about the length distribution of the population being fished and it is only possible to model the relative retention probability given the population encountered during testing (Revill & Holst, 2004). In general, the relative retention probability model can be of arbitrary shape and, for example, may not be monotone.

The response variable is the proportion of fish retained by one gear versus the other

$$\frac{n_{1,h,l}}{n_{1,h,l} + n_{2,h,l}}.$$

In a catch comparison study, the modeled proportion of the total catch retained and sampled in gear 1 versus gear 2 is the probability of entering gear 1, being retained in gear 1, and being sampled over the probability of entering, being retained, and being sampled in either gear. It is always modeled with a logit link:

$$\phi_{compare,h,l} = \frac{p_h s_{1,h,l} r_1(l)}{p_h s_{1,h,l} r_1(l) + (1 - p_h) s_{2,h,l} r_2(l)}$$

where $r_1(l)$ and $r_2(l)$ are the absolute retention probabilities of the two gears, but it is not possible to estimate them separately. Holst and Revill (2009) showed that the expected proportion can be approximated with a polynomial with a logit link and an offset to account for any subsampling. They showed that the relative fishing pressure $p_h$ can be absorbed by the intercept and that it may vary randomly between hauls and accounted for by including a random intercept of haul as in a mixed effects model. In general, the relative retention model in a catch comparison analysis can be formulated as either a polynomial (Holst & Revill, 2009) or a spline (Miller, 2013); see the catch comparison case studies in Supplements C and D for a demonstration of how this can be done with `selfisher`.

# Handling extra variability

## Overdispersion

Overdispersion is the presence of variability in the proportions that is in excess of the variability specified under a binomial model as is typical for selectivity models. Overdispersion can arise due to between-haul variability whereby the retention model varies from haul to haul. To a lesser extent it can also arise due to within-haul variability due to schooling behavior that violates the assumption that all fish behave independently. The accepted approaches for including this variability are the use of bootstrapping (Millar, 1993) and the use of mixed effects models (Fryer et al. 2003, Millar et al., 2004). Both of these methods are available in `selfisher`. All of the case studies demonstrate how to bootstrap as described below. See the covered codend and paired gear case studies (Supplements A and B) for examples of using mixed effects in `selfisher`. In this text, we do not get into the details of random effects because it is a large topic; however, note that in `selfisher`, they are implemented in the same way as `glmmTMB` (Brooks et al., 2017) using the Laplace approximation, which is a standard method commonly used in modern mixed modelling. Laplace approximation should work well provided there is a sufficient sample size to inform the estimates for each haul (Ogden 2021), but the sample sizes needed are an open area of research.

## Quantify uncertainty by bootstrapping

For any statistical model, it is important to compare predictions from the model - and the uncertainty around those predictions - with observed data to ensure that the model reasonably represents the data. A bootstrapping procedure was developed by Millar (1993) to account for variation between and within hauls and calculate appropriately wide confidence intervals. The bootstrapping method can account for overdispersion in data due to variability among hauls (as described above in the *Overdispersion* section); because of this, it is not necessary to include a random effect of haul in models to be bootstrapped. The bootstrapping method is also sometimes referred to as a “double bootstrap” in fishing gear selectivity literature, but this term has another meaning in statistics (e.g. Kuk, 1989). This method first resamples the same number of hauls from the observed set, with replacement. Then for each resampled haul, the method resamples observed fish within the haul. Then it refits the model to the resampled data set and the refit model is used to produce values such as predictions or parameter estimates. It is typically repeated one thousand times or more. This bootstrapping method is implemented in `selfisher` in a way that maintains all variables associated with each observed data point, not just length class, e.g. sampling fractions or total catch. This is facilitated by specifying the `haul` argument in the `selfisher` model fitting function; followed by a call to the `bootSel` function.

It is also possible to do resampling from pools of hauls, so that every bootstrapped dataset has the same number of hauls in each pool as in the original data (Herrmann et al., 2017). That is, if the original dataset has $H_A$ hauls of type A and $H_B$ hauls of type B, then it is possible to bootstrap in a way such that simulated data sets have $H_A$ hauls of type A and $H_B$ hauls of type B. This is done using the `pool` argument to the `selfisher` function. The inner part of the bootstrapping method (resampling fish within hauls) is the same as in the regular bootstrapping method. This is useful when hauls of the two gears or tactics being compared are unpaired. See the gillnet case study in Supplement C for an example of specifying pools of hauls.

# Bootstrapping and mixed modeling

Mixed modelling is a formal method that takes into account possible sources of variability in the data such as variation between hauls, enabling sound hypothesis testing and model selection. However, fitting mixed models can be computationally intensive. Moreover, the researcher is typically interested in obtaining overall selectivity predictions, rather than at the haul level, because these are relevant to the selectivity applied to the fishery. These could be obtained by taking a pseudo-Bayesian approach, that is, by simulating over all sources of variability in the fitted mixed model, including random effects and parameter uncertainty. However, this would require complex bespoke programming. A more straightforward alternative would be to fit the best candidate model, but leaving out all random effects, and to perform a bootstrap of this fixed-effects model by resampling the data in a way that emulates all additional sources of variability (Millar, 2011). Bootstraps can then be used to obtain confidence intervals for estimated quantities such as predicted retention curves or $l_{50}$ and SR. See case studies A and B for examples.

