Fisheries New Zealand

Tini a Tangaroa

An alternate-tow net A vs. net B comparison

New Zealand Fisheries Assessment Report 2021/52

M. S. Chambers
D. A. J. Middleton
D. Moran
G. Janssen

ISSN 1179-5352 (online)
ISBN 978-1-99-100996-8 (online)

September 2021

New Zealand Government

Requests for further copies should be directed to:

Publications Logistics Officer
Ministry for Primary Industries
PO Box 2526
WELLINGTON 6140

Email: brand@mpi.govt.nz
Telephone: 0800 00 83 33
Facsimile: 04-894 0300

This publication is also available on the Ministry for Primary Industries websites at:
http://www.mpi.govt.nz/news-and-resources/publications
http://fs.fish.govt.nz go to Document library/Research reports

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Chambers, M.S.[1]; Middleton, D.A.J.[2]; Moran, D.[3]; Janssen, G.[3] (2021). An alternate-tow net A vs. net B comparison.

***New Zealand Fisheries Assessment Report 2021/52.* 43 p.**

Experimental fishing was undertaken off the south Canterbury coast in January 2019 to compare the performance of an MHS1480 specification Modular Harvesting System (MHS) with a conventional 4 inch mesh trawl net typically used in the South Island inshore trawl fishery. The purpose of the study was to test an alternate-tow approach to assess the performance of an MHS trawl relative to a conventional trawl net.

A key feature of the trial design was to ensure, as far as possible, that each tow made with one gear type was matched with a corresponding tow made with the other gear type under similar circumstances. The *alternate-tow* design generates data that can be analysed using efficient *paired-comparison* type methodologies.

The criteria for assessing the performance of new trawl technologies require consideration of a number of factors including catch rates and catch composition. Catch rates are relevant to the benthic impacts associated with a given catch. Higher catch rates attain a given yield with less trawling effort and consequently lower benthic impacts. Catch rates of barracouta achieved by the two gear types were compared using a Bayesian Generalised Linear Mixed Model (GLMM) and a non-parametric bootstrap approach.

Methods for comparing the length distributions of red gurnard, tarakihi and sea perch retained by the two nets were also explored. Splines fitted to bootstrapped proportions of catch at length on MHS tows provide a measure of MHS selectivity relative to the conventional net. Potential differences between the gear types in the proportions of fish retained below a specific cut-off length, such as a minimum legal size (MLS), were investigated using a parametric test and non-parametric bootstrapping.

Despite a modest sample size of 11 tow-pairs, data collected during the trial were sufficient to demonstrate some clear differences between the two gear types:

- the barracouta catch rate of the conventional 4 inch mesh net is significantly higher than the MHS net;
- the MHS net retains fewer tarakihi less than 20 cm; and
- given the fish available to the trial, MHS catches included a lower proportion of tarakihi below the 25 cm minimum legal size.

Red gurnard retained by the two gear types on the trial had very similar size distributions. However, small gurnard were not caught by either gear so the trial was not informative about expected relative gurnard retention in areas where small gurnard are available. Sampling of sea perch lengths was insufficient to draw any conclusions about relative length-specific sea perch retention.

The analyses presented indicate that the alternate-tow design tested on the trial provides a sound basis for comparing new trawl designs with current gear. We consider that:

[1]Braw Research
[2]Pisces Research
[3]Plant & Food Research

1. the GLMM and bootstrapping approaches are both suitable for comparing catch rates. The GLMM approach may be slightly more flexible when comparing catch rates from tows targeting different species;
2. for comparisons of relative size-based retention of the two gear types, relative selectivity is more informative and more generally applicable than the approaches that compare proportions below a specific length;
3. for species where a minimum legal size is specified, it is nevertheless helpful to have an indication of the proportion of catch above or below the MLS. The non-parametric bootstrap for proportion below the MLS is recommended for this purpose.

The additional data collected on the trial (over and above the statutory data) on fish lengths and total catch of all species was important for undertaking these analyses. Although 11 tow-pairs was sufficient to draw some clear conclusions regarding catch rate and size retention of specific species from the trial, the sample sizes required in future comparisons will depend on the magnitude of the differences between gears and the level of certainty required by decision makers. The Statistics, Assessments and Methods Working Group has suggested that an allowance for up to 30 tow-pairs is considered in planning future trials, with real-time monitoring and analyses of data to understand what, if any, differences are emerging between gears.

The analysis methods recommended above can be adapted to consider data from multiple trials, including the use of multiple vessels to increase sample sizes, if required.

1. INTRODUCTION

The Modular Harvest System (MHS), developed by Precision Seafood Harvesting (PSH) as an alternative to conventional trawl gear, has recently been approved for use in specific New Zealand fisheries:

- from 24 May 2018, an MHS has been approved for targeting hoki, hake and ling in the HOK 1, HAK 1, HAK 4, HAK 7, LIN 3, LIN 4, LIN 5, LIN 6 and LIN 7 Quota Management Areas when midwater or bottom trawling (Fisheries New Zealand 2018a); and
- from 1 July 2019 an MHS has been approved for fishing in Fisheries Management Areas (FMAs) 1, 2, 8 and 9 when targeting snapper, tarakihi, trevally, red gurnard and John dory (Fisheries New Zealand 2019). The approvals for use of MHS in North Island inshore fisheries are subject to certain spatial and depth restrictions.

The current approval process for new types of trawl nets is set out in regulation 71A of the Fisheries (Commercial Fishing) Regulations 2001, as a result of amendments in the Fisheries (Trawling) Amendment Regulations 2017 and Fisheries (Commercial Fishing) Amendment Regulations 2018. These regulatory changes resulted from the Ministry for Primary Industries' *Enabling innovative trawl technology* (EITT) programme that aims to facilitate the trialing and use of innovative trawl technologies.

A new trawl net ('net A') can be approved if it performs "at least as well as" a specified net ('net B') in providing for the utilisation of fisheries resources while ensuring sustainability (i.e., in meeting the purpose of the Fisheries Act 1996). The specified net (net B) used for comparative purposes is a trawl net that is already approved for use and that the chief executive considers appropriate for making the comparison, taking account of the types of nets that are currently approved for use in the fishery of interest.

Comparisons between nets A and B must be made by assessing a range of matters laid out in regulation 71B, including:

- species composition;
- size composition;
- impact on protected species; and
- impact on benthic species.

Regulation 71B appears to recognise that gaining comprehensive data on all of the relevant matters may not be possible, and permits the chief executive to consider either how nets compare or how they are "likely to compare".

The original approvals for PSH gear in inshore (and also certain deepwater) fisheries drew on a range of analyses of both experimental data and data from commercial fishing operations conducted under the auspices of Special Permits. The parallel development of the regulatory regime led to uncertainty regarding the evidence required to meet the standard for approval. Having achieved initial approvals for the use of the MHS, there is a desire from both PSH and Fisheries New Zealand to identify an efficient approach for evaluating whether an MHS can be deployed in new fisheries (i.e., in additional areas or for targeting different species).

O'Driscoll & Millar (2017) successfully compared the selectivity of hoki, and other species, for an MHS and conventional trawl by using length frequency data collected from twin trawl gear fitted with MHS and conventional mesh codends. However, key aspects of that study cannot be readily transferred to an inshore fishery scenario. Inshore vessels are smaller than the vessel used in the hoki investigation, limiting the number of additional personnel that can be carried to undertake data collection. More importantly, twin trawl gear is not commonly used in New Zealand inshore fisheries so the opportunity to fish two different codends at the same time is not available.

O'Driscoll & Millar (2017) estimated the *absolute* selectivity of hoki for the two gear types using the SELECT method (Millar 1992), and making use of data collected from a third gear type: a conventional net fitted with a 40 mm liner assumed to be non-selective for the fish sizes encountered. However, for the net A versus net B comparisons required, the relative performance of the two gears can be assessed without estimating the absolute selectivity of either gear.

Minimum Legal Sizes (MLS) for retained fish are currently a common management setting in New Zealand inshore fisheries. As a result, the proportion of snapper (in particular) above and below the MLS was an important consideration in the comparison of MHS and conventional trawl performance in northern inshore fisheries. Catch weights of snapper above and below the MLS, recorded as part of the statutory catch data, facilitated these comparisons for northern inshore fisheries where catch-at-length data were not readily available. However, the intended approach of using focused trials to collect data for future comparisons of MHS and conventional gear allows catch-at-length data to be collected for key species and permits more detailed comparisons of relative selectivity to be made.

This report considers data from two voyages, conducted off the south Canterbury coast in January 2019, where an MHS1480 trawl was compared with a conventional mesh trawl net as typically used in the South Island east coast fishery. The trial used an alternating tow approach where the same tow line was fished first with one gear, then subsequently with the other.

We consider whether the approach used on these trips has the potential to find evidence of meaningful differences in the performance of two net types when such a difference exists, and if it could therefore provide a suitable 'standard' approach for net A vs. net B comparisons in the future.

2. METHODS

Tows using a current inshore MHS net (net A; MHS1480) and a 100 mm (4 inch) mesh trawl (net B) were made by the participating vessel operating under special permit SP-682. The mesh trawl was a Kenton wing trawl as normally used by the vessel in the South Island east coast fishery. This is a two panel trawl with a 42 m ground rope, 100 m sweeps and 50 m bridles. It has a design headline height of 5 m and, with the normal 2.4 m^2 doors, usually achieves a door spread of 116 m (Hamill 2019). The MHS used in the trial was originally designed to match the performance of 5 inch mesh trawls used in the SNA 1 fishery and has been approved for restricted use in North Island inshore fisheries. Tows using the MHS1480 trawl were conducted by the vessel, prior to the voyages considered here, to allow the crew to familiarise themselves with the gear.

2.1 Allocation of tows

The comparison involved changing between the normal 100 mm mesh codend and the MHS codend after the end of the first day and every second day thereafter. The vessel's crew were asked to fish the same towlines with the two gears on consecutive days at the same time of day. A Plant & Food Research scientist and a Fisheries New Zealand observer were present on the trip and undertook data collection and sampling with the assistance of the vessel's crew. Catches from each tow were separated and binned by species and catch weights were estimated from the bin counts.

Planning documents prepared prior to the first voyage emphasised that sampling would focus on red cod, tarakihi and red gurnard with other species sampled opportunistically, although it was intended that the approach could be applied on trips targeting whatever species would have been targeted on a regular commercial fishing trip.

A total of 23 tows was undertaken during the two trial trips (Table 1). Barracouta was the reported target species for all tows. The 23 tows comprised 11 sets of 'paired' tows and one unpaired tow that was undertaken using the conventional mesh net. The unpaired tow was not part of the study design but was

Table 1: Basic details of tows undertaken during January 2019 and considered in this report. The *Tow* column describes the order in which the operations were undertaken and is identical to those used in previous reports describing the same data (e.g., Hamill 2019). The *Tow-pair* column was added for the purpose of this report to make this essential element of the design clearer. Species codes are defined in Table 2.

						Estimated catch (kg)				
Tow-pair	Tow	Gear	Date	Start time	Duration (min)	BAR	GUR	SPE	TAX	TAR
A	1	Conv.	18 Jan.	06:20	125	3090	–	–	–	–
A	5	MHS	19 Jan.	06:25	120	1140	–	–	–	–
B	2	Conv.	18 Jan.	09:30	165	2490	–	–	–	–
B	6	MHS	19 Jan.	10:25	125	1470	1.5	–	–	–
C	3	Conv.	18 Jan.	14:50	110	10	60	3	0.5	1
C	7	MHS	19 Jan.	14:30	130	90	90	5	–	–
D	4	Conv.	18 Jan.	18:15	105	690	40	80	0.5	40
D	8	MHS	19 Jan.	18:00	110	150	15	540	5	60
E	9	MHS	20 Jan.	05:30	210	1950	165	15	35	75
E	13	Conv.	22 Jan.	05:40	210	4200	210	30	30	30
F	10	MHS	20 Jan.	09:55	115	360	75	10	10	1
F	14	Conv.	22 Jan.	10:00	110	180	120	15	20	1
G	11	MHS	20 Jan.	13:35	120	90	150	150	20	150
G	15	Conv.	22 Jan.	13:35	125	1980	120	30	7	1
H	12	MHS	20 Jan.	16:15	125	180	180	2	7	30
H	16	Conv.	22 Jan.	16:25	125	2820	180	–	5	30
I	17	Conv.	23 Jan.	05:30	120	1170	180	45	12	30
I	21	MHS	24 Jan.	05:35	120	330	180	–	5	20
J	18	Conv.	23 Jan.	09:00	120	510	180	1	7	3
J	22	MHS.	24 Jan.	09:00	120	240	180	15	1	1
K	19	Conv.	23 Jan.	12:00	120	360	780	–	0.5	1
K	23	MHS	24 Jan.	12:10	120	570	960	–	–	–
–	20	Conv.	23 Jan.	15:15	45	360	60	1	–	–

unmatched because of operational constraints on the second voyage. Operational details are described in greater depth in the voyages report (Hamill 2019).

The approach for alternating trawls was intended to achieve a serviceable sample design while avoiding extraneous gear changes. Accordingly, four tows were made by one of the two gear types each day with the gear changed every second day. In general, the four tows undertaken on the first day after a gear change were intended to match the four tows made with the other gear on the previous day and then an additional four tows were made on the second day that were the first tows of four new tow-pairs.

Among the 11 tow-pairs from the trial, the first tows of seven tow-pairs were made using the conventional mesh net and the MHS was used to make the first tow of the other four pairs. As a result of tows occurring over two trips, there are some tow-pairs where the second tow was made two days after the first, but still at similar times of the day (Table 1).

The pattern of replicating tows in the same location using each gear type is referred to as a 'paired tow strategy' by Hamill (2019). Similar studies have been undertaken elsewhere and referred to as 'alternate-tow' studies, although most alternate-tow studies have compared the catch of one test gear with a non-selective gear rather than two gears with unknown selectivities (see e.g., Wileman et al. 1996).

We have adopted the 'alternate-tow' terminology here to more clearly distinguish the approach from twin trawl studies, such as the hoki study undertaken on the FV *Rehua* (O'Driscoll & Millar 2017). Alternate haul, parallel haul, twin trawl and trouser trawl selectivity studies can all be considered *paired-gear* trials (Sistiaga et al. 2009).

2.2 Data collection

All gurnard, red cod and tarakihi caught on the trips were measured except on two tows where subsamples of large gurnard catches were measured for length. On the two tows where subsampling occurred, gurnard were placed in bins with capacity for approximately 30 kilogram of fish, then every fourth bin was sampled until a total of six bins of fish had been measured.

Data were recorded on sampling sheets and transcribed to a spreadsheet. Data recorded from each tow included:

- tow details (e.g., gear type, date, time, location, duration, etc.);
- estimated catch weights in kilograms of all species;
- length sampling tally data.

Fish were measured to the nearest centimetre rather than the standard fisheries approach of rounding down to the nearest whole centimetre. For the binary selectivity analyses described below, fish measured at the threshold length were therefore assigned at random to either the above or below threshold categories.

In addition to this independent data collection, the statutory catch and effort data are available. Video footage was also collected from within the trawl and the on-deck fish sorting operation, but has not been used in this report.

2.3 Assessing differences in catch rate

Catch rates of the target species achieved by net A and net B are compared to gauge potential differences in the benthic impacts that might occur as a result of using one gear rather than the other. High target species catch rates are preferred because these allow a given catch to be achieved with fewer tows (Suuronen et al. 2012, McConnaughey et al. 2020).

The catch rate analyses described here focus on barracouta, the target species on the two trips that comprise the trial. Comparisons of catch rates of non-target species may provide some information on how the catch rates of a particular species are likely to compare when that species is targeted. However, the magnitude of differences in catch rates of any particular species are likely to depend on the target species.

2.3.1 Generalised linear mixed model of barracouta catch rate

The alternate-tow design of the trial generates data well suited to analysis using generalised linear mixed models (GLMMs) with tow-pair random effects. The GLMMs provide a framework that allows the effects on catch of multiple factors to be estimated without requiring an experimental design that controls for each factor.

Difficulties faced when fitting GLMMs using maximum likelihood based approaches are discussed by Breslow & Clayton (1993). For the trial data, we used the `brms` package (Bürkner 2017) to fit Bayesian models for the quantity of catch taken on a tow, with duration and gear type as fixed effects, and random effects for tow-pair. Because the tows comprising each pair fished, as far as possible, the same ground at the same time of day on consecutive days, this feature of the trial data reduces the need to include 'nuisance' variables in the GLMMs to control for their effects.

Barracouta were caught on all tows (Table 1), so the 23 estimates of tow-level catch can be modelled as continuous, positive data. The fitted model was specified using R code as shown in Equation (1). The `gear` factor is required and a `tow_pair` random effect is included to account for differences in catch attributable to unmodelled variability between tow-pairs. The indicator variable `first` is included to account for a possible difference in tow-level barracouta catch resulting from a tow being either the first or second within its tow-pair.

Interest centres on the possibility that one gear type catches distinctly more of the target species. Evidence from a fitted GLMM of a difference would be reflected by an estimated `gear` coefficient that was significantly different from zero. Tow durations were generally quite consistent, particularly within tow-pairs (Figure 1). However, some variation in tow-level barracouta catch might be explained by differences in tow duration.

A lognormal model was compared with a gamma model with a log link function using the leave-one-out information criterion (Vehtari et al. 2017). The gamma model was chosen according to this criterion and is specified as:

$$\text{bar\_kg} \sim \text{offset}(\text{log\_minutes}) + \text{gear} + \text{first} + (1|\text{tow\_pair}). \qquad (1)$$

Specifying the tow duration variable, `log_minutes` as an offset on the log scale effectively means that barracouta catch rate, in kg per minute, is modelled in terms of the other explanatory variables. A preliminary model was fitted with the `log_minutes` estimated as a spline to ensure it was reasonable to assume a directly proportion relationship between tow duration and barracouta catch.