# Case studies

## A. Covered codend analyses of four codends catching haddock

This case study uses the haddock data from an experiment that employed the covered codend method to investigate the selective performance of four codends (O'Neill et al., 2016). The codends were made from netting materials with different twine bending stiffnesses and mesh sizes.

We begin the case study by looking at just one gear type to demonstrate different link functions that can be used and to show how to account for subsampling, which in this example varies with length. The default link is the logistic, but we also consider the probit and Richards curves, and a spline. The spline is fit using the `bs()` function from the `splines` package to construct a basis for a polynomial spline with 3 degrees of freedom so that the model has an intercept plus three nodes located based on quantiles of fish length. Penalized splines are not yet implemented so the degrees of freedom are fixed. Having chosen a model, we bootstrap to estimate 95% confidence interval for the proportion retained by the codend.

We then analyse all four gear types together and investigate the influence of length, mesh size, bending stiffness and catch size. We assume the principle of geometric similarity (as used by Tokai et al. (1996) to investigate grid selection). Geometric similarity is the assumption that $l_{50}$ and SR are proportional to mesh size and this corresponds to using a `I(length/meshsize)` term in the `selfisher` formula interface for the retention model (Baranov, 1948). We explore a number of models and choose the best fit using Bayesian information criterion (BIC). When choosing the best model, we include a random effect of haul to account for between-haul variation. As in the original publication, we show that selection is dependent on all three parameters. Before bootstrapping, we drop the random effect of haul from the best model because the bootstrapping method accounts for between-haul variability and the random effect would slow it down considerably. We bootstrap to estimate confidence intervals for the proportion retained by each gear and numerically solve for $l_{50}$ and SR dependent on covariates. See Supplement A for details and code.

Results were similar to the original manuscript, which also found a dependence of haddock $l_{50}$ on mesh size, twine bending stiffness, and total codend catch weight (Fig. 1) but that SR is constant (O'Neill et al., 2016). We should not, however, expect the results to be identical as there are many differences between the analyses. Originally, it was assumed that the slope and intercept of the logistic link functions vary randomly from haul to haul and that $l_{50}$ and log (SR) mean selection curve were linearly dependent on the explanatory variables using separate models. Here, we have used a single model which assumed geometric similarity for mesh size and that overall retention is related to the explanatory variables.

# B. Paired gear analyses of codend selectivity dependent on mesh size

This case study draws on a subset of data from the German research project CRANNET (Santos et al., 2018). The experimental method consisted of fishing with two identical beam trawls, simultaneously and in parallel on the same shrimp population. One of the trawls had a small-mesh (11 mm) control codend with very limited selectivity, assumed to be nonselective on the range of shrimp lengths available to the trawl. The second trawl had a test codend. The subset of data analyzed here consists of catch data from 87 hauls, during which 13 diamond-mesh codends varying in mesh size ranging from 19.1 mm to 36.3 mm were tested. The goal was to model $l_{50}$ and SR as a function of mesh size, and to quantify any effect of two additional haul covariates, sea state and catch weight.

The statistical modeling of selectivity begins with a mixed model to formally assess the effect of mesh size, sea state and catch weight, while controlling for additional random variation among hauls using a random intercept of haul. *A priori*, we assumed geometric similarity for the mesh size (ms) covariate as in case study A above, by including the term `I(length/ms)` in models. Sea state and test catch weight each had one value per haul in the data set. So that, for a given paired haul of the test and control nets, the same value of sea state was used as a predictor of retention probability of the test net for all length classes. The default model was compared to several others and found to be preferred (using BIC), and neither sea state nor catch weight had a significant effect.

Having chosen geometric similarity (with respect to mesh size) as the preferred model, this model was refitted without random effects so as to estimate size selectivity at the population level. Bootstrapping was used to obtain appropriate confidence intervals on $l_{50}$ and SR for any given mesh size. In addition, this case study demonstrates the use of `psplit=TRUE` (unequal fishing power of the paired codends), the use of sampling ratios, use of the `inits()` function to specify good starting values (because without it some models converged to local minima that did not make any sense), and shows how $l_{50}$ and SR can be obtained directly from the model fitted by `selfisher`. See Supplement B for details and code.

Estimates of $l_{50}$ and SR dependent on mesh size were very close to those from the original analysis (Fig. 2). The `selfisher` model predicted higher values of SR, with larger uncertainty than Santos et al. (2018). This is a plausible result considering the different model structures applied, and the large variation of the by-haul estimates.

## C. Catch comparison analyses of unpaired hauls of gillnets avoiding an unwanted crab

This example deals with data from an experiment originally published by Savina et al. (2017). Two soak tactics (12h at day and 12h at night) were compared in the Danish gillnet plaice fishery to estimate the role that the choice of a soak tactic plays in the catch efficiency of both target and unwanted species. This is a subset of the original dataset (one species, two soak tactics) where we are looking at the unwanted invertebrate edible crab (*Cancer pagurus*).

We use the method developed by Herrmann et al. (2017) which was developed for assessing the effect of changing the gear design on the relative length-dependent catch efficiency. This example is representative of experimental fishing where the catch data obtained for two gears or tactics were not collected in pairs, and can allow for a different number of deployments.