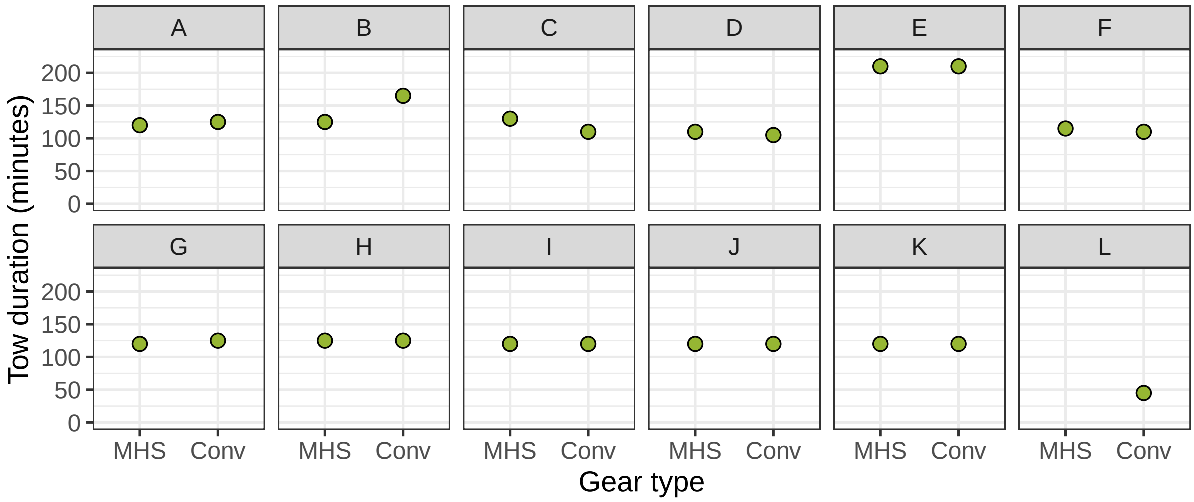


Figure 1: Durations of tows made on the trial by tow-pair and gear type. Note there was no MHS tow made on tow-pair L.

Inferences made from these analyses of catch rates are model-based and their validity requires that the model assumptions are not seriously violated. Model diagnostics appropriate for Bayesian models are used for this purpose (see Appendix D). A alternative, non-parametric, bootstrap analysis of barracouta catch rates is described in Appendix A.

2.4 Assessing differences in size composition

The size data collected during the trial can be analysed in a variety of ways to compare the retention or selectivity characteristics of the two gear types. A preliminary randomisation test (Ernst 2004) was carried out using length frequency measurements made on each tow-pair. For each species, the randomisation test is used to assess whether the collected data provide significant evidence to reject the null hypothesis that the two gears have the same selectivity. Details of the randomisation tests are described in Appendix B.

The result of a randomisation test will be either a conclusion that significant evidence of a difference in selectivity exists or, alternatively, a conclusion that significant evidence of a difference does not exist. The absence of a significant difference does not, of course, imply the two gears have the same selectivities, and more specific information is likely to be required irrespective of the results of the randomisation test.

We carried out two types of further analyses on size compositions: (i) proportions above and below a specified cut-off size (for example, a minimum legal size), similar to the Undersized Catch per Kilo (UCK) measure used in comparisons of snapper off the North Island (Jones & Millar 2018), and (ii) analyses similar to standard estimates of length-based selectivity like those presented by O'Driscoll & Millar (2017).

2.4.1 Proportional catch-at-length

Catch-at-length data from alternate-tow studies can be modelled using Millar's SELECT approach (Wileman et al. 1996). The trial considered here did not collect data from a 'non-selective' gear and so the more normal approach of estimating absolute selectivity is not available. However, it is possible to characterise the *relative selectivity* (see Huse et al. 2000) of the two gear types as the relative proportion of catch taken by MHS for each length class.

For a given species, the proportion of catch of length class l by the MHS net is:

$$R(l) = \frac{N_{lm}}{N_{lm} + N_{lc}}, \tag{2}$$

where N_{lm} is the total number of individuals with length class l caught by MHS and N_{lc} is the number caught by conventional trawl. Equation 2 can be interpreted as a measure of relative selectivity if the same population of fish is assumed to be available to each gear type. The populations are likely to be similar because of the alternate-tow design used in the trial.

Uncertainty in the proportion of catch by MHS in each length class was estimated using a double bootstrapping procedure (Millar 1993). Sets of fish lengths were obtained by first sampling (with replacement) the sets of tow-pairs where the species of interest was caught, then sampling lengths, again with replacement, from each resampled tow.

Defining N^*_{lmb} as the number of fish of length l resampled from MHS tows in double bootstrap resample b and N^*_{lcb} as the number fish of the same length class sampled from conventional trawl tows on the same double bootstrap resample, then bootstrap estimate $R^*_b(l)$ of the proportional catch-at-length, l, for bootstrap resample b is given by:

$$R^*_b(l) = \frac{N^*_{lmb}}{N^*_{lmb} + N^*_{lcb}}. \tag{3}$$

A large number, B, of bootstrap resamples are made and the bootstrap point estimator is the mean of the B resamples of $R^*_b(l)$ as defined in Equation 3. The standard error of the bootstrap estimator is the standard deviation of the B resamples of $R^*_b(l)$. The influence of each tow-pair on the distributions of relative catch of separate length classes can be examined using jackknife-after-bootstrap diagnostic plots (Appendix D).

Equations 2 and 3 give separate estimates of the proportion of catch-at-length for each length class, and each estimate is based solely on the catch of fish in that length class. Length-based selectivity of towed gear is usually assumed to be a smooth function of fish length. Parametric models of selectivity are more difficult to justify for relative selectivity, but it can be reasonably assumed that the proportion of catch-at-length (i.e., relative selectivity) varies smoothly with length.

A smooth curve could be generated by extending the bootstrapping procedure to fit a simple spline to each bootstrap sample and then inferring uncertainty in the proportional catch-at-length from the full distribution of bootstrapped splines (see e.g., Hastie et al. 2009, §8.2.1). As an alternative, we fitted Bayesian mixed model splines (Marley & Wand 2010) to the bootstrapped length-specific proportions of catch-at-length (Plummer 2019) as a separate step. JAGS code used to fit the model is given in Appendix E.

2.4.2 Differences in binary selectivity

Binary selectivity refers to situations where there is interest in a specific size breakpoint, such as the proportion of fish above and below the MLS. This can be analysed in terms of catch numbers or catch weight, depending on the data available. In the trial there was specific interest in the proportions of tarakihi above and below the MLS; that is, in the difference in probabilities that tarakihi less than 25 cm would be retained by net A and net B.

It is commonly found that the variation in length of fish within a tow is less than the variation between tows. In the alternate-tow design it is expected that similar length distributions of fish are available to the tows in the pair. An estimator for a difference in the proportions of fish above and below a cut-off length, accommodating dependencies in the length classes of fish caught in the same tow and in different tows within the same tow-pair, adapted from Fleiss et al. (2003, §15.2), is described in Appendix F. Using this estimator, we calculated point estimates and standard errors for differences in the proportions of fish above or below selected length cut-offs. The significance of the estimated differences was tested using a test statistic based on the standard normal distribution as described in Appendix F. Confidence intervals based on the normal approximation suggested by the method were also calculated.

2.4.3 Bootstrapping binary selectivity

The proportions of fish below the MLS, or other cut-off lengths, can also be analysed using the same two-step bootstrapping procedure applied to the proportional catch-at-length estimator. Tow-pairs are resampled in the first step and then fish lengths are resampled within each tow. For each bootstrap sample, the overall proportions of fish lengths above or below the MLS are calculated for each gear type. Denoting the difference in proportions of fish below the reference length between MHS gear and conventional trawl on bootstrap sample b as $\Delta\Phi_b^*$, we have:

$$\Delta\Phi_b^* = \frac{\sum_{i\in\tau_b} Y^*_{ibm^+}}{\sum_{i\in\tau_b} n_{im}} - \frac{\sum_{i\in\tau_b} Y^*_{ibc^+}}{\sum_{i\in\tau_b} n_{ic}}. \tag{4}$$

where τ_b is the set of tow-pairs generated for bootstrap sample b and n_{im} and n_{ic} are the number of fish measured from the MHS and conventional tows, respectively, on tow-pair i. Here $Y^*_{ibm^+}$ and $Y^*_{ibc^+}$ are the number of fish measurements smaller than the cut-off length for tow-pair i on bootstrap resample b by MHS and conventional trawl, respectively.

The bootstrap estimator is the mean of the full set of $\Delta\Phi_b^*$ calculated for B bootstrap samples and the standard error is estimated as the standard deviation of the same set of bootstrap samples.

2.4.4 GLMMs assessing size selectivity

Various GLMMs fitted to the summaries of the sampled length data may also be useful for understanding aspects of differences in selectivity between gear types. As with analyses of catch data, the main advantage of GLMMs for comparisons of length data is the ability to estimate the effects of multiple factors on the observed size composition of the catch. Therefore, we use GLMMs to examine the effect of total catch weight and tow order on simple models related to selectivity to consider whether these might affect selectivity more generally.

Millar (2018) modelled catch weight of sub-MLS snapper per kilogram of snapper above MLS. We apply a similar approach to the tarakihi data from the trial, except we model catch numbers instead of weights. Accordingly, we fitted a negative binomial model to the number of small individuals, with the logarithm of the number of large individuals caught as an offset. Using the `brms` package (Bürkner 2017), the model was specified as:

$$\text{tax\_num} \sim \text{offset}(\log(\text{tar\_num})) + \text{gear} + \text{first} + \log(\text{tot\_catch\_kg}) + (1|\text{tow\_pair}). \qquad (5)$$

Models of this form could not be fitted to the length observations of gurnard and sea perch from the trial because not all fish caught were measured. However, if weight data are collected for the different size categories then these could be modelled using a similar approach.

3. RESULTS

Prior to the trip, it was considered that catches of large quantities of spiny dogfish may result in the accumulation of sand and mud in the MHS. This was monitored and no sign of accumulation was seen during the trial (Hamill 2019). One small albatross and one sooty shearwater were caught during the trial, both while using the mesh codend. Both birds were caught in the trawl wings.

The tows recorded as part of the trial were the only tows recorded by the vessel in its statutory catch and effort data over the period of the trial. There is a reasonable match between the trial data records of tow data and the statutory effort data, although it was noted that the vessel recorded all tows using the 'BT' method code. Recorded tow durations were comparable other than for one tow where the same start and end time was recorded in the statutory data. Information from the vessel's plotter demonstrated that a high degree of consistency was achieved between the towlines fished on the paired tows (Figure 2).

The catch composition data recorded by the onboard science staff included all species caught, whereas the estimated catches from the vessel on the Trawl Catch Effort and Processing Return form record only the top four species plus sub-MLS tarakihi (TAX). As a result, the trial dataset includes catch data for species that do not appear in the statutory estimated catch data (Table 2). For the key species that were recorded in the statutory estimated catch data, catch quantities were broadly similar in the Plant & Food (PF) Research data. However, for some minor species, differences arise because of the 'top 4' recording; for example, estimated catches of STA are recorded on 19 tows in the trial dataset, but only appear in the statutory estimated catch data for 5 tows.

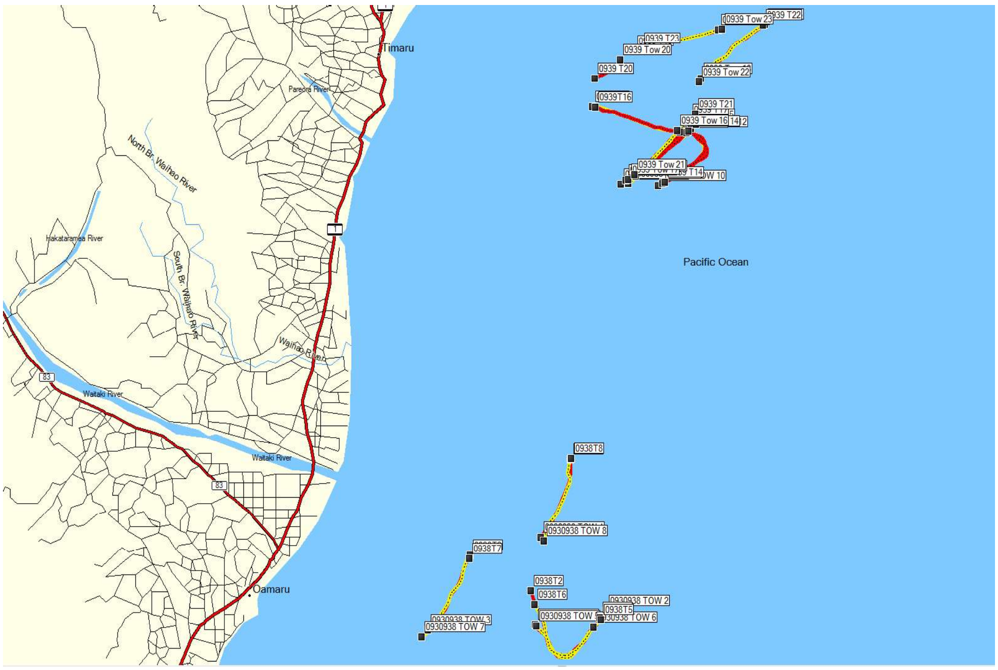


Figure 2: Tow tracks from the trial trips. Yellow points indicate where the MHS codend was on the net and red points where the mesh codend was used. From Hamill (2019, figure 1).

Table 2: Aggregate catches (kg) by species recorded in the statutory catch estimates (CEL data) and trial dataset (PF data). Species with overall catches of less than 20 kg are grouped under the OTH code.

			CEL data		PF data	
Code	Common name	Scientific name	BT	MHS	BT	MHS
BAR	Barracouta	*Thyrsites atun*	17 880	6 570	17 970	6 573
SPD	Spiny dogfish	*Squalus acanthias*	3 240	425	4 055	405
GUR	Red gurnard	*Chelidonichthys kumu*	1 590	1 890	1 930	1 996
CAR	Carpet shark	*Cephaloscyllium isabellum*	560	360	780	475
STA	Giant stargazer	*Kathetostoma* spp.	210	180	471	636
SQU	Arrow squid	*Nototodarus* spp.	540	270	660	315
SPE	Sea perch	*Helicolenus* spp.	90	720	205	737
GSH	Ghost shark	*Hydrolagus novaezealandiae*	180		610	77
TAR	Tarakihi	*Nemadactylus macropterus*	95	210	220	420
JMA	Jack mackerel	*Trachurus* spp.	510		525	71
SSK	Smooth skate	*Dipturus innominatus*	30	60	245	99
SCH	School shark	*Galeorhinus galeus*	90		177	147
SDO	Silver dory	*Cyttus novaezealandiae*			301	12
WIT	Witch	*Arnoglossus scapha*	120	90	204	104
ELE	Elephant fish	*Callorhinchus milii*	105	60	145	105
MOK	Blue moki	*Latridopsis ciliaris*		210	10	210
SKI	Gemfish	*Rexea* spp.	30		191	19
HAP	Hāpuka	*Polyprion oxygeneios*	30	60	72	137
RSK	Rough skate	*Zearaja nasuta*			47	146
SPO	Rig	*Mustelus lenticulatus*			5	160
RCO	Red cod	*Pseudophycis bachus*			64	68
SCG	Scaly gurnard	*Lepidotrigla brachyoptera*			106	20
SWA	Silver warehou	*Seriolella punctata*			60	40
SCC	Sea cucumber	*Stichopus mollis*			45	8
KIN	Kingfish	*Seriola lalandi*			16	35
LEA	Leatherjacket	*Meuschenia scaber*			25	21
LIN	Ling	*Genypterus blacodes*			12	31
PIG	Pigfish	*Congiopodus leucopaecilus*			30	7
OTH	Other species				51	49

3.1 Catch rates

The total barracouta catch from tows by the conventional mesh net over the course of the trial was considerably greater than that from MHS tows (Figure 3) but MHS catches were higher for some species (Figures 3 and 4). The conspicuous difference between gears in spiny dogfish catch (SPD; Figure 3) is mostly due to a catch of three tonnes on tow 20 using the conventional trawl net; there was no MHS tow paired to tow 20 (Table 1).

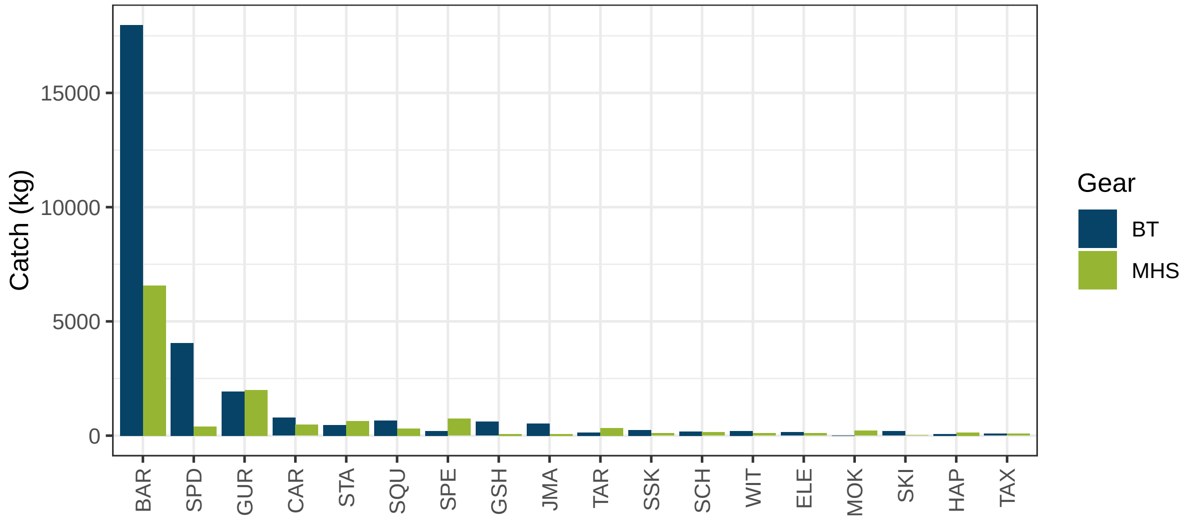


Figure 3: Aggregate catches by species and gear, for the key commercial species caught on the trial using the trial dataset (where key commercial species are species where estimated catches were also recorded in the statutory data). Species codes are defined in Table 2.

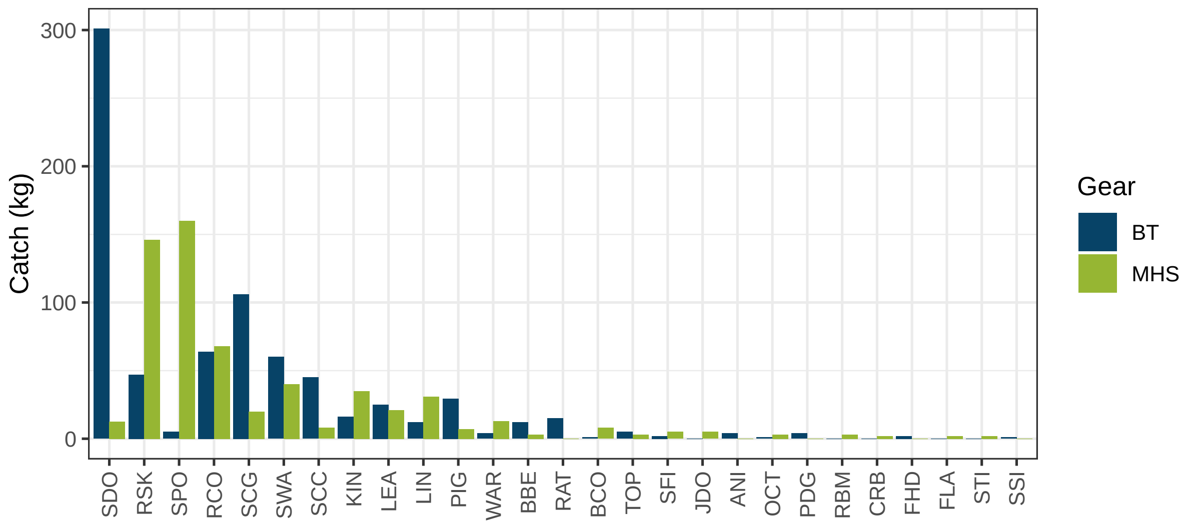


Figure 4: Aggregate catches by species and gear, for the other species caught on the trial using the trial dataset (those species not recorded in the statutory data). Species codes are defined in Table 2.

3.1.1 Catch rate of barracouta

Catches of barracouta show considerable variation among tow-pairs (Figure 5). However, within most tow-pairs, the catch of barracouta made by conventional trawl was clearly higher than the catch by MHS (Figure 5). Preliminary models fitted to barracouta catch suggested that including the effect of tow duration as an offset was reasonable, effectively modelling the catch rate of barracouta in kilograms per minute.

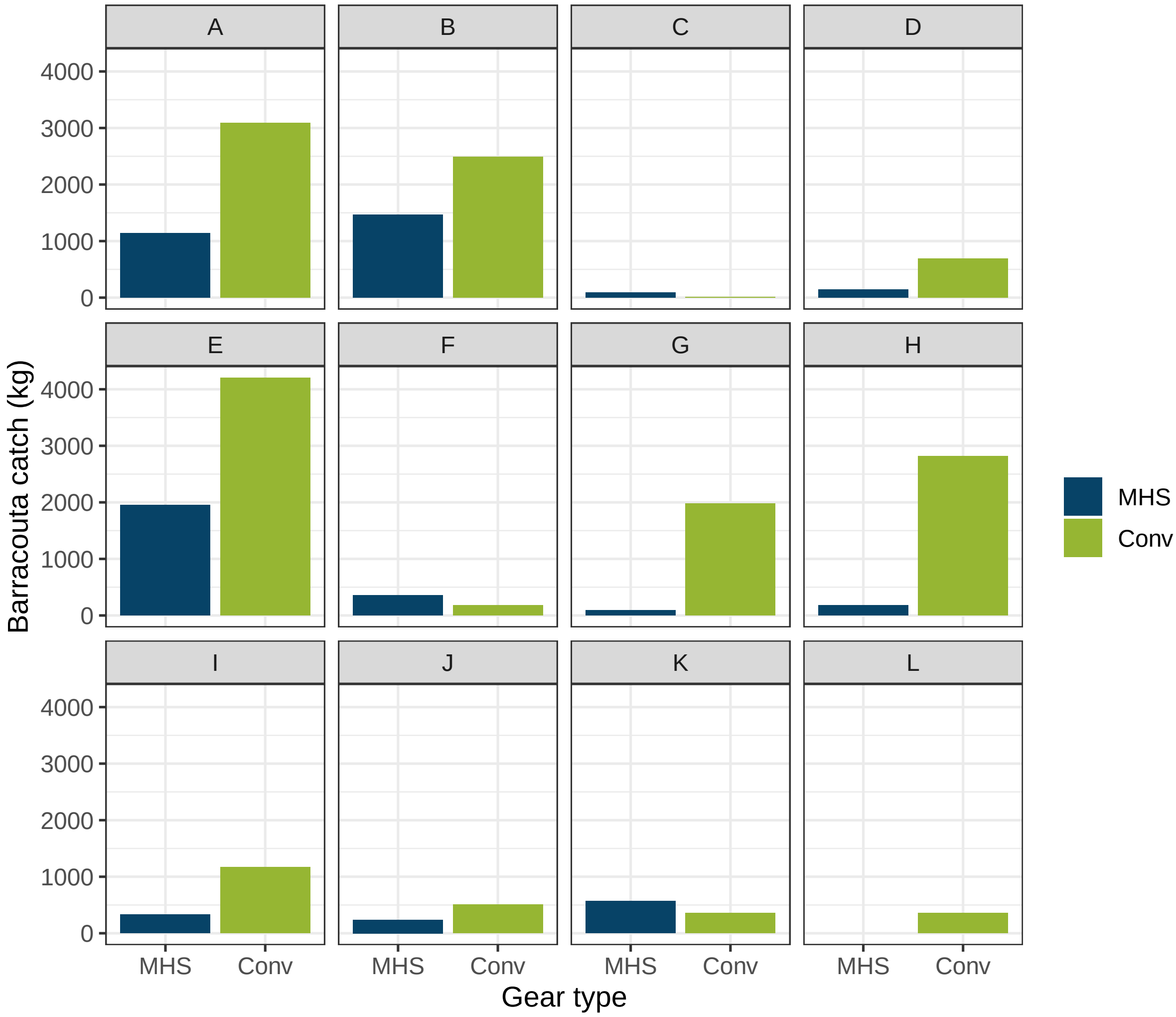


Figure 5: Catches of barracouta by MHS and conventional gear tow-pairs. Note: there was no MHS tow made on tow-pair L.

According to the fitted model, use of MHS instead of conventional trawl has an effect of -1.047 on log-scale barracouta catch (Table 3). The 95% credible interval of the log-scale MHS effect, (-1.9, -0.191), is exclusively negative. This provides evidence that catch rates of barracouta by the MHS net are significantly lower than by the conventional trawl net.

The exponentiated estimates can be interpreted as multiplicative effects on the natural scale as shown in Figure 6a. This indicates that the multiplicative effect of using the MHS compared with the conventional net lies between exp(-1.9) = 0.15 and exp(-0.191) = 0.826.

Although there appear to have been differences among tow-pairs in the barracouta catch (Figure 5), estimates of random tow-pair effects are uncertain (Figure 7); in effect, two tows per set do not provide much information.

Furthermore, there is a degree of confounding between estimates of tow-pair random effects, the effects of using MHS gear (Figure 6a), and the effect of the first tow of the tow-pair relative to the second (Figure 6b). The standard deviation of the tow-pair random effects is also quite uncertain (Table 4). The evidence for a difference in barracouta catch rates between tow-pairs is therefore not overwhelming.

A complementary bootstrapping analysis of barracouta catch rates is described in Appendix A, providing results that are broadly consistent with the GLMM results.

Table 3: Posterior summaries of log-gamma barracouta catch rate model fixed effects coefficients.

	Estimate	Est.Error	Q2.5	Q97.5
Intercept	2.618	0.452	1.774	3.591
gear = MHS	-1.047	0.432	-1.900	-0.191
first = yes	-0.427	0.422	-1.239	0.431

Table 4: Posterior summaries of log-gamma barracouta catch rate model shape parameter and random effects standard deviation.

	Estimate	Est.Error	Q2.5	Q97.5
sd(random tow pair effect)	0.607	0.362	0.041	1.421
shape parameter	1.337	0.482	0.611	2.499

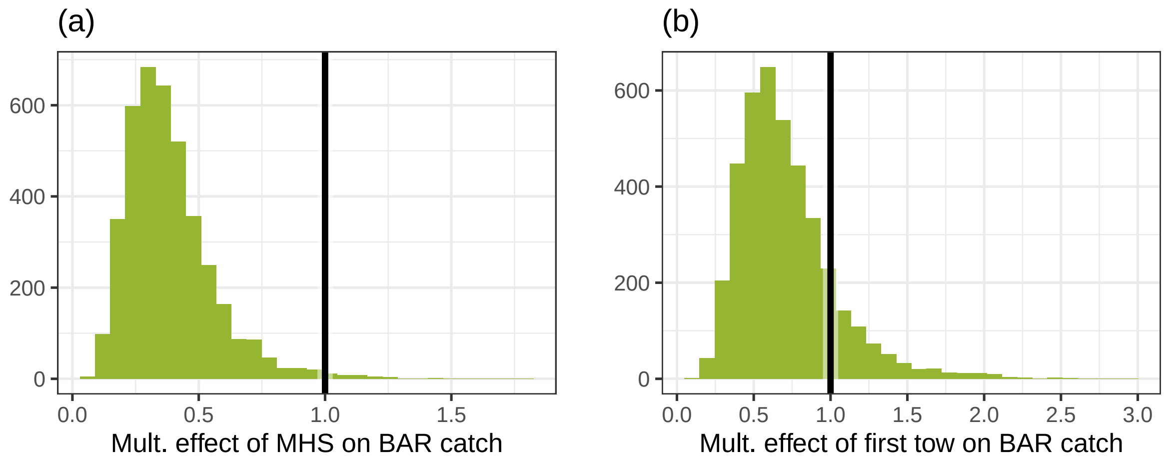


Figure 6: Histogram summaries of the posterior distributions of the multiplicative effects of (a) using MHS instead of conventional trawl and (b) the first tow of a tow-pair on barracouta catch.

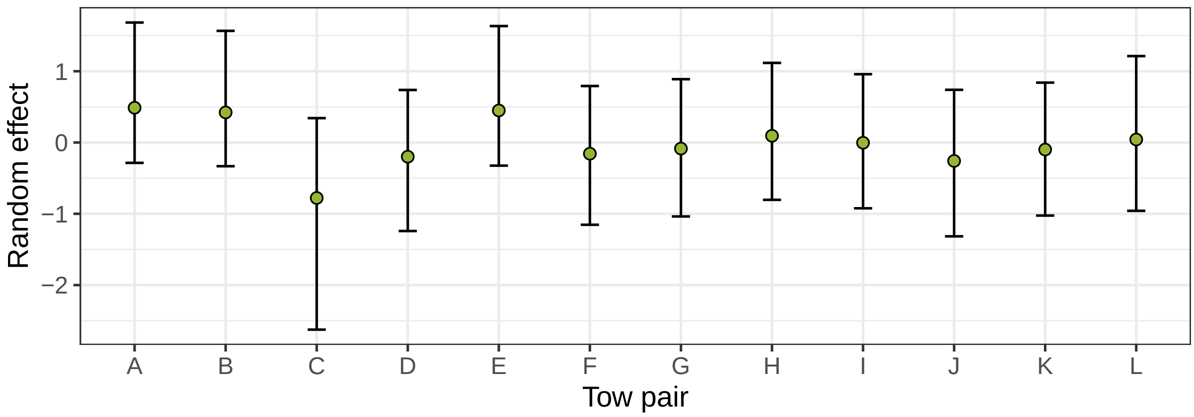


Figure 7: Random tow-pair effects on barracouta catch rate.

3.2 Size composition

Although all tows on the trial targeted barracouta, the species of primary interest were gurnard, tarakihi and red cod. There was minimal sampling of barracouta lengths (Table 5). All tarakihi and red cod caught were sampled, but catch of red cod on the trial was minimal (Table 2).

Analyses of size composition are limited to the key commercial species where sufficient lengths were measured to support statistical comparisons of size between gears: GUR, TAR and SPE. Gurnard was caught on most tows and gurnard lengths were measured on all tows where gurnard was caught, although on two tows where the largest gurnard catches occurred only a sample of the catch was measured. A moderate sample of sea perch lengths was collected during the trial but there were some tows where sea perch was caught and not measured (Hamill 2019).

Table 5: Total fish by species with length measurements. Species codes are defined in Table 2.

Species code	Number measured
GUR	3430
TAR	1663
GSH	653
SPE	504
BAR	345
SWA	162
RCO	98
LEA	67
MOK	64

3.2.1 Red gurnard

A total of 1588 gurnard caught by MHS gear and 1842 by conventional 4 inch mesh gear were measured in the trial (Figure 8). All gurnard caught were measured, except on MHS tow 23 and conventional tow 19 where subsamples were measured. Measured gurnard caught by MHS had a mean length of 37.7 cm fork length compared with a mean of 37.6 cm for gurnard caught by conventional trawl.

There is no MLS for gurnard and length at sexual maturity is thought to be around 23 cm (Fisheries New Zealand 2018b). The smallest gurnard caught by either gear was 26 cm and the great majority of gurnard measured were greater than 30 cm (Figure 8). More gurnard were measured on the trial than any other species so, to illustrate the methods for analysis of binary selectivity, we define a reference length of 35 cm (Figure 9).

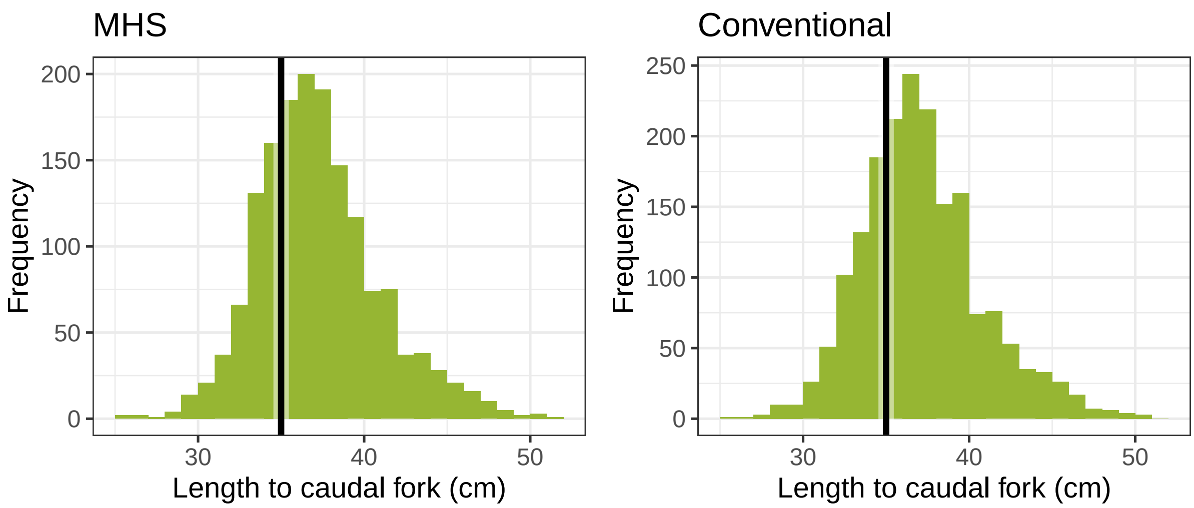


Figure 8: Histograms of lengths of gurnard on tows using MHS and conventional trawl. The vertical lines depict an arbitrary 35 cm cut-off.

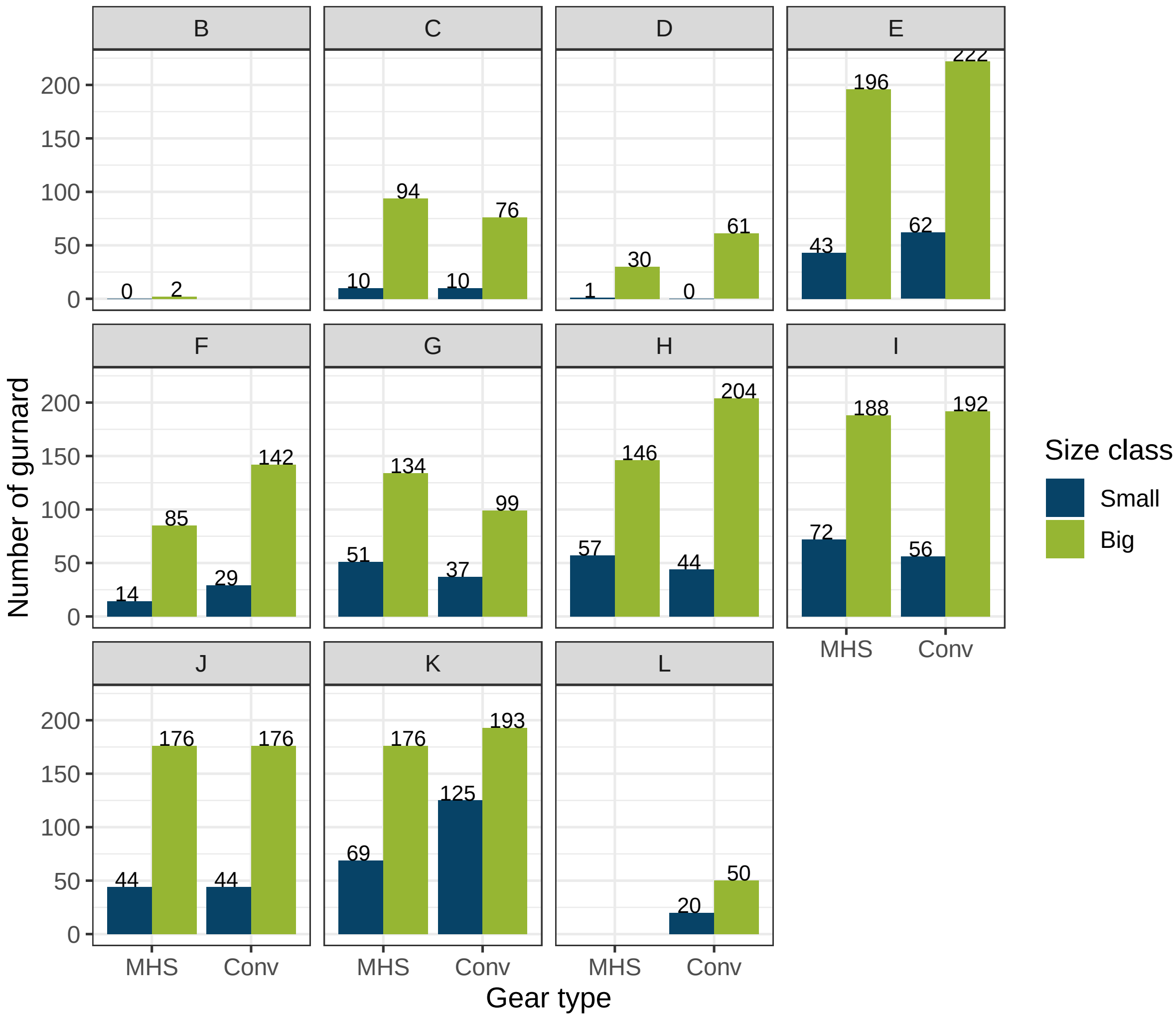


Figure 9: Number of gurnard less than and greater than 35 cm length by gear type and tow-pair.