This case study is a typical model for catch comparison of multiple haul data without subsampling using a spline. To get confidence intervals on predictions, we bootstrapped from two pools according to tactic (night vs day) using the argument `pool=tactic` to the `bootSel` function. See Supplement C for details and code.

The estimated catch comparison curves and relative confidence intervals were very similar to the original analysis (Fig. 3). The catch comparison curves properly reflected the trend in the experimental points. The experimental rates were subject to increasing binomial noise outside the length classes representing the main bulk of the catches. The results for edible crab showed significantly higher catches at night than at day with no strong indication of a length dependency.

# D. Catch comparison analyses of paired hauls of *Nephrops* twin-rigged trawls

The example is based on the data from Melli et al. (2018). An anterior gear modification, namely the counter-herding device FLEXSELECT, was tested in a twin-rig configuration, where two identical trawls were towed in parallel. One trawl was equipped with FLEXSELECT, referred to as the test trawl, while the other worked as baseline. The aim of the study was to determine if FLEXSELECT could reduce the fish bycatch in a *Nephrops*-directed fishery. The data used in the example are from haddock, which was found to be strongly affected by the counter-herding device.

Following the steps of the published paper, we conducted a catch comparison analysis, modeling the relative retention as a 4th-order polynomial. In addition, we used a spline with 4 degrees of freedom using the `splines` package and performed model selection to determine if it fitted the data better. Considering that part of the hauls were conducted in day-time and part in night-time, "time" was included in the model as an explanatory variable to determine if the length-based efficiency of FLEXSELECT presented diel differences. We predicted both catch comparison rates and catch ratio with bootstrapped confidence intervals using the `predict` and `bootSel` functions from `selfisher`. See Supplement D for details and code.

Model selection using AIC indicated that a spline with five degrees of freedom was only slightly more parsimonious than a fourth-order polynomial (0.4 deltaAIC). Because the polynomial has fewer parameters with nearly equal AIC (df=12 compared to df=10) we plot results from the polynomial here to show that results from `selfisher` were nearly equal to the analyses from the original publication which used `SELNET` (Fig. 4.).

# Discussion

We have introduced an open source R package for estimating fishing gear selectivity of both towed and passive gear, making it easier for anyone to analyze fishing gear selectivity data without writing extensive amounts of code. We have demonstrated its broad applicability in four case studies spanning a range of experimental designs. The case studies have shown that results from `selfisher` are comparable to previously published results and that `selfisher` is more flexible than some methods (e.g. a single model to quantify the effect of changing mesh size on $l_{50}$ and SR). Some of the features of `selfisher` that were demonstrated in the case studies are summarized in Table 2. The case studies aim to demonstrate best practices based on current knowledge. However, this is an active area of research and with a new powerful model fitting tool, best practices may change. Even with (or especially with) a powerful tool, analyses require careful thought and checking of results. For example, in complicated models such as paired gear models which contain two submodels (retention and relative fishing power), it may be necessary to be cautious about identifiability of parameters and local optima encountered during maximum likelihood estimation, but better starting values help avoid those issues as demonstrated in Supplement B (Bolker et al., 2013). If some parts of a model are not identifiable, then the summary output will contain "NA"s. The `TMBhelper` package (Thorson, 2020) contains a function to identify which parts of a `selfisher` model (e.g. `mod`) may be problematic by typing `check_estimability(mod$obj)`.

We have several advancements for the package either planned for the future or already underway. Thus, we recommend installing the latest version and keeping up to date with advancements by following the guidance in Appendix 1. We plan to add functions to calculate discard ratio indices and indicator functions (e.g. Wienbeck et al., 2014; Santos et al., 2016; Veiga-Malta et al., 2019). We may add a function to facilitate model averaging, although it is already possible to piece this functionality together with the existing features. We have not tried to fit structured non-monotonic curves (e.g. bell-shaped curves of gillnet absolute selectivity based on geometric similarity, Baranov 1948) with `selfisher`, but we will explore this possibility in the future. We will investigate how to choose starting values of parameters in models that have Richards link, to increase robustness. We plan to implement a general method to extract $l_{50}$ and SR from complex models as demonstrated in case studies A and B. To handle overdispersion more elegantly, we plan to add the option of having a beta-binomial distribution for the response (Miller, 2013). We are already in the process of developing a Shiny app, which will facilitate simple standard analyses without the need for writing code; this will help to bridge the gap for scientists or managers with extensive experience in gear development but little experience with R. As an open source package, code developers are encouraged to contribute improvements through GitHub such as those listed here.

Having access to a free and open source software should benefit this field of research in several ways. It allows researchers to share code and thereby foster a community for discussion and repeatability. The free nature of the software will enable researchers and managers with limited budgets - such as those in developing countries - to perform analyses themselves. It gives statistical methods of retention modelling a way into classrooms containing the next generation of fisheries scientists who are already learning modern regression methods as part of a general scientific curriculum.

# Acknowledgements

We thank William T. Stockhausen for the trick of fixing $p_h$ to any number in a paired gear analysis. We thank the helpful crews of both research and commercial vessels, as well as the fisheries technicians of the research institutes providing essential support during the fishing trips.

# Competing interests statement

Competing interests: The authors declare there are no competing interests.

# Contributors' statement

Conceptualization: MB, LK, JF; Methodology, MB, RM; Software, MB, TVM; Validation: MB, VM, TVM, ES, JS, BON; Data Curation: MB, VM, ES, JS, BON; Writing - Original Draft: MB, VM, TVM, ES, RM, JS, BON; Writing - Review & Editing: MB, VM, TVM, ES, RM, JS, BON, JF Funding acquisition: LK, JF

# Funding Statement

We are grateful for funding from the European Maritime and Fisheries Fund (EMFF) and the Ministry of Environment and Food of Denmark (Miljø- og Fødevareministeriet) as part of the projects FastTrack – Sustainable, cost effective and responsive gear solutions under the landing obligation (33112-P-15-013), FastTrack II – Sustainable, cost effective and responsive gear solutions under the landing obligation (33112-P-18-051), and Flexselect - Innovative and Flexible solutions for the Nephrops Fisheries (33113-I-16-068).

# Data Availability Statement

Data used in the case studies is available as part of the `selfisher` package on GitHub. See the *Installation* instructions in Appendix 1 as well as the online supplementary material.

# References

Baranov, F. I. (1948). Theory and assessment of fishing gear. In Theory of fishing with gillnets. Chap. 7. Pishchepromizdat, Moscow. (Translation from Russian by Ontario Dept of Lands, Maple, Ont., 45 pp.)

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Bolker, B. M., Gardner, B., Maunder, M., Berg, C.W., Brooks, M., Comita, L., Crone, E., Cubaynes, S., Davies, T., de Valpine, P., Ford, J., Gimenez, O., Kéry, M., Kim, E.J., Lennert-Cody, C., Magnusson, A., Martell, S., Nash, J., Nielsen, A., Regetz, J., Skaug, H. and Zipkin, E. (2013), Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. *Methods Ecol Evol*, 4: 501-512. https://doi.org/10.1111/2041-210X.12044

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H.J., Machler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal*, *9*(2), 378-400. https://doi.org/10.3929/ethz-b-000240890

Erickson, D. L., Perez-Comas, J. A., Pikitch, E. K., & Wallace, J. R. (1996). Effects of catch size and codend type on the escapement of walleye pollock (*Theragra chalcogramma*) from pelagic trawls. *Fisheries Research*, *28*(2), 179-196. https://doi.org/10.1016/0165-7836(96)00497-3

Ferro, R. S. T., Özbilgin, H., & Breen, M. (2008). The potential for optimizing yield from a haddock trawl fishery using seasonal changes in selectivity, population structure and fish condition. *Fisheries research*, *94*(2), 151-159. https://doi.org/10.1016/j.fishres.2008.08.018

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., & Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, *27*(2), 233-249. https://doi.org/10.1080/10556788.2011.597854

Fryer, R. J. (1991). A model of between-haul variation in selectivity. *ICES Journal of Marine Science*, *48*(3), 281-290. https://doi.org/10.1093/icesjms/48.3.281

Fryer, R. J., Zuur, A. F., & Graham, N. (2003). Using mixed models to combine smooth size-selection and catch-comparison curves over hauls. *Canadian Journal of Fisheries and Aquatic Sciences*, *60*(4), 448-459. https://doi.org/10.1139/f03-029

He, P. (1993). Swimming speeds of marine fish in relation to fishing gears. *ICES Mar. Sci. Symp*, *196*, 183-189.

Herrmann, B., Sistiaga, M., Rindahl, L., & Tatone, I. (2017). Estimation of the effect of gear design changes on catch efficiency: methodology and a case study for a Spanish longline fishery targeting hake (*Merluccius merluccius*). *Fisheries Research*, *185*, 153-160. https://doi.org/10.1016/j.fishres.2016.09.013

Holst, R., & Revill, A. (2009). A simple statistical method for catch comparison studies. *Fisheries Research*, *95*(2-3), 254-259. https://doi.org/10.1016/j.fishres.2008.09.027

Kenelly, S.J., & Broadhurst, M.K. 2021. A review of bycatch reduction in demersal fish trawls. Reviews in Fish Biology and Fisheries, https://doi.org/10.1007/s11160-021-09644-0

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., Bell, B.M. (2016). TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software*, 70(5), 1-21. http://dx.doi.org/10.18637/jss.v070.i05

Kuk, A. Y. (1989). Double bootstrap estimation of variance under systematic sampling with probability proportional to size. *Journal of Statistical Computation and Simulation*, *31*(2), 73-82.

Larsen, R. B., Herrmann, B., Sistiaga, M., Brinkhof, J., & Grimaldo, E. (2018). Bycatch reduction in the Norwegian deep-water shrimp (*Pandalus borealis*) fishery with a double grid selection system. *Fisheries Research, 208*, 267-273. https://doi.org/10.1016/j.fishres.2018.08.007

Melli, V., Karlsen, J. D., Feekings, J. P., Herrmann, B., & Krag, L. A. (2018). FLEXSELECT: counter-herding device to reduce bycatch in crustacean trawl fisheries. *Canadian Journal of Fisheries and Aquatic Sciences*, *75*(6), 850-860. https://doi.org/10.1139/cjfas-2017-0226

Melli, V., Krag, L. A., Herrmann, B., & Karlsen, J. D. (2019). Can active behaviour stimulators improve fish separation from Nephrops (*Nephrops norvegicus*) in a horizontally divided trawl codend?. *Fisheries Research*, *211*, 282-290. https://doi.org/10.1016/j.fishres.2018.11.027

Millar, R. B. (1993). Incorporation of between-haul variation using bootstrapping and nonparametric estimation of selection curves. *Fishery Bulletin*, *91*, 564-572.