A randomisation test of relative selectivity suggested the null hypothesis of constant MHS proportion for catch of gurnard at all lengths cannot be rejected (two sided p-value = 0.938). Similarity in the length distributions of red gurnard caught by MHS and conventional gear is apparent in Figures 8 and 9 and further evidenced by the plot of relative gurnard selectivity by MHS (Figure 10b).

Point estimates for the proportions of gurnard less than 35 cm taken by MHS and conventional gear are 0.227 and 0.232, respectively. The test statistic for a significant difference in the sub-35 cm proportions of gurnard caught by the two gear types is -0.17, which indicates no significant difference between the gears (p = 86.7%). Equivalently, the parametric confidence interval of the difference (-0.057, 0.048) includes zero (see Figure 11a).

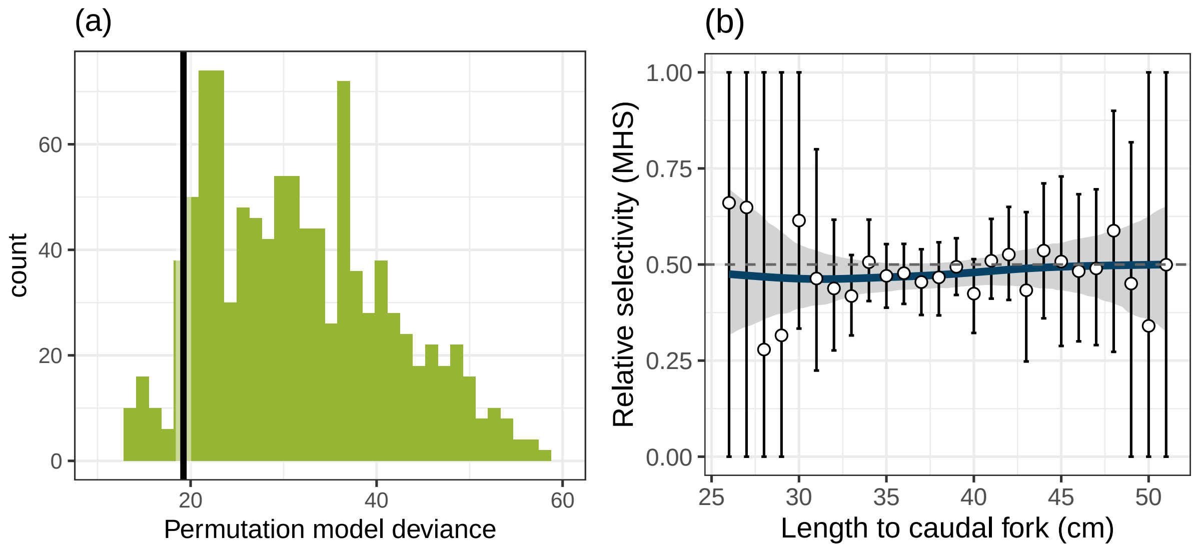


Figure 10: (a) Model deviances of null binomial model of proportion MHS gurnard catch fitted to randomised observations and (b) bootstrapped estimates of MHS relative gurnard selectivity from the trial. The thin vertical bars depict 95% bootstrapped confidence intervals for the length class specific relative selectivities. Circles depict bootstrapped mean relative selectivity. The smooth curve is a Bayesian spline fitted through the length-specific relative selectivities and the shaded region is generated from 95% pointwise credible intervals of the spline fit.

The bootstrapped estimate of the difference in proportions of gurnard less than 35 cm is in good agreement with the parametric estimate (Figure 11a). The confidence intervals estimated by the parametric and bootstrapped approaches are in good agreement (Figure 11).

Although more measurements were made of gurnard than of any other individual species (Table 5), the bootstrap estimate of the difference in sub-35 cm gurnard as a proportion of sub-35 cm caught by conventional trawl is not very precise (Figure 11b). The data suggest that MHS is likely to catch between 20 % less and 20 % more sub-35 cm gurnard than the conventional trawl under similar conditions as the trial.

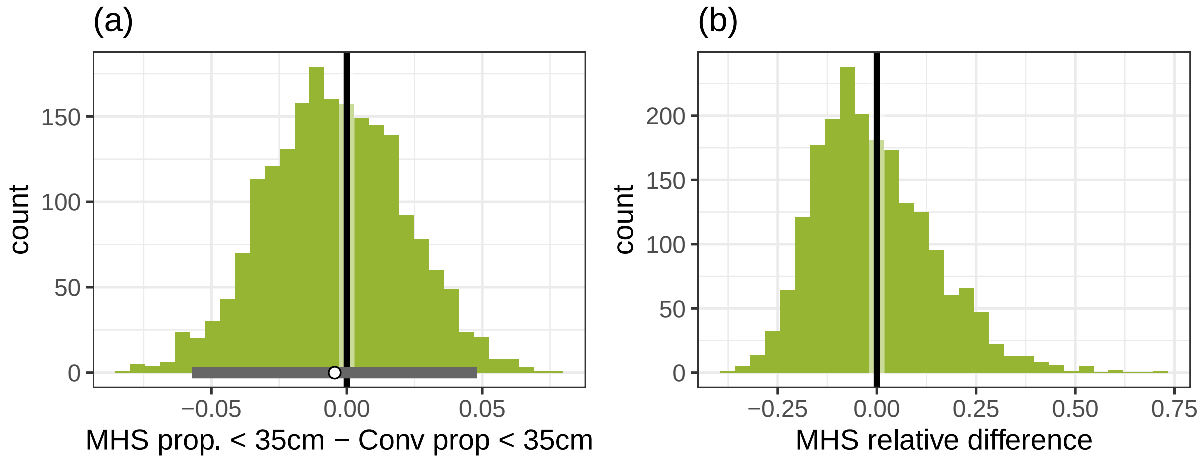


Figure 11: Difference in (a) absolute proportion of gurnard caught by MHS less than 35 cm and (b) difference relative to proportion of conventional trawl less than 35 cm. Estimates around the vertical lines are consistent with no difference between gear types. The open circle and horizontal bar on panel (a) give the parametric mean and 95% confidence intervals for the difference respectively.

3.2.2 Tarakihi

A total of 872 tarakihi were caught by MHS gear and 791 by conventional 4 inch mesh gear (Figure 12). Tarakihi caught by MHS had a mean length of 25.8 cm fork length compared with a mean length of 23.3 cm for tarakihi caught by conventional trawl. All tarakihi caught on the trial were measured.

The MLS for tarakihi is 25 cm, and fish below this length are referred to using the code TAX. Similarities in the numbers of tarakihi caught, as well as the proportions of TAX and TAR on the two alternate tows in the same tow-pairs (Figure 13), are suggestive of within-tow-pair dependence in both quantities. For example, relatively high numbers of tarakihi were captured by both MHS gear and conventional trawl on tow-pairs D, E, F and G. Furthermore, the relative proportions of TAR and TAX also appear to be similar within tow-pairs. For example, both gear types caught predominantly TAR on tow-pair D, but predominantly TAX on tow-pair F.

A randomisation test of the proportional catch-at-length (Figure 14a) indicates that the null hypothesis that the MHS catches a constant proportion of tarakihi at all lengths can be rejected (two sided p-value = 0.02) providing evidence that the two gears have different selectivities for tarakihi.

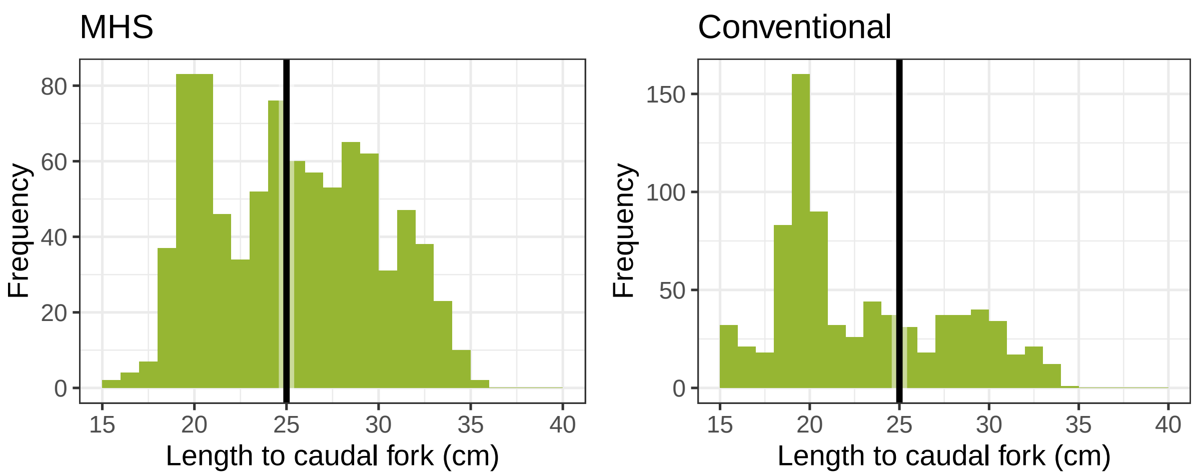


Figure 12: Histograms of lengths of tarakihi on tows using MHS and conventional trawl. The vertical lines depict the 25 cm MLS.

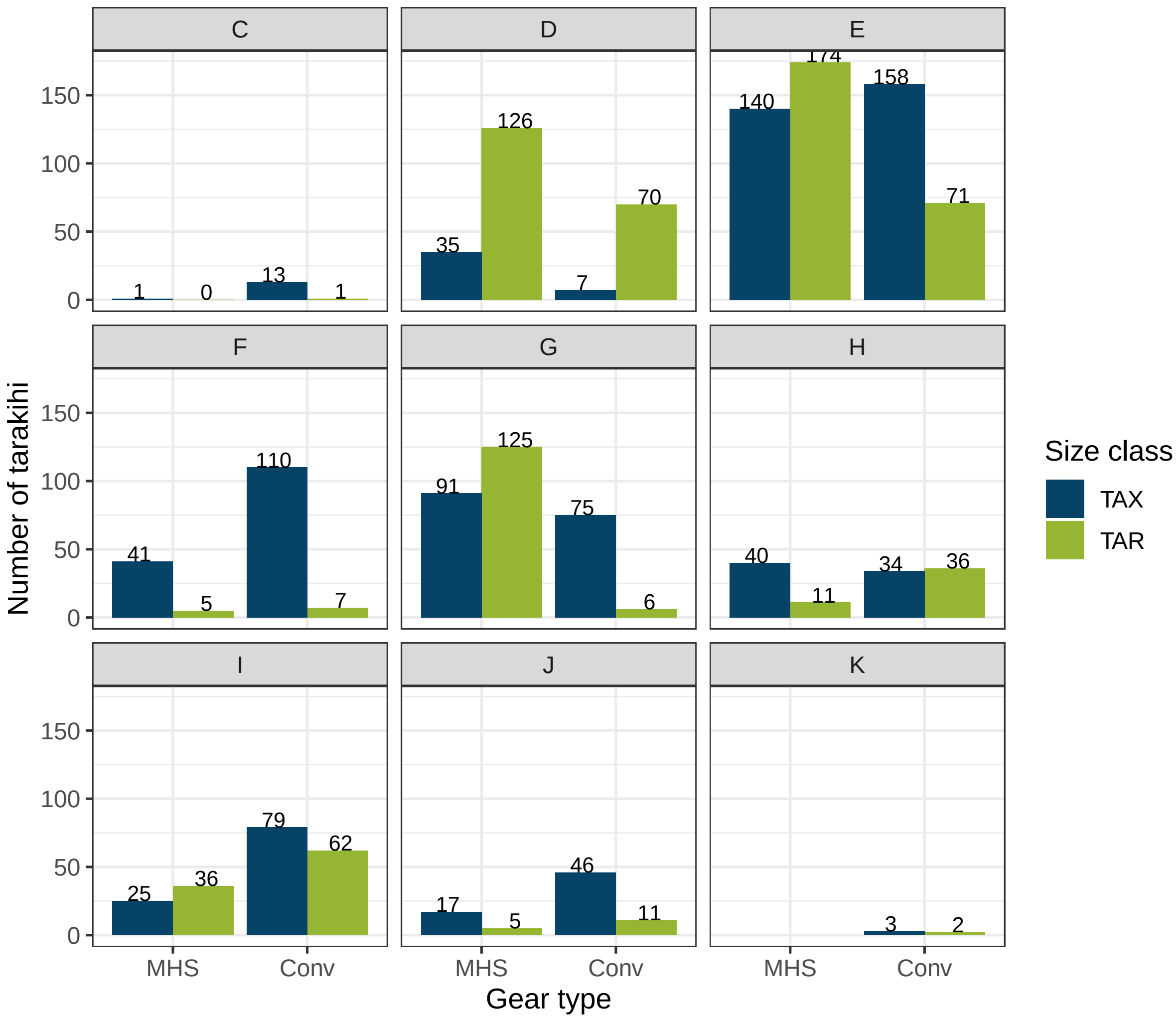


Figure 13: Number of sub-MLS (TAX) and 25 cm + (TAR) tarakihi caught by gear and tow-pair.

Bootstrap estimates of tarakihi selectivity by the MHS relative to the conventional trawl are shown in Figure 14b. The spline fitted to these bootstrapped estimates gives strong evidence that the MHS retains markedly lower proportions of tarakihi in each size class up to 20 cm. There is also some evidence that the MHS achieves higher catch rates of tarakihi greater than 22 cm.

In the parametric analysis of the proportion of TAX, the intra-tow correlation in tarakihi size-class for MHS gear is estimated to be 0.044 and for conventional gear it is estimated to be 0.158. Within tow-pair correlation between the size classes of tarakihi caught by MHS and tarakihi caught by conventional gear is 0.072. The z test statistic has a value of -3.84 indicating that the proportion of TAX retained by MHS (0.447) is significantly lower than the proportion (0.664) retained by conventional trawl (p = 0.013%). The 95% confidence interval (-0.327, -0.106) for the difference is strictly negative (Figure 15a).

The bootstrap estimates of the differences in proportions of TAX are in reasonable agreement with the parametric estimate (Figure 15a). The bootstrapped confidence interval for the difference is slightly wider than the parametric interval. The parametric confidence interval depends on small-sample estimates of the between tow variance in TAX proportions from MHS and conventional gear tows as well as within tow-pair covariance. However, bootstrapping two stages of variability can overstate parameter uncertainty in some situations (Dixon 1993).

Summaries of GLMMs for the numbers of TAX per TAR are provided in Appendix C. The posterior mean estimate of the multiplicative effect of MHS is consistent with an MHS rate of TAX per TAR of roughly 42 percent of the conventional net catch rate.

Although the posterior distribution of the multiplicative effect of MHS on TAX per TAR is mostly less than unity (Figure C-1a), the 95% credible interval is not exclusively less than one. There is no evidence that tow order affects the number of TAX per TAR (Figure C-1b). Similarly, there is little evidence that total catch size affects the number of TAX per TAR (Figure C-1c).

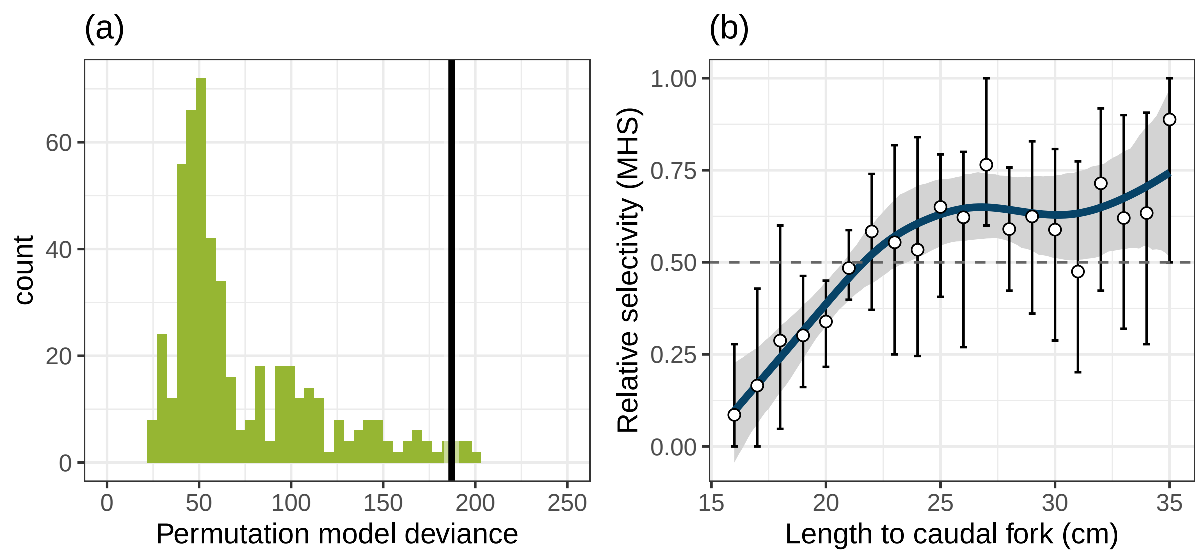


Figure 14: (a) Model deviances of the null binomial model of the proportion MHS catch of tarakihi fitted to randomised observations and (b) Bootstrapped estimates of MHS relative tarakihi selectivity from the trial. The thin vertical bars depict 95% bootstrapped confidence intervals for the length class specific relative selectivities. Circles depict bootstrapped mean relative selectivity. The smooth curve is a Bayesian spline fitted through the length-specific relative selectivities and the shaded region is generated from 95% pointwise credible intervals of the spline fit.

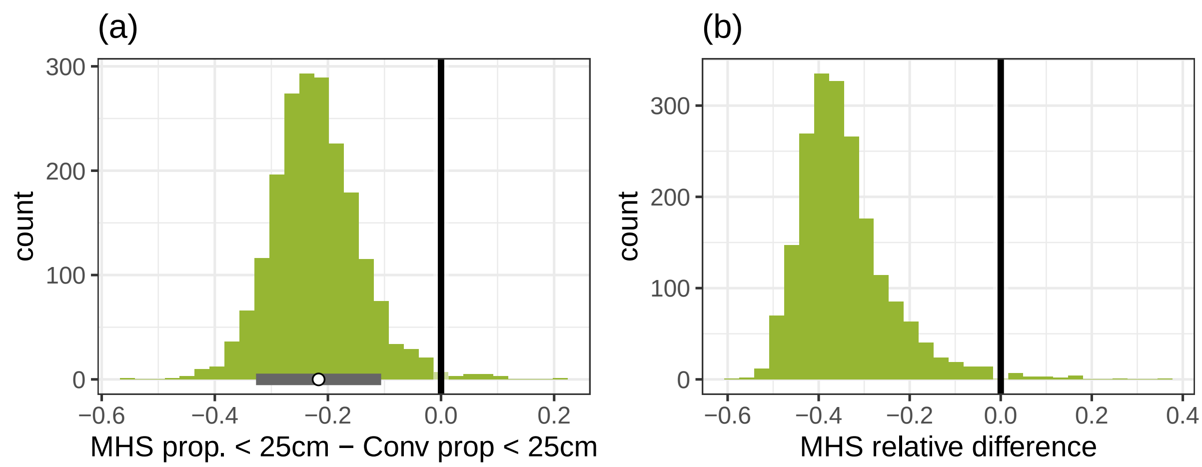


Figure 15: Bootstrapped estimates of (a) the difference in the proportion of tarakihi caught by MHS less than the 25cm MLS and (b) the difference in the proportion relative to the conventional net proportion. Estimates about the vertical lines are consistent with no difference. The open circle and horizontal bar on panel (a) give the parametric mean and 95% confidence intervals for the difference respectively.

3.2.3 Sea perch

Unlike gurnard and tarakihi, sea perch length measurements were not sampled on all tows where the species was caught. A total of 162 sea perch caught by MHS gear and 342 by the conventional mesh net were measured (Figure 17). The greater number of sea perch sampled from conventional tows occurred despite a total sea perch catch weight from the MHS that was more than three times that from the conventional mesh net (Table 2). Sampled sea perch caught by MHS had a mean length of 24.7 cm fork length[4] compared with a mean of 24.1 cm for sampled sea perch caught by conventional trawl.

Sea perch length at 50 percent sexual maturity has been reported as occurring between 15 and 20 cm for females and between 19 and 25 cm for males. With this in mind, we chose a threshold of 22 cm to classify small and large sea perch. Numbers of sea perch varied between tow-pairs as did the proportions of large and small fish (Figure 16). It is evident that a substantial proportion of sea perch measured were below 22 cm and the conventional trawl gear, at least, retained some sea perch smaller than 15 cm (Figure 17).

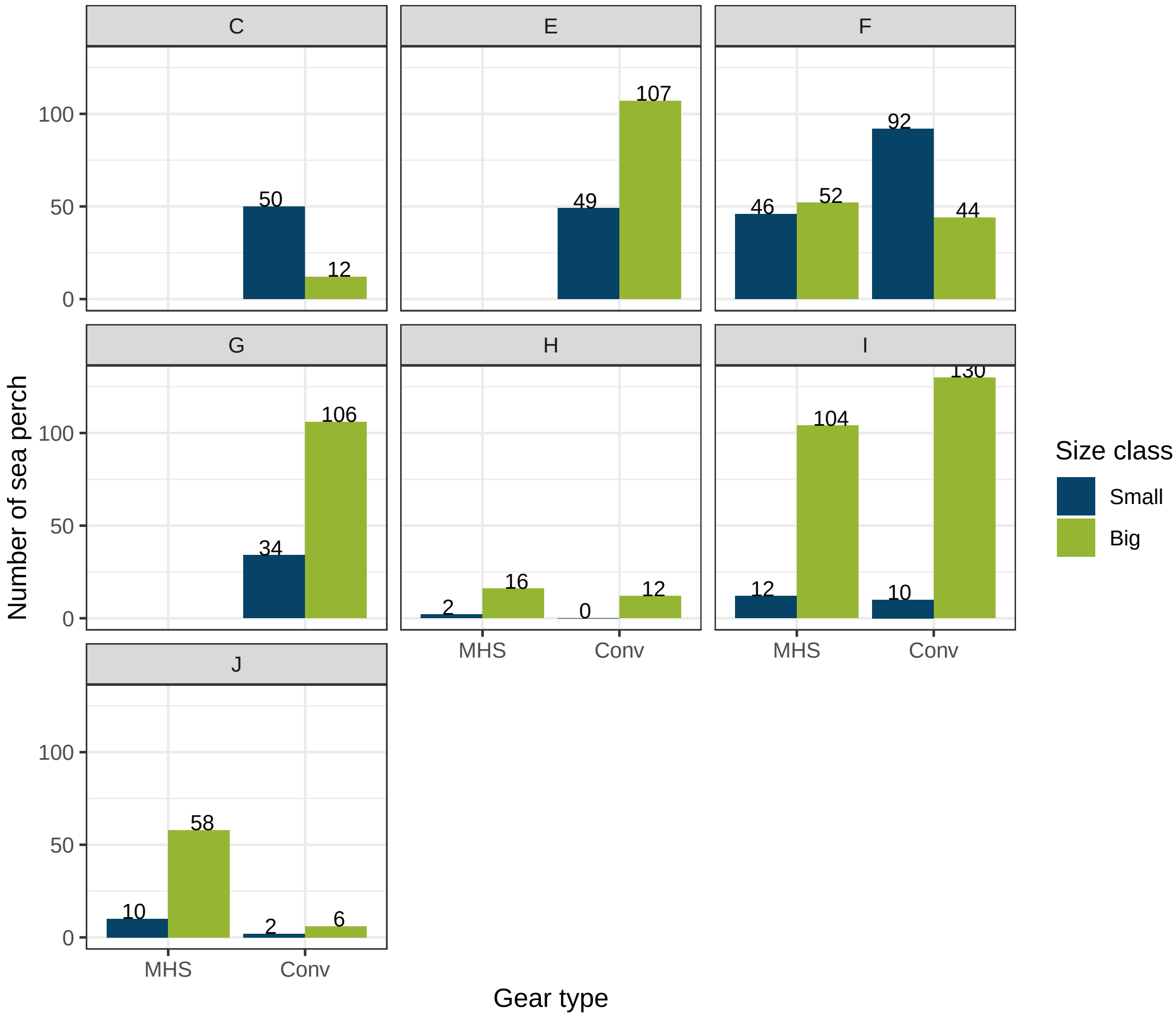


Figure 16: Number of measured sea perch less and greater than 22 cm length by tow-pair and gear type.

A randomisation test of relative selectivity (Figure 18a) suggested the null hypothesis of a constant proportion of MHS catch for all lengths could not be rejected (two sided p-value = 0.9) implying there is not significant evidence the two gears have different sea perch selectivities. Analyses of proportions of catch-at-length hint that the highest relative sea perch selectivity of MHS compared with the conventional mesh net (Figure 18b) might occur at intermediate lengths around 23 cm. However,

[4] for sea perch, the lack of a tail fork implies that fork length is equivalent to total length

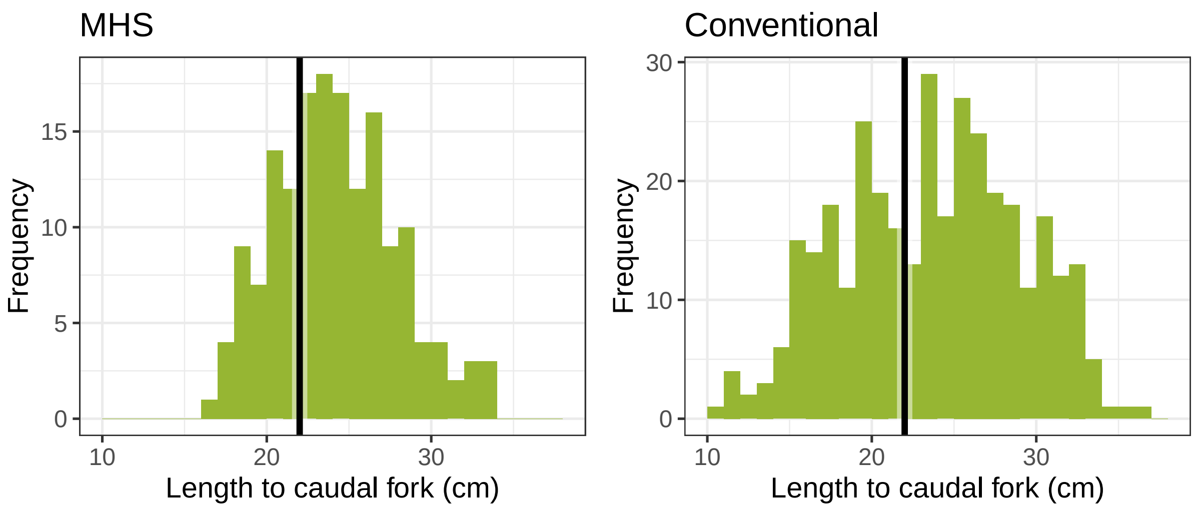


Figure 17: Histograms of lengths of sea perch on tows using MHS and conventional trawl. The vertical lines at 22 cm depict an approximate length at 50 percent sexual maturity.

Figure 18b may be misleading because of inconsistent sampling of sea perch. For instance, the mean of the spline fitted proportions of catch-at-length is below 0.5 for all lengths suggesting MHS has lower sea perch fishing mortality. However, the apparent lower fishing mortality is an artefact of the inconsistent sampling. The shape of the proportional catch-at-length might be less misleading than its scale, but comparisons based on four paired tows should be interpreted with caution.

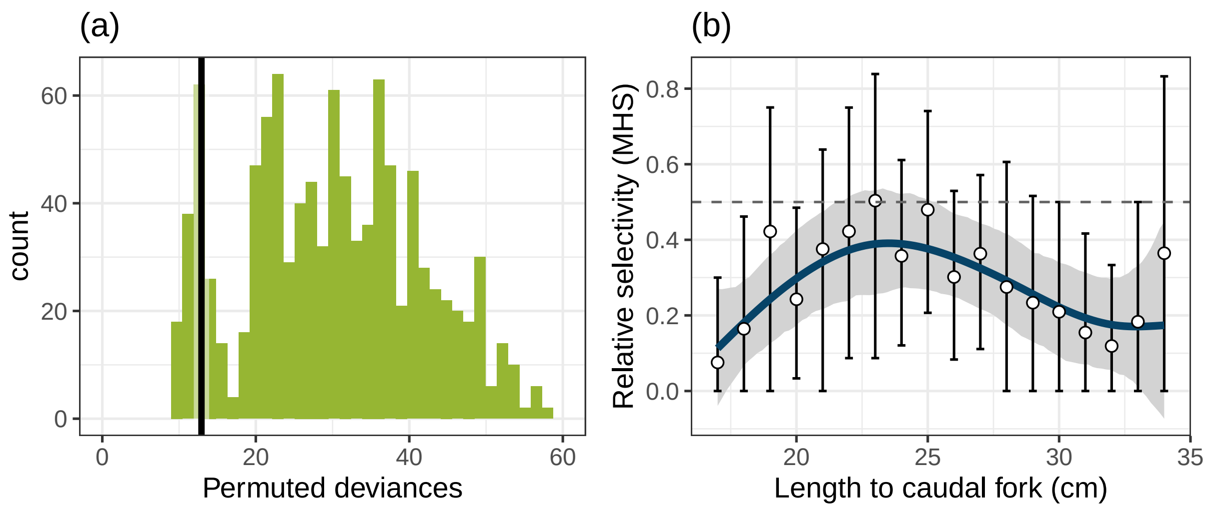


Figure 18: (a) Model deviances of the null binomial model of the proportion MHS catch of sea perch fitted to randomised observations and (b) Bootstrapped estimates of the proportion MHS of sea perch catch-at-length from the trial. The thin vertical bars depict 95% bootstrapped confidence intervals for the length class specific relative selectivities. Circles depict bootstrapped mean relative selectivity. The smooth curve is a Bayesian spline fitted through the length-specific relative selectivities and the shaded region is generated from 95% pointwise credible intervals of the spline fit.

The point estimates for the proportions of sea perch taken by MHS and conventional gear that were less than 22 cm are 0.233 and 0.362 respectively. The test statistic to assess the null hypothesis of a zero difference in the sub-22 cm proportions of sea perch caught by the two gear types is -2.29, which indicates a significant difference between the gears ($p = 2.2\%$). The parametric 95% confidence interval (-0.239, -0.019) is strictly negative (Figure 19a).

Bootstrapped 95% confidence intervals for the difference in the sea perch proportion below 22 cm ranges between MHS catching 78 percent less small sea perch and 10 percent more. However, the bootstrap procedure applied to the sea perch length frequency data had to be modified so that an MHS sample was taken in at least one of the seven tow-pairs sampled with replacement in each bootstrap resample. Intuitively, the four MHS tows that were sampled is simply too few to be confident that the distribution of catch length can be reliably modelled.

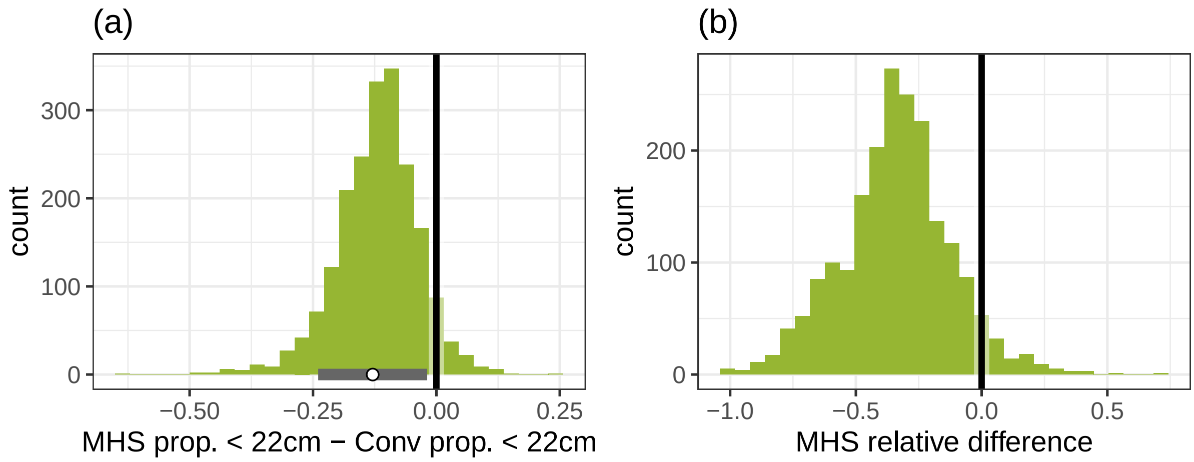


Figure 19: (a) Difference in the proportion of sea perch caught by MHS less than the 22 cm and (b) difference in the proportion of sea perch caught by MHS less than 22 cm relative to the conventional net proportion. Estimates around the vertical lines are consistent with no difference. The open circle and horizontal bar on panel (a) give the parametric mean and 95% confidence intervals for the difference, respectively.

4. DISCUSSION

The alternate-tow design adopted in the trial has been shown to be a useful approach for net A vs. net B trials. Despite a modest sample size (11 tow-pairs), data collected during the trial were sufficient to demonstrate a statistically significant difference between gear types in barracouta catch rate and length-based tarakihi selectivity.

Analyses provide strong evidence that the conventional trawl net used achieves higher catch rates of barracouta than the tested MHS net. There was no evidence that the first tow of a tow-pair has a higher expected catch rate of barracouta than the second tow, although the trial provides insufficient data to reliably test this possibility. The Bayesian credible interval for the effect of MHS on catch rates accounts for uncertainty due to tow order.

Observations of in-trawl video recorded during the trial included footage of barracouta freely escaping from the rectangular holes in the MHS (Hamill 2019).

It follows that more trawling using the MHS net would be required to catch barracouta compared with the mesh net tested, implying greater benthic impacts if fishers used MHS to target barracouta (all else being equal).

Smaller differences in catch rate will also be of interest, but would require more data to estimate precisely. The proportional catch-at-length analyses indicated that MHS caught more tarakihi for all length classes above 22 cm. However, exploratory GLMMs fitted to catch weights of TAR (not included in the report) did not find statistically significant differences in TAR catch rates between gears. Catch rate data from multiple trials may be required to attain precise estimates of the effect of gear type on catch rate of some species, especially if differences are small.

The analysis of the proportional catch-at-length (or relative selectivity) has a number of advantages over the binary selectivity comparisons, making fuller use of the collected data and – more importantly – providing information on the relative performance of the two gears that are not dependent upon the size distribution of fish encountered in the trial.

Comparisons of binary measures such as the proportion of catch below MLS are expected to vary somewhat among areas depending on the local size distribution of fish (Hurst et al. 2017, McKenzie & Millar 2018), whereas the relative selectivities estimated in one area are characteristics of the gear and can be expected to apply in other areas as well.

Jones & Millar (2018) suggest that the use of a fine mesh net in paired gear trials provides greater statistical power to detect differences in selectivity than an alternate-tow approach. However, analyses of length frequency data collected from the trial were sufficient to allow estimates of the proportional catch-at-length for tarakihi and red gurnard that were reasonably precise over well-sampled length classes.

Both parametric and bootstrap estimators of the proportion of tarakihi below 25 cm indicated that MHS caught a significantly lower proportion of TAX than the conventional mesh net. Analyses of the proportional catch-at-length suggest MHS is likely to retain markedly lower proportions of tarakihi smaller than 20 cm. For length classes between 20 and 25 cm the difference in the proportions retained by the two gear types became smaller. Analyses comparing the lengths of tarakihi sampled by observers in North Island inshore fisheries suggested fishers using the same MHS net caught smaller proportions of sub-MLS tarakihi than fishers using conventional trawl nets (Chambers 2019).