Millar, R. B. (1994). Sampling from trawl gears used in sized selectivity experiments. *ICES Journal of Marine Science*, *51*(3), 293-298. https://doi.org/10.1006/jmsc.1994.1030

Millar, R. B. (2011). Maximum likelihood estimation and inference: with examples in R, SAS and ADMB. Wiley, London. 357 p. ISBN: 978-0-470-09482-2

Millar, R. B., & Fryer, R. J. (1999). Estimating the size-selection curves of towed gears, traps, nets and hooks. *Reviews in Fish Biology and Fisheries*, *9*(1), 89-116. https://doi.org/10.1023/A:1008838220001

Millar, R. B., Broadhurst, M. K., & Macbeth, W. G. (2004). Modelling between-haul variability in the size selectivity of trawls. *Fisheries Research, 67*(2), 171-181. https://doi.org/10.1016/j.fishres.2003.09.040

Miller, T. (2013). A comparison of hierarchical models for relative catch efficiency based on paired-gear data for US Northwest Atlantic fish stocks. *Can. J. Fish. Aquat. Sci. 70*(9), 1306–1316. https://doi.org/10.1139/cjfas-2013-0136

Munro, P. T., & Somerton, D. A. (2001). Maximum likelihood and non-parametric methods for estimating trawl footrope selectivity. *ICES Journal of Marine Science*, *58*(1), 220-229. https://doi.org/10.1006/jmsc.2000.1004

O’Neill F.G., Kynoch R.J., Blackadder L., Fryer, R.J., Eryaşar A.R., Notti, E., & Sala A. (2016). The influence of twine tenacity, thickness and bending stiffness on codend selectivity. *Fisheries Research, 176*, 94–99. https://doi.org/10.1016/j.fishres.2015.12.012

O'Neill, F. G., & Kynoch, R. J. (1996). The effect of cover mesh size and cod-end catch size on cod-end selectivity. *Fisheries Research*, *28*(3), 291-303. https://doi.org/10.1016/0165-7836(96)00501-2

Ogden, H. (2021). On the error in Laplace approximations of high-dimensional integrals. Stat, 10( 1), e380. https://doi.org/10.1002/sta4.380

Özbilgin, H., Tosunoğlu, Z., Tokaç, A., & Metin, G. (2007). Seasonal variation in the trawl codend selectivity of picarel (*Spicare smaris*). *ICES Journal of Marine Science*, *64*(8), 1569-1572. https://doi.org/10.1093/icesjms/fsm115

Pérez Roda, M.A., Gilman, E., Huntington, T., Kennelly, S.J., Suuronen, P., Chaloupka, M., Medley, P. (2019) A third assessment of global marine fisheries discards. FAO Fisheries and Aquaculture Technical Paper No. 633, FAO, Rome, 78pp

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Revill, A.S., & Holst, R. (2004) Reducing discards of North Sea brown shrimp (*C. crangon*) by trawl modification. *Fisheries Research,* 68, 113-122. https://doi.org/10.1016/j.fishres.2004.02.001

Ryer, C. H., & Barnett, L. A. (2006). Influence of illumination and temperature upon flatfish reactivity and herding behavior: potential implications for trawl capture efficiency. *Fisheries Research*, *81*(2-3), 242-250. https://doi.org/10.1016/j.fishres.2006.07.001

Santos, J., Herrmann, B., Mieske, B., Stepputtis, D., Krumme, U., & Nilsson, H. (2016). Reducing flatfish bycatch in roundfish fisheries. *Fisheries Research*, *184*, 64-73. https://doi.org/10.1016/j.fishres.2015.08.025

Santos, J., Herrmann, B., Stepputtis, D., Günther, C., Limmer, B., Mieske, B., and Schultz, S. et al. (2018). Predictive framework for codend size selection of brown shrimp (*Crangon crangon*) in the North Sea beam-trawl fishery. *PLoS ONE* 13(7): e0200464. https://doi.org/10.1371/journal.pone.0200464

Savina, E., Krag, L.A., Frandsen, R.P., & Madsen, N. (2017). Effect of fisher's soak tactic on catch pattern in the Danish gillnet plaice fishery. *Fisheries Research, 196*, 56-65. https://doi.org/10.1016/j.fishres.2017.08.009

Skalski, J. R., & Perez-Comas, J. (1993). Nonparametric maximum likelihood estimation of mesh size selectivity. *Fisheries research*, *18*(3-4), 321-334. https://doi.org/10.1016/0165-7836(93)90160-9

Somerton, David A., Weinberg, Kenneth L., Goodman, Scott E., & Chen, Yong. (2013) Catchability of snow crab (*Chionoecetes opilio*) by the eastern Bering Sea bottom trawl survey estimated using a catch comparison experiment. *Canadian Journal of Fisheries and Aquatic Sciences, 70*(12), 1699-1708. https://doi.org/10.1139/cjfas-2013-0100

Suuronen, P., & Millar, R. B. (1992). Size selectivity of diamond and square mesh codends in pelagic herring trawls: only small herring will notice the difference. *Canadian Journal of Fisheries and Aquatic Sciences*, *49*(10), 2104-2117. https://doi.org/10.1139/f92-234

Thorson, J. (2020). TMBhelper: Package for basic helper functions that are not worth putting in a specialized contributed package. R package version 1.3.0.

Tokai, T., Omoto, S., Sato, R., & Matuda, K. (1996). A method of determining selectivity curve of separator grid. *Fisheries Research*, *27*(1-3), 51-60. https://doi.org/10.1016/0165-7836(95)00471-8

Veiga-Malta, T., Feekings, J. P., Frandsen, R. P., Herrmann, B., & Krag, L. A. (2020). Testing a size sorting grid in the brown shrimp (*Crangon Crangon* Linnaeus, 1758) beam trawl fishery. *Fisheries Research, 231*, 105716. https://doi.org/10.1016/j.fishres.2020.105716

Veiga-Malta, T., Feekings, J., Herrmann, B., & Krag, L. A. (2019). Industry-led fishing gear development: Can it facilitate the process?. *Ocean & Coastal Management*, *177*, 148-155. https://doi.org/10.1016/j.ocecoaman.2019.05.009

Walsh, S. J., & Hickey, W. M. (1993). Behavioural reactions of demersal fish to bottom trawls at various light conditions. *ICES Mar. Sci. Symp*,196, 68-76.

Wienbeck, H., Herrmann, B., Feekings, J. P., Stepputtis, D., & Moderhak, W. (2014). A comparative analysis of legislated and modified Baltic Sea trawl codends for simultaneously improving the size selection of cod (*Gadus morhua*) and plaice (*Pleuronectes platessa*). *Fisheries Research*, *150*, 28-37. https://doi.org/10.1016/j.fishres.2013.10.007

Wileman, D., Ferro, R. S. T., Fonteyne, R., & Millar, R. B. (1996). Manual of methods of measuring the selectivity of towed fishing gears. *ICES Cooperative Research Report*, *215*.

# Figures

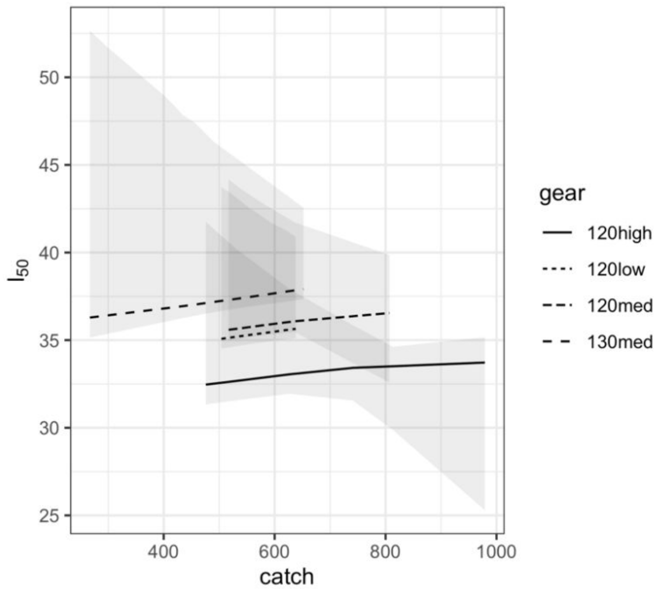


Figure 1. Estimated $l_{50}$ as a continuous function of catch weight (kg) for four types of gear using a single `selfisher` model. The lines represent the median prediction from 1000 bootstraps of the `selfisher` model, while the bands show the 2.5 and 97.5 percentiles.