No evidence was found of a difference in gurnard selectivity between the two nets. The two nets appeared to have very similar selectivities of gurnard between about 33 and 43 cm. Comparisons of gurnard catch-at-length from areas where small gurnard are caught would be more informative for assessing whether MHS was no worse than conventional gear in terms of gurnard size composition. However, potential differences in selectivity between gear types of small gurnard will be unimportant if fishing only takes place in areas inhabited by larger gurnard.

Length measurements of sea perch catch indicate that conventional mesh retains a proportion of sea perch below the length at 50 percent sexual maturity. However, sampling of sea perch was not sufficient to allow a reasonable comparison of relative selectivity between the gear types.

Catch of protected species is rare and sporadic and, except in extreme cases, is unlikely to be resolved sufficiently precisely from small scale trials to permit a meaningful comparison of gear types.

We conclude that the experimental design and analysis approach illustrated in this trial is suitable for adoption as the future basis for net A vs. net B comparisons. Designs for future trials should be clear about the key metrics to be tested (i.e., catch rates, size composition, etc.), while recognising that comprehensive data collection may facilitate additional comparisons.

There are a number of improvements that can be made in the implementation of future trials:

- trial data should ideally be managed in standardised databases rather than in spreadsheets;
- sampling instructions should encourage sampling from ungraded catches before sorting, rather than after the fish have been sorted by size;
- measurements of fish length should follow the conventional practice of recording values to the nearest centimetre below the measured length, rather than the nearest centimetre;
- there should be a clearer rationale for the selection of species for length measurement, with clear instructions around whether a species needs to be measured from all tows where it is caught and when sampling from larger catches is acceptable.

Sampled tows should be made in areas where the range of size classes available to the trawl gear includes size classes for which the selectivity is of interest to managers. In the trial considered here, for instance, tarakihi below the minimum legal size were clearly available, allowing comparisons that suggest MHS may be more selective against undersized tarakihi than the conventional trawl nets tested.

In future trials it will be desirable that the intended target species is among the focal species for sampling. In this case, some tow-level catches may be too large for all individuals to be measured. Measuring a representative sample of tow catches is common in analyses of selectivity. Ideally, the same numbers of fish or the same fraction of the total catch should be sampled from each tow (Wileman et al. 1996). Alternate-tow studies should ensure as far as possible that each gear is the first gear fished on an equal number of tow-pairs.

It may be that opportunities for sampling particular species arise during the trial. There may be value in sampling these species to support informal analysis to assess future sampling opportunities. However, small-scale, haphazard sampling is unlikely to permit convincing analyses without the addition of data collected from a planned study.

5. ACKNOWLEDGEMENTS

The trial fishing was carried out by the FV *Ikawai*. Dave Woods (Precision Seafood Harvesting) provided the data and initial reports from the trial, and helpfully elaborated on a number of aspects. Analyses were improved by suggestions from the Statistics, Assessment and Methods Working Group. Additional advice provided by Russell Millar was particularly helpful.

6. REFERENCES

Breslow, N.E.; Clayton, D.G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association 88 (421)*: 9–25.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software 80 (1)*: 1–28. doi:10.18637/jss.v080.i01.

Chambers, M.S. (2019). Modelling lengths sampled by onboard observers. Presentation to the Statistics, Assessments and Methods Working Group, 26 February 2019. Fisheries New Zealand.

Dixon, P.M. (1993). The bootstrap and the jackknife: Describing the precision of ecological indices. *In*: S.M. Scheiner; J. Gurevitch (Eds.), Design and Analysis of Ecological Experiments (2nd ed., Chap. 14, pp. 267–288). Oxford University Press.

Efron, B. (1982). The jackknife, the bootstrap, and other resampling plans. *CBMS-NSF Regional Conference Series in Applied Mathematics 38*. 85 p.

Ernst, M.D. (2004). Permutation methods: A basis for exact inference. *Statistical Science 19 (4)*: 676–685.

Fisheries New Zealand (2018a). Approval for the use of a Precision Seafood Harvesting Modular Harvest System Trawl Net under the Fisheries (Commercial Fishing) Regulations 2001. Regulation 71A Approval. Retrieved from https://www.fisheries.govt.nz/dmsdocument/29003-precision-seafood-harvesting-modular-system-trawl-net-deepwater-fisheries-24-may-2018.

Fisheries New Zealand (2018b). Fisheries Assessment Plenary: stock assessments and stock status. Compiled by the Fisheries Science, Wellington, New Zealand.

Fisheries New Zealand (2019). Approval for the inshore use of a Precision Seafood Harvesting Modular Harvest System trawl net under regulation 71A of the Fisheries (Commercial Fishing) Regulations 2001. Retrieved from https://www.fisheries.govt.nz/dmsdocument/34527-precision-seafood-harvesting-modular-system-trawl-net-north-island-inshore-fisheries-12-april-2019.

Fleiss, J.L.; Levin, B.; Paik, M.C. (2003). Statistical Methods for Rates and Proportions (3rd ed.). John Wiley & Sons.

Hamill, J. (2019). Voyage Report IKA0936- 0939 FV San Ikawai 13-24 January 2019. (Unpublished report for Precision Seafood Harvesting). Plant & Food Research, Nelson, New Zealand.

Hastie, T.; Tibshirani, R.; Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). Springer Science+Business Media, New York.

Hurst, R.J.; Bagley, N.W.; O'Driscoll, R.L.; Millar, R.B. (2017). Sampling Protocols Hoki Fishery PSH Selectivity Trials. (Unpublished NIWA client report no. 2016163WN for Precision Seafood Harvesting). NIWA, Wellington, New Zealand.

Huse, I.; Løkkeborg, S.; Soldal, A.V. (2000). Relative selectivity in trawl, longline and gillnet fisheries for cod and haddock. *ICES Journal of Marine Science 57 (4)*: 1271–1282.

Jones, E.; Millar, R. (2018). SNA 1 Fishery Selectivity Trials Report. (Unpublished NIWA client report no. 2017343WN). NIWA, Auckland, New Zealand.

Marley, J.; Wand, M. (2010). Non-standard semiparametric regression via BRugs. *Journal of Statistical Software 37 (5)*: 1–30.

McConnaughey, R.A.; Hiddink, J.G.; Jennings, S.; Pitcher, C.R.; Kaiser, M.J.; Suuronen, P.; Sciberras, M.; Rijnsdorp, A.D.; Collie, J.S.; Mazor, T.; O Amorosa, R.; Parma, A.M.; Hilborn, R. (2020). Choosing best practices for managing impacts of trawl fishing on seabed habitats and biota. *Fish and Fisheries 21*: 319–337.

McKenzie, J.; Millar, R. (2018). Implications of the Precision Seafood Harvesting 'Modular Harvest System' on snapper stock yield relative to standard trawl. (Unpublished NIWA client report no. 2017370AK). NIWA, Auckland, New Zealand.

Millar, R.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.

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. (2018). Equivalence analysis of UCK in MHS and conventional gear. (Unpublished report to Precision Seafood Harvesting). Retrieved from https://www.mpi.govt.nz/dmsdocument/34578/direct.

O'Driscoll, R.L.; Millar, R.B. (2017). Hoki fishery PSH selectivity trials: FV Rehua (REH1701). (Unpublished NIWA client report no. 2017209WN for Precision Seafood Harvesting). NIWA, Wellington, New Zealand.

Plummer, M. (2019). rjags: Bayesian Graphical Models using MCMC. R package version 4-10. Retrieved from https://CRAN.R-project.org/package=rjags.

Sistiaga, M.; Herrmann, B.; Larsen, R.B. (2009). Investigation of the paired-gear method in selectivity studies. *Fisheries Research 97 (3)*: 196–205.

Suuronen, P.; Chopin, F.; Glass, C.; Løkkeborg, S.; Matsushita, Y.; Queirolo, D.; Rihan, D. (2012). Low impact and fuel efficient fishing—looking beyond the horizon. *Fisheries Research 119*: 135–146.

Vehtari, A.; Gelman, A.; Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing 27 (5)*: 1413–1432.

Venables, W.N.; Ripley, B.D. (2002). Modern applied statistics with s. Springer, New York.

Wileman, D.A.; 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 No. 215*. 126 p.

APPENDIX A: BOOTSTRAP ANALYSIS OF BARRACOUTA CATCH RATES

The GLMM approach described in Section 2.3.1 makes parametric assumptions about the distribution of barracouta catch rates by MHS and conventional trawl catch conditional upon the explanatory variables in model 1. An alternative non-parametric bootstrapping (Efron 1982) was used to supplement the GLMM approach.

We considered the possibility of within-pair dependence in the barracouta catch rates by the two gear types and therefore we specified a bootstrap procedure that resampled tow-pairs with replacement, rather than individual tows or fish lengths. The sample pairwise correlation between barracouta catch rate (kg/hr) using MHS and conventional trawl across the 11 paired tows was approximately 0.49.

For each bootstrap replicate we sampled 12 tow-pairs with replacement (i.e., the unmatched conventional tow was included). The total barracouta catch and total trawl duration corresponding with the 12 sampled tows from each tow gear were summed separately and overall catch rates calculated for each gear type by their quotient. Where the unmatched tow-pair was sampled, values of zero were assumed for both the MHS catch and MHS tow duration.

Let τ_b be the 12 tow-pairs sampled on bootstrap resample b including duplicates and let C_{im} and C_{ic} be the estimated barracouta catch by MHS and conventional trawl, respectively, on tow-pair i. Let E_{im} and E_{ic}, respectively, be the duration in hours of the MHS and conventional trawl tows on tow-pair i. Then the estimated difference in catch rate on bootstrap resample b was calculated as:

$$\Delta CPUE_b^* = \frac{\sum_{i \in \tau_b} C_{im}^*}{\sum_{i \in \tau_b} E_{im}^*} - \frac{\sum_{i \in \tau_b} C_{ic}^*}{\sum_{i \in \tau_b} E_{ic}^*}. \tag{A-1}$$

The distribution of $\Delta CPUE_b^*$ over the full set of bootstrap resamples defines the distribution of the bootstrap estimator. The asterisk notation is conventionally used to denote a bootstrap resample (Venables & Ripley 2002). Additionally, the effect of the unpaired tow can be assessed from the jackknife after bootstrap diagnostic plot for observation 12 in Figure D-2. For comparison, the procedure was repeated using only the 11 complete tow-pairs (A – K) to ensure the inclusion of the unmatched tow did not seriously influence the results. It should also be noted that the bootstrapping procedure models the difference in average catch rates with the influence of individual tows weighted by tow duration, whereas the GLMM models the expected catch of each tow given gear type and tow-pair. Tow durations were generally similar so this is unlikely to make an important difference in the comparison.

The bootstrapping comparison of barracouta catch rates suggested the conventional trawl net can be expected to achieve average catch rates that are about 500 kg per hour higher than the MHS configuration tested (Figure A-1). This represents a catch rate approximately 60 percent lower than conventional trawl which is consistent with the GLMM estimate. A jackknife after bootstrap diagnostic plot for the barracouta catch data reveals that observation ‘8’ (tow-pair H) has the largest absolute jackknife value (Figure D-2). Although an absolute standardised jackknife value of around -2 is not particularly extreme, the diagnostic plot suggests this tow-pair increases the variance of the bootstrap estimator of the mean difference in catch. There are too few tow-pairs to conclude whether the tow-pair is aberrant. It does not seem particularly extreme and the most conservative approach would be to include this tow-pair in the analysis.

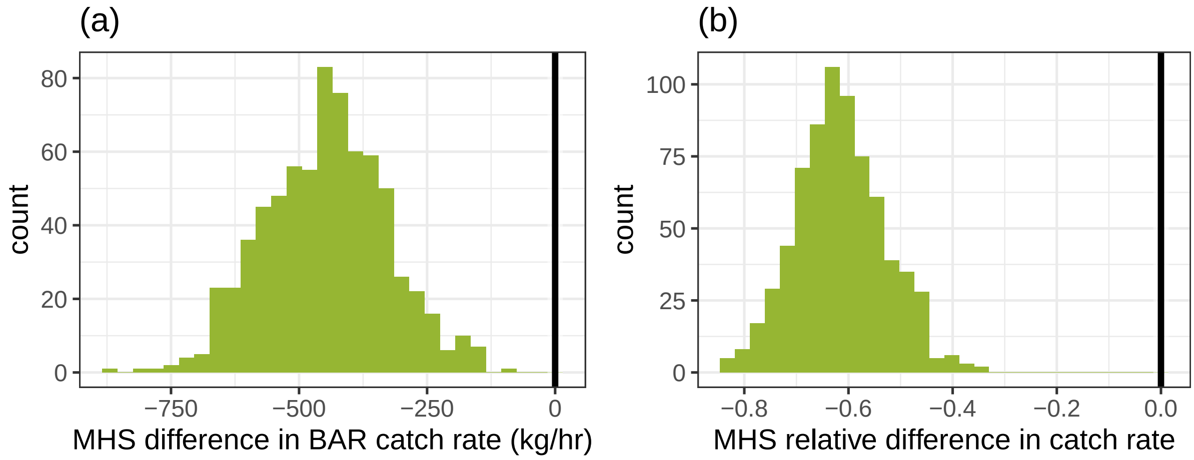


Figure A-1: Bootstrapped estimates of (a) the difference in the absolute BAR catch rate of MHS relative to conventional trawl and (b) the difference in the MHS BAR catch rate from conventional trawl as a proportion of the conventional trawl catch rate.

APPENDIX B: RANDOMISATION TEST OF THE PROPORTIONAL CATCH-AT-LENGTH

The null hypothesis that the MHS net and the conventional net had the same selectivity was tested using randomisation (see e.g., Ernst 2004). If the two gears have the same selectivities, the probability that a fish chosen at random from the combined sampled catch was caught on an MHS tow should be the same irrespective of its length class. Equivalently, if the selectivities are the same, the number of measured fish of length class l that were caught on MHS tows, N_{ml}, is a random variable with Binomial(N_l, π) distribution, where N_l is the total number of fish measured in length class l and π is a constant between zero and one. For each species, we fitted a null binomial generalised linear model (GLM) to the set of length classes l that were caught on the trial. If the two gears have different selectivities, the N_{ml} would be expected to be poorly fitted by the null model which would be reflected by a high model deviance.

Following O'Driscoll & Millar (2017), the randomisation procedure involved permuting the treatment labels (i.e., 'mhs' and 'conv') on the length frequency distributions measured from the two tows in each tow-pair. For each tow-pair there were two possible states where the observed catches were either (a) assigned to the correct gear type or (b) where the labels were switched. Therefore, in the case of tarakihi, for example, where length frequency distributions were available on nine tow-pairs there were $2^9 = 512$ permutations that could be compared. The p value of the randomisation test is the probability that the test statistic, T, derived from a random permutation of the labels is at least as extreme as the observed test statistic if the null hypothesis, H_0, were true. Mathematically,

$$p = P(T \geq t_{obs}|H_0) = \frac{\sum_{i=1}^{2^N} I(t_i \geq t_{obs})}{2^N}. \tag{B-1}$$

In Equation B-1, t_{obs} is the deviance of the null binomial model fitted to the measured length distributions, t_i is the deviance of the same model fitted to the length distributions resulting from permutation i and N is the number of tow-pairs. Note that the total number of fish in each length class, N_l, remains the same in each permutation. The $I(\cdot)$ notation denotes an indicator function equal to one when the condition inside the brackets is true and zero otherwise.

APPENDIX C: NEGATIVE BINOMIAL MODEL OF TARAKIHI SIZE COMPOSITION

Table C-1: Posterior summaries of negative binomial TAX per TAR model fixed effects coefficients.

	Estimate	Est.Error	Q2.5	Q97.5
Intercept	6.767	3.750	-0.509	14.262
gear = MHS	-1.053	0.583	-2.188	0.108
log(Total_catch) coeff.	-0.724	0.499	-1.707	0.259
first = yes	-0.430	0.546	-1.552	0.635

Table C-2: Posterior summaries of negative binomial TAX per TAR model shape parameter and random effects standard deviation.

	Estimate	Est.Error	Q2.5	Q97.5
sd(random tow pair effect)	0.915	0.523	0.070	2.099
shape parameter	1.511	0.782	0.528	3.616

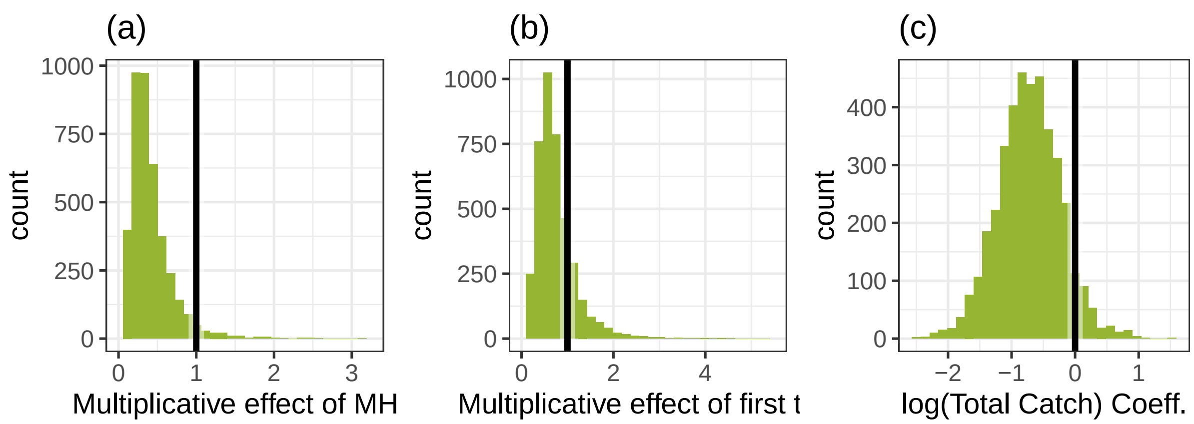


Figure C-1: Histograms of posterior distributions of (a) the multiplicative effect of MHS, (b) the multiplicative effect of first tow in tow-pair and (c) log (Total catch) coefficient.