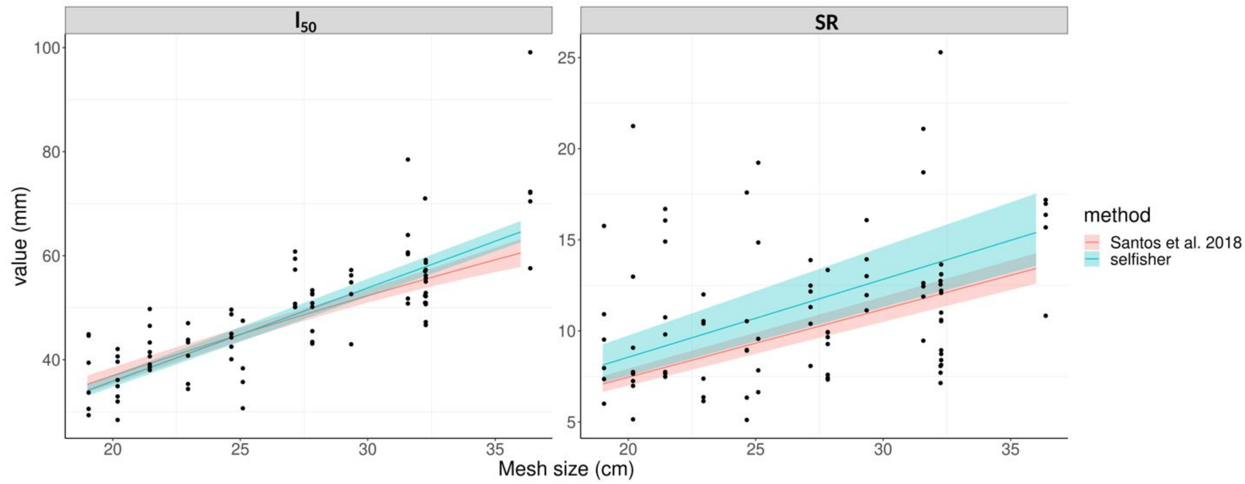


Figure 2. Estimated $l_{50}$ and SR dependent on mesh size. The black points in the figure represent values of $l_{50}$ and SR estimated at the haul level in the original study (Santos et al. 2018) and the red line and ribbon represent the analysis which combined those haul-specific estimates. The green line represents the median prediction from 1000 bootstraps of the `selfisher` model, while the band shows the 2.5 and 97.5 percentiles.

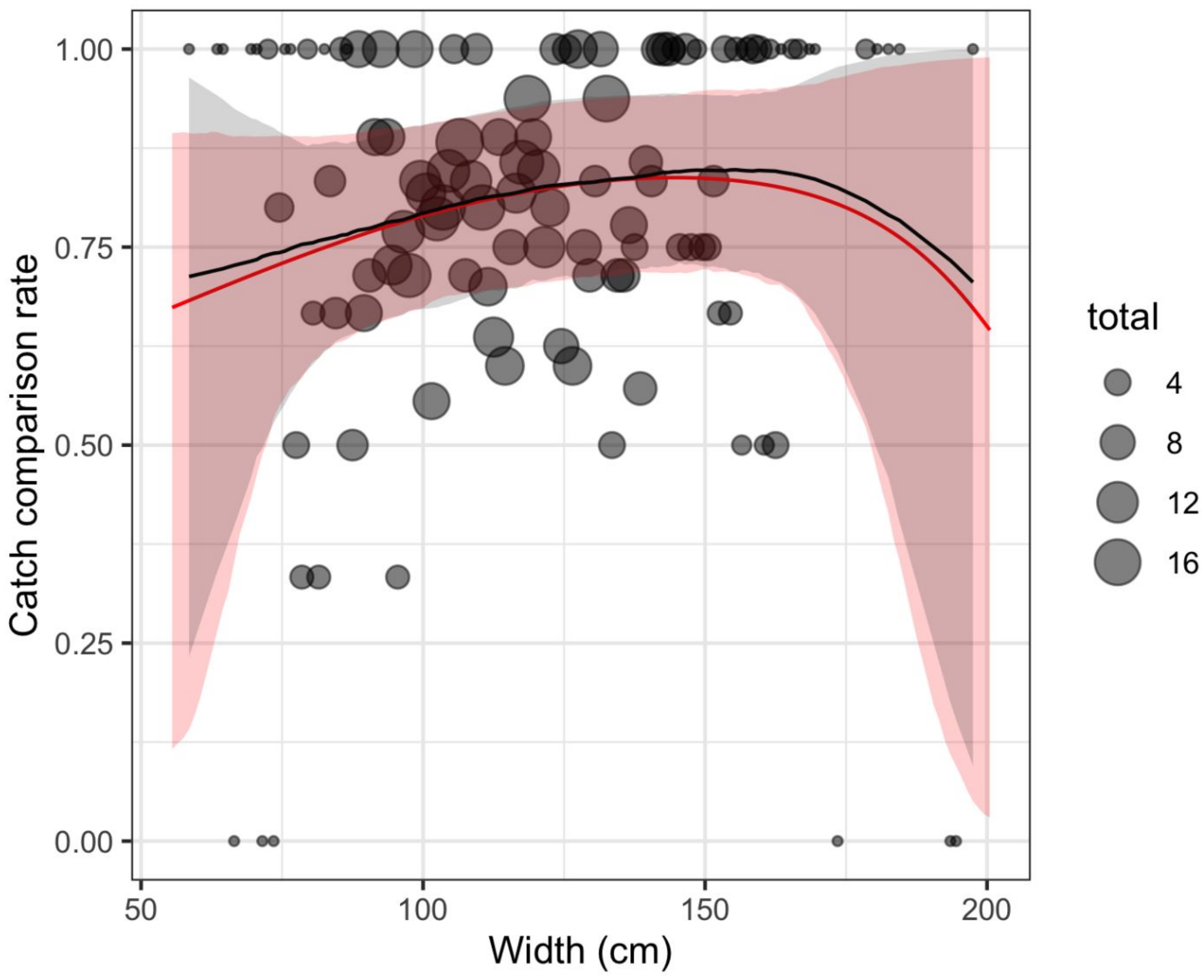


Figure 3. Catch comparison rate of night vs day fishing tactics for an unwanted crab in a gillnet fishery. The black line and grey ribbon represent the median, 2.5, and 97.5 percentiles from 1000 bootstraps from a `selfisher` model using a spline with three degrees of freedom. The red line and ribbon are predictions from `SELNET` using model averaging. Circles are total observations from the raw data, the size of which represents the total number of crabs caught with either tactic.