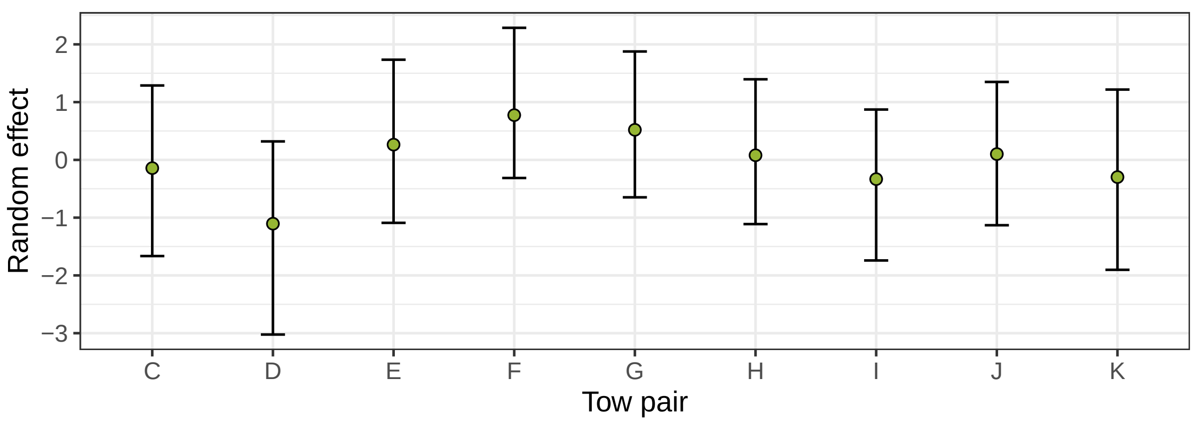


Figure C-2: Random tow-pair effects on TAX per TAR.

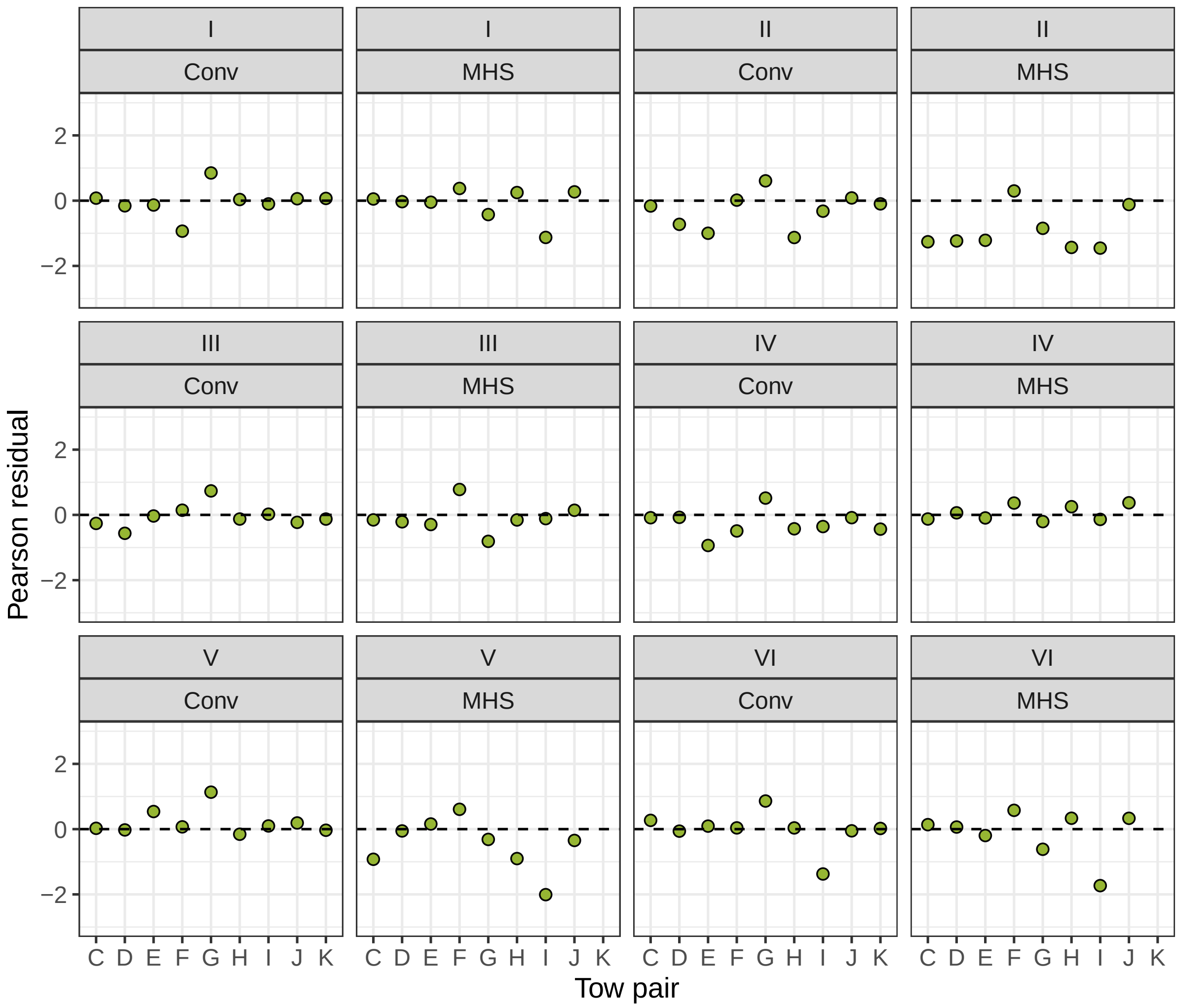


Figure C-3: Realised deviance residuals from GLMM for TAX per TAR by tow-pair and gear type from six random MCMC samples.

APPENDIX D: ADDITIONAL DIAGNOSTIC PLOTS

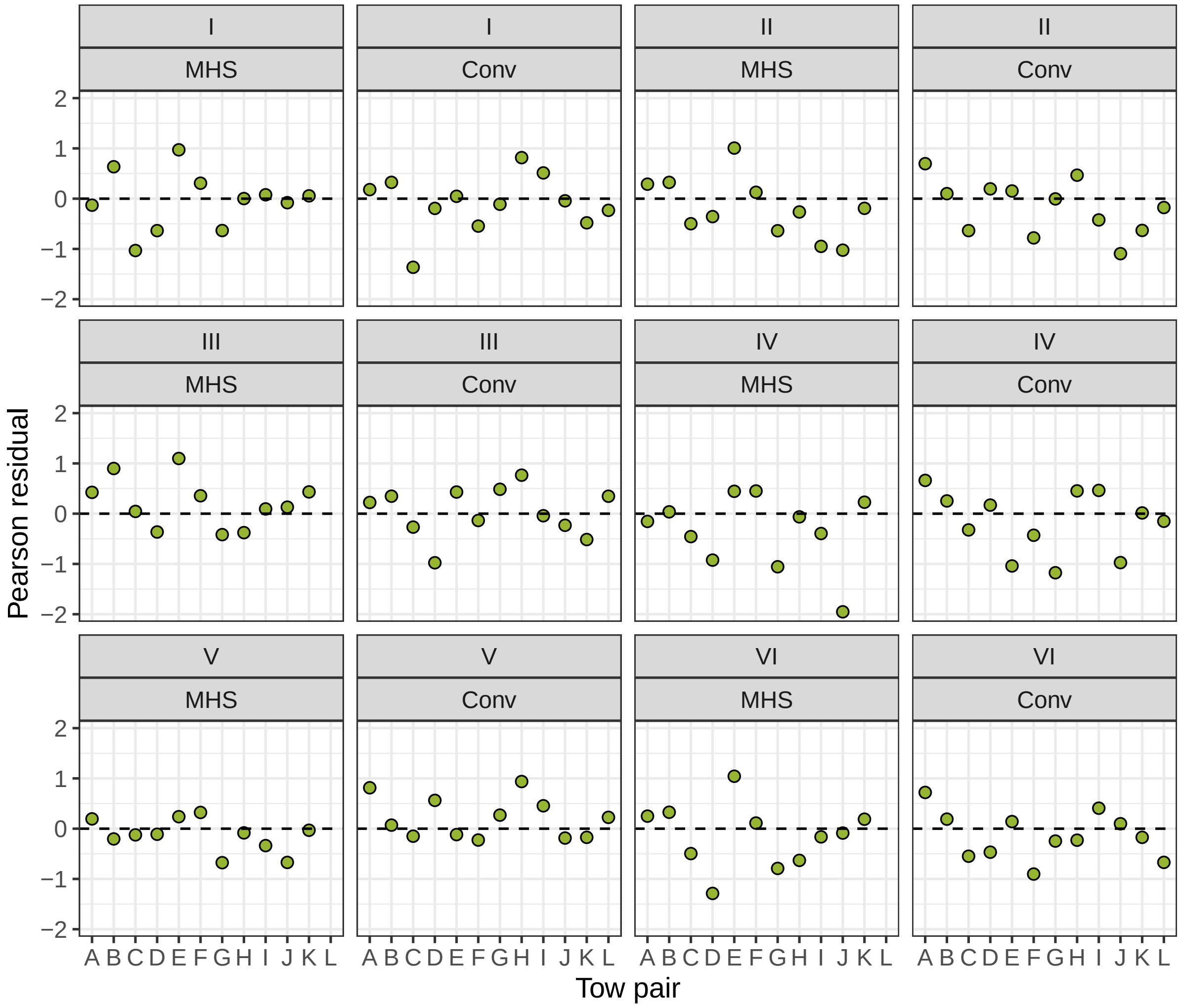


Figure D-1: Realised deviance residuals from GLMM for barracouta catch by tow-pair and gear type from six random MCMC samples.

The bootstrapping estimators used in this report all included resampling of tow-pairs. Jackknife after bootstrap diagnostic plots can be used to examine the influence of each tow-pair on the overall estimate. For each tow-pair, j, there is a subset, $B_{\neg j}$, of the bootstrap resamples in which tow-pair j is not selected. Let $T^*_{\neg j}$ be the bootstrapped estimates of the test statistic T based on the resample $B_{\neg j}$. The jackknife after bootstrap plot summarises the distributions $T^*_{\neg j}$ for each of the 11 tow-pairs. Percentiles of the centred distributions of $T^*_{\neg j} - t_{\neg j}$ are plotted against standardised jackknife influence values, where $t_{\neg j}$ is the arithmetic mean of $T^*_{\neg j}$. The standardised jackknife influence values are the 11 jackknife estimates of bootstrap bias, $t_{j0} - t_{\neg j}$, less the mean of the 11 bias estimates and divided by their standard deviation.

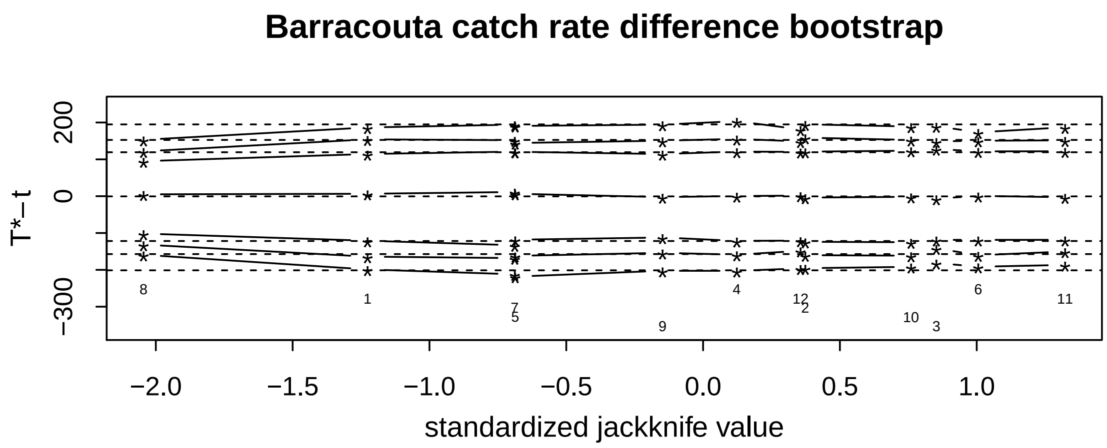


Figure D-2: Jackknife after bootstrap plot for barracouta catch rate. Points are 5th, 10th, 16th, 50th, 84th, 90th and 95th percentiles of jackknife distribution for each tow-pair of the bootstrap.

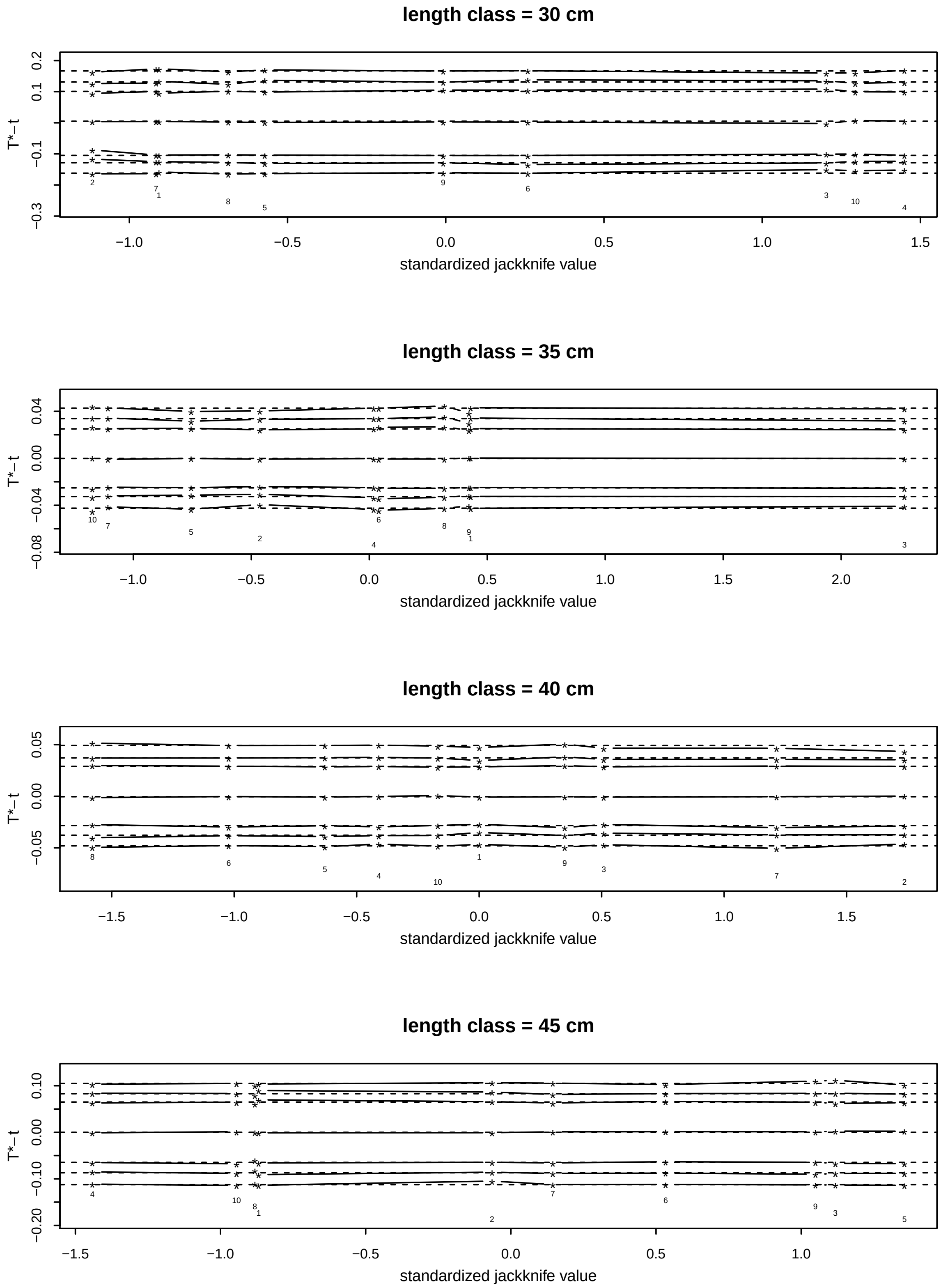
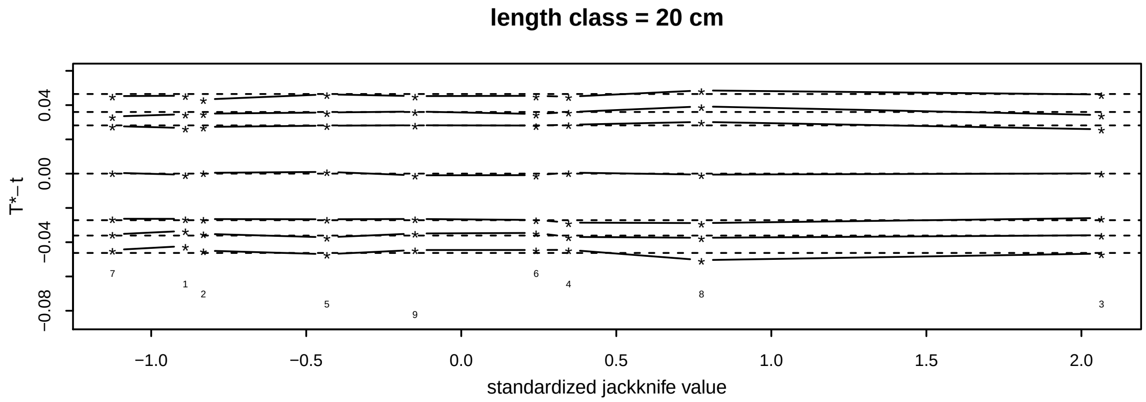
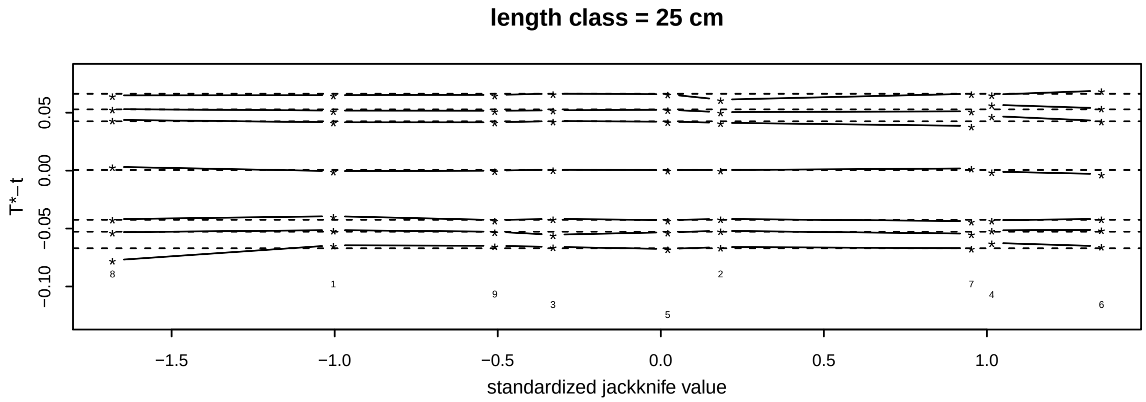


Figure D-3: Jackknife after bootstrap plot for relative gurnard selectivity at 30 cm, 35 cm, 40 cm and 45 cm length classes. Points are 5th, 10th, 16th, 50th, 84th, 90th and 95th percentiles of jackknife distribution for each tow-pair of the bootstrap.





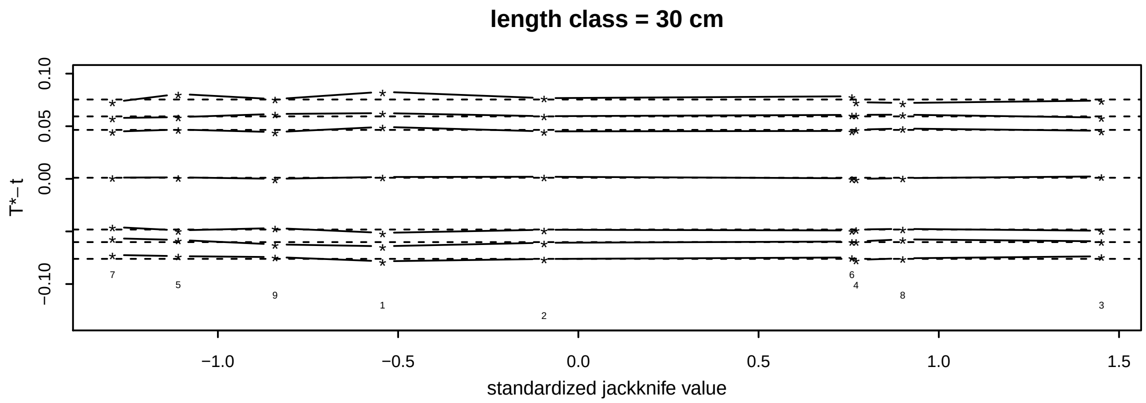


Figure D-4: Jackknife after bootstrap plot for relative tarakihi selectivity at 20 cm, 25 cm and 30 cm length classes. Points are 5th, 10th, 16th, 50th, 84th, 90th and 95th percentiles of jackknife distribution for each tow-pair of the bootstrap.

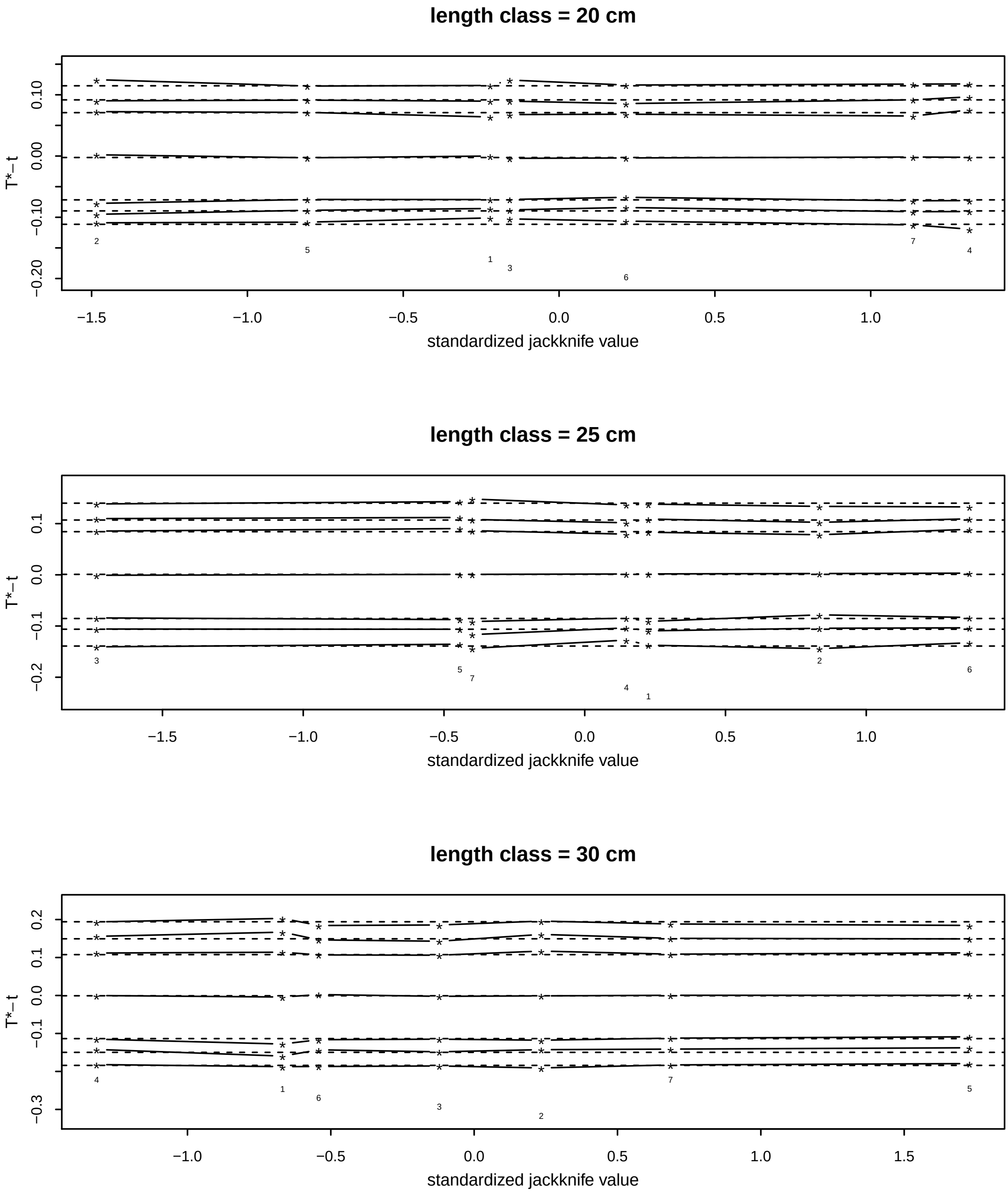


Figure D-5: Jackknife after bootstrap plot for relative sea perch selectivity at 20 cm, 25 cm and 30 cm length classes. Points are 5th, 10th, 16th, 50th, 84th, 90th and 95th percentiles of jackknife distribution for each tow-pair of the bootstrap.

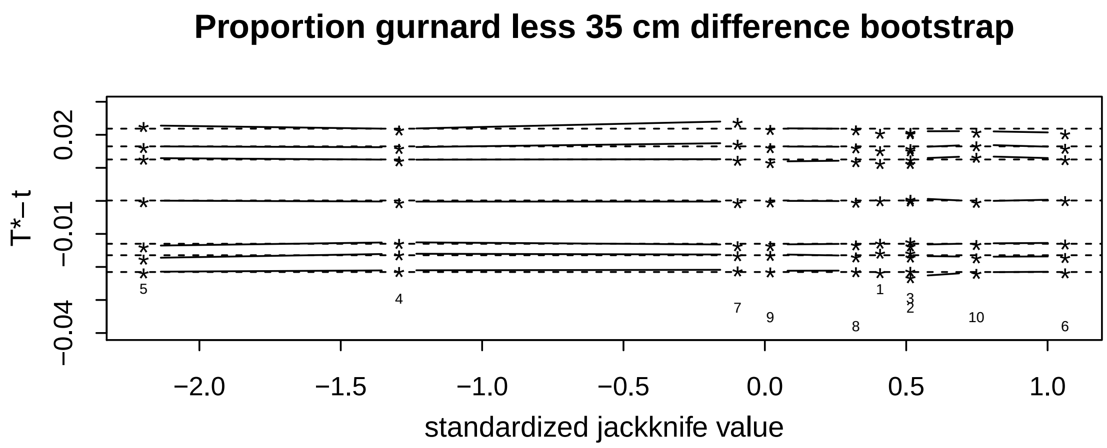


Figure D-6: Jackknife after bootstrap plot for difference in proportion of gurnard less than 35 cm. Points are 5th, 10th, 16th, 50th, 84th, 90th and 95th percentiles of jackknife distribution for each tow-pair of the bootstrap.

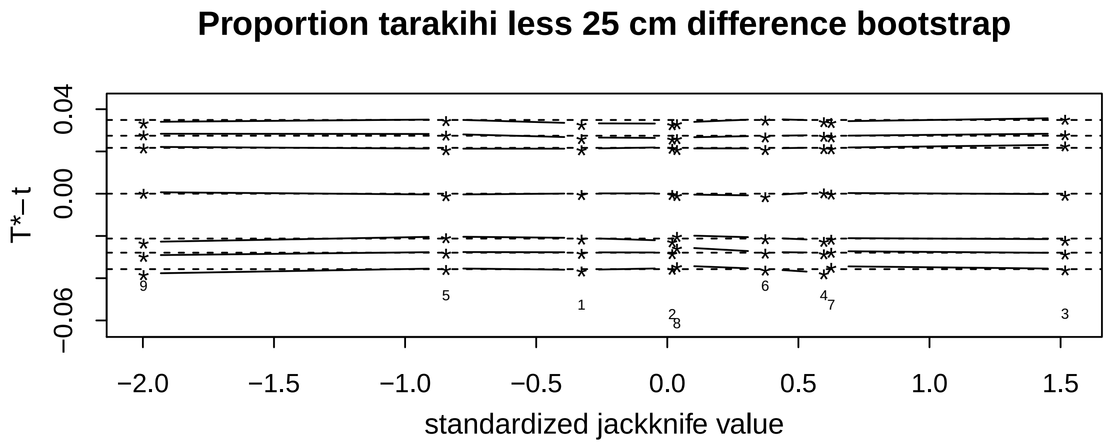


Figure D-7: Jackknife after bootstrap plot for difference in proportion of tarakihi less than 25 cm. Points are 5th, 10th, 16th, 50th, 84th, 90th and 95th percentiles of jackknife distribution for each tow-pair of the bootstrap.

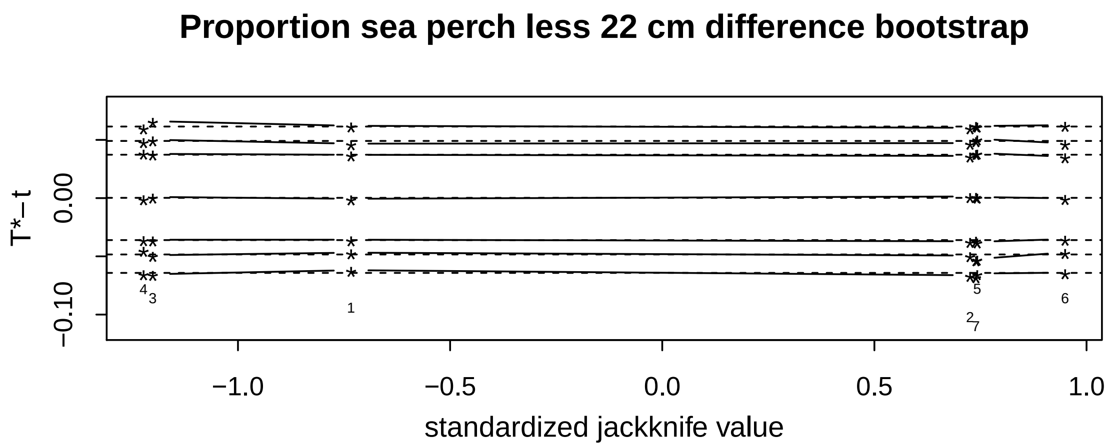


Figure D-8: Jackknife after bootstrap plot for difference in proportion of sea perch less than 22 cm. Points are 5th, 10th, 16th, 50th, 84th, 90th and 95th percentiles of jackknife distribution for each tow-pair of the bootstrap.

APPENDIX E: JAGS CODE TO FIT BAYESIAN MIXED-MODEL SPLINES

The following code fits a spline to length-specific bootstrapped proportions approximated as normal distributions. The `x[i]` in the script denotes the length of the ith fitted length class and `Z[i,]` denotes the O'Sullivan spline basis for that length given `numKnots` knots. The approach also uses R code available from Marley & Wand (2010) including R code to calculate O'Sullivan spline bases.

```
model{
   for(i in 1:N)
   {
mu[i] <- beta0 + beta1*x[i] + inprod(u[],Z[i,])
mean_diff[i] ~ dnorm(mu[i],tau_diff[i])
    }
   for(k in 1:numKnots)
   {
      u[k] ~ dnorm(0,tauU)
   }
beta0 ~ dnorm(0,1.0E-3)
beta1 ~ dnorm(0,1.0E-3)
  sigma_u ~ dunif(0,25)
tauU <- 1/(sigma_u*sigma_u)
   }
```

APPENDIX F: PARAMETRIC ESTIMATION OF DIFFERENCES IN BINARY SELECTIVITY

To take advantage of the alternate-tow design where all fish caught were measured we employ the methods described by Fleiss et al. (2003, §15.2) for the analysis of correlated binary data.

Following Fleiss et al. (2003), we define X_{ij} as a binary indicator variable that specifies whether the measured fish j captured on tow-pair i was taken with MHS gear:

$$X_{ij} = \begin{cases} 0 & \text{if fish } j \text{ from tow-pair } i \text{ was caught using conventional trawl,} \\ 1 & \text{if fish } j \text{ from tow-pair } i \text{ was caught using the modular harvesting system;} \end{cases} \quad \text{(F-1)}$$

and Y_{ij} as a binary indicator response variable indicating whether the fish was less than the MLS:

$$Y_{ij} = \begin{cases} 0 & \text{if fish } j \text{ from tow-pair } i \text{ had length} \geq \text{MLS ,} \\ 1 & \text{if fish } j \text{ from tow-pair } i \text{ had length} < \text{MLS .} \end{cases} \quad \text{(F-2)}$$

Expressed using these indicator variables, the total number of sub-MLS fish caught by MHS gear on tow-pair i is

$$Y_{im^+} = \sum_{j=1}^{n_i} X_{ij} Y_{ij} \quad \text{(F-3)}$$

and number of sub-MLS fish caught by conventional gear is

$$Y_{ic^+} = \sum_{j=1}^{n_i} (1 - X_{ij}) Y_{ij} \quad \text{(F-4)}$$

where n_i is the total number of fish measured from tow-pair i.

The total number of fish measured in the trial can be separated into four classes a, b, c and d:

	$< MLS$	$>= MLS$
MHS	a	b
BT	c	d

Expressed in terms of the indicator variables, for K tow-pairs:

$$a = \sum_{i=1}^{K} \sum_{j=1}^{n_i} X_{ij} Y_{ij}$$

$$b = \sum_{i=1}^{K} \sum_{j=1}^{n_i} X_{ij} (1 - Y_{ij})$$

$$c = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (1 - X_{ij}) Y_{ij}$$

$$d = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (1 - X_{ij})(1 - Y_{ij})$$

.

The probability a fish is less than the reference size (for simplicity, denoted here as the MLS) given it is caught using MHS gear is denoted by $P_m = P(Y = 1|X = 1)$ and the corresponding probability given the fish was captured using conventional gear is $P_c = P(Y = 1|X = 0)$.

The point estimates are $p_m = a/(a + b)$ and $p_c = c/(c + d)$, and the significance of the difference between these point estimates is assessed with the test statistic, z:

$$z = \frac{p_m - p_c}{\sqrt{\mathrm{Var}(p_m - p_c)}}$$

where the variance is calculated as:

$$\mathrm{Var}(p_m - p_c) = \frac{P_m Q_m f_m}{a + b} + \frac{P_c Q_c f_c}{c + d} - \frac{2\,\mathrm{Cov}(a, c)}{(a + b)(c + d)} \qquad \text{(F-5)}$$

where Q_x is shorthand for $1 - P_x$ and the variance inflation factors for the groups of tow-pairs with MHS (m) and conventional (c) gear are:

$$f_x = 1 + \left\{ \frac{s_x^2}{\bar{n}_x} + (\bar{n}_x - 1) \right\} \rho_x$$

with

$$\bar{n}_m = \frac{\sum_{i=1}^{K} X_i n_i}{\sum_{i=1}^{K} X_i}$$

$$s_m^2 = \frac{\sum_{i=1}^{K} X_i (n_i - \bar{n}_i)^2}{\sum_{i=1}^{K} X_i}$$

$$\rho_m = \sum_{i=1}^{K} \left\{ Y_{im^+}(Y_{im^+} - 1) - 2p_m(n_i - 1)Y_{im^+} + n_i(n_i - 1)p_m^2 \right\}$$

for the MHS tow-pairs, and with the corresponding quantities for the conventional tow-pairs obtained by replacing X_i with $(1 - X_i)$, Y_{im^+} with Y_{ic^+}, and p_m with p_c.

The covariance between a and c given X_{ij} is given by:

$$\mathrm{Cov}\,(a, c) = \rho_{mc} \sqrt{P_m Q_m \, P_c Q_c} \sum_{i=1}^{K} k_i(n_i - k_i) \qquad \text{(F-6)}$$

where k_i is the number of fish caught by MHS gear on tow-pair i.

A consistent estimator of the within-pair correlation between observed binary size class frequencies caught by MHS and conventional trawl is given by:

$$\hat{\rho}_{mc} = \frac{\sum_{i=1}^{K} \{Y_{im^+} - k_i p_m\} \{Y_{ic^+} - (n_i - k_i)\, p_c\}}{\sum_{i=1}^{K} k_i (n_i - k_i) \sqrt{p_m q_m \, p_c q_c}} \tag{F-7}$$

.