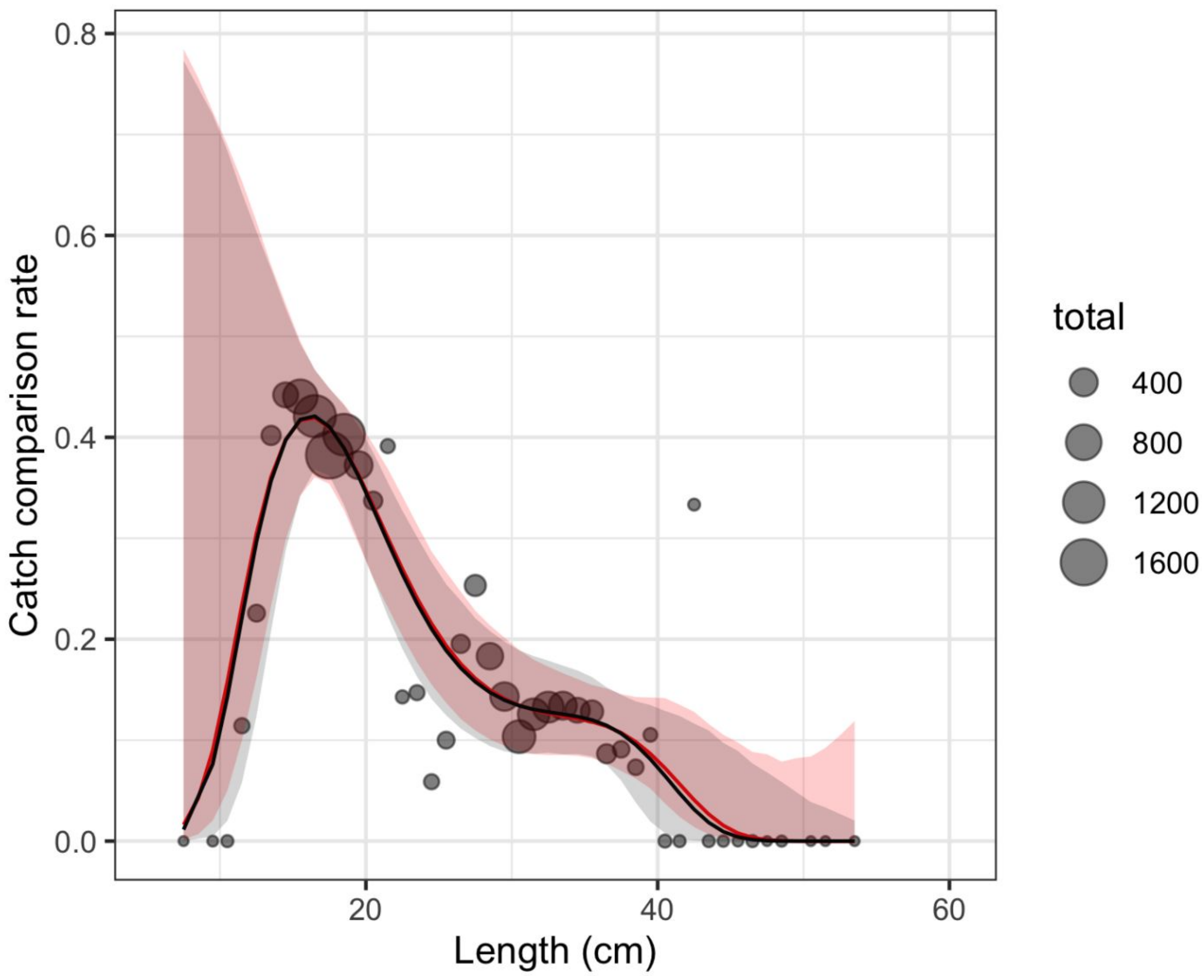


Figure 4. Catch comparison rates. Lines and ribbons are estimates and confidence intervals from `SELNET` in red using model averaging and from `selfisher` in black using a 5th order polynomial. Circles are observations from the raw data, aggregated over hauls, the size of which represents the number of haddock counted.

Table 1. Symbols used in the text.

| **Symbol** | **Meaning** |
|---|---|
| $n_{i,h,l}$ | Number of fish sampled in compartment *i*, haul *h*, in length class *l* |
| $s_{i,h,l}$ | Proportion of fish sampled in compartment *i*, haul *h*, in length class *l* |
| $q_{h,l} = s_{i,h,l}/s_{j,h,l}$ | Subsampling ratio for compartment *i* vs *j* in haul *h*, in length class *l* |
| $r(l)$ | Retention probability as a function of length *l* |
| $p_h$ | Relative fishing power of test vs control gear in haul *h* |
| $\phi_{paired,h,l}$ | Probability model in a paired gear model |
| $\phi_{compare,h,l}$ | Probability model in a catch comparison model |
| $\phi_{covered,h,l}$ | Probability model in a covered codend model |

Table 2. A list of `selfisher` features demonstrated in the four case studies with code in Supplements A through D.

| **Feature** | **Demonstrated in case studies** |
|---|---|
| Random effects | A, B |
| Fixed effects other than length or size | A, B |
| Splines | C, D |
| Residual plot | A |
| Specifying initial values | B |
| Model selection via information criteria | A, B, C, D |
| Link functions other than default logit | A |