Review of the Atlantic halibut longline survey index of exploitable biomass

Stephen J. Smith

Science Branch, Maritimes Region
Fisheries and Oceans Canada
Bedford Institute of Oceanography
1 Challenger Drive
Dartmouth, Nova Scotia
B2Y 4A2

2016

Canadian Technical Report of
Fisheries and Aquatic Sciences 3180

Fisheries and Oceans Canada
Pêches et Océans Canada
Canada

Canadian Technical Report of Fisheries and Aquatic Sciences

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature. Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

Rapport technique canadien des sciences halieutiques et aquatiques

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. II n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques.*

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 456 de cette série ont été publiés à titre de Rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de Rapports techniques de la Direction générale de la recherche et du développement, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de Rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Canadian Technical Report of Fisheries and Aquatic Sciences 3180

2016

Review of the Atlantic Halibut longline survey index of exploitable biomass

by

Stephen J. Smith[1]

Science Branch, Maritimes Region
Fisheries and Oceans Canada
Bedford Institute of Oceanography
1 Challenger Drive
PO Box 1006
Dartmouth, NS
B2Y 4A2

[1] 383 Portland Hills Drive, Dartmouth, Nova Scotia, B2W 6R4. Smithsj@eastlink.ca

Cat. No. Fs 97-6/3180E-PDF ISBN 978-0-660-06219-8 ISSN 1488-5379

Correct citation for this publication:

Smith, S. J. 2016. Review of the Atlantic Halibut longline survey index of exploitable biomass. Can. Tech. Rep. Aquat. Sci. 3180: v + 56 p

TABLE OF CONTENTS

TABLES

FIGURES

ABSTRACT

Smith, S. J. 2016. Review of the Atlantic Halibut longline survey index of exploitable biomass. Can. Tech. Rep. Fish. Aquat. Sci. 3180: v + 56 p.

This report contains the results of the review commissioned by The Atlantic Halibut Council and the Department of Fisheries and Oceans Canada (DFO) of the joint Industry/DFO annual longline survey. This survey has been used to monitor Atlantic halibut exploitable biomass since 1998. The survey was originally stratified based on areas of Low, Medium and High catch based on data from commercial fishing logs (1995–1997). Starting in 2009, a standardized catch rate calculated from a Negative Binomial generalized linear model replaced the stratified estimate of mean weight per standard set. The review recommended replacing the Negative Binomial model with a Multinomial model that not only accounts for the number of halibut caught but also the number of hooks occupied by other species or missing bait. It was not possible to fully assess the impact of covariates collected during the survey because of the observational nature of the data, but based on literature review standardization of protocols to reduce the variation in covariates is recommended. A new sample-to-strata allocation plan was presented that has potential to improve the precision of the estimates.

RÉSUMÉ

Smith, S. J. 2016. Review of the Atlantic Halibut longline survey. Can. Tech. Rep. Fish. Aquat. Sci. 3180: v + 56 p.

Le présent rapport contient les résultats de l'examen commandé par l'Atlantic Halibut Council et Pêches et Océans Canada (MPO) concernant le relevé à la palangre annuel réalisé conjointement par l'industrie et le MPO. Ce relevé est utilisé pour surveiller la biomasse exploitable du flétan de l'Atlantiqué depuis 1998. Il était à l'origine stratifié selon des zone aux taux de prise faibles, moyens et élevés, d'après le données de journaux de pêche commerciale (datant de 1995 à 1997). À compter de 2009, un taux de prise normalisé calculé à partir d'un modèle linéaire généralisé utilisant une distribution de l'erreur binomiale négative a remplacé les estimations stratifiées du poids moyen par trait standard. L'examen a recommandé de remplacer le modèle binomial négatif par un modèle multinomial qui tient compte non seulement du nombre de flétans capturés, mais aussi du nombre d'hameçons occupés par d'autres espèces ou sans appât. Il n'a pas été possible d'évaluer pleinement l'incidence des covariables recueillies au cours du relevé, car les données reposaient sur des observations; toutefois, selon l'analyse documentaire, la normalisation des protocoles est recommandée afin de réduire la variation des covariables. Un nouveau plan de répartition des échantillons par strates susceptible d'améliorer la précision des estimations a été présenté.

INTRODUCTION

Annual surveys are used to monitor stock status and annual changes for many commercial exploited marine species throughout the world (Gunderson 1993). Most of these surveys use fishing gear that has been standardized with respect to construction, operation and sampling protocol. Longline gear has been used to survey a number of demersal species such as Pacific halibut (Pelletier and Parma 1994; Soderlund et al. 2012), Greenland halibut (Nygaard 2014), and sablefish (Sigler 2000). There has been extensive research reported in primary journals and agency reports on models for interpreting longline catch rates (e.g., Rothschild 1967; Ricker 1975; Somerton and Kikkawa 1995; Kimura and Zenger 1997; Baum and Blanchard 2010; Etienne et al. 2013), factors affecting longline catches through field surveys (e.g., Hamley and Skud 1978; Skud 1978a,b; Bjordal 1983; Hovgård and Lassen 2000; Rodgveller et al. 2008; Soderlund et al. 2012), use of hook timers (Somerton et al. 1988; Somerton and Kikkawa 1995; Sigler 2000) and underwater observation using submersibles or cameras (High 1980; He 1996; Kaimmer 1999).

The Atlantic halibut longline survey was implemented in 1998 to provide annual indices of abundance and biomass as well as information on size composition, diet, tagging/movement and bycatch species for stock assessment purposes (Zwanenburg and Wilson 2000). Previously, abundance indices had been derived from the Department of Fisheries and Oceans (DFO) annual research groundfish trawl surveys but catches of commercial size halibut (fork length $>$ 81 cm) were quite low in the trawl gear and this gear was considered inadequate to provide enough information to monitor the stock. Longline gear is the main gear used by the halibut fishery and this gear was chosen for the survey as it was expected to be more successful than trawl gear for capturing commercial size halibut. The original design of the survey was defined to be stratified random in Zwanenburg and Wilson (2000). Strata were defined based on maps of fishing effort or catch for the 1995 to 1997 period and survey stations were fixed locations to be consistently fished each year. The original design called for the number of stations to be proportionally allocated as 23%, 32% and 45% (ratio of 5:7:10) of the total number of planned stations to three strata designated as Low, Medium and High catch areas, respectively (Zwanenburg and Wilson 2000; den Heyer et al. 2015). Stratified estimates were used for the survey mean numbers and weights of halibut caught (e.g., Armsworthy et al. 2006) until the assessment by Trzcinski et al. (2009) when the stratification system was no longer used for the estimates although the strata were still part of the survey design. Beginning with the 2008 fishing year, a generalized linear model has been used to provide standardized survey catch rates of mean halibut weight per standard longline set (Trzcinski et al. 2009; den Heyer et al. 2015).

The Atlantic halibut longline survey index based on the generalized linear model is one of two fishery-independent indices used in both the 2010 and 2014 assessment models and the current stock assessment model developed for a recent Canadian Scientific Advisory Secretariat (CSAS) Halibut framework meeting (Cox et al. 2016). Full assessments are currently planned to be conducted every five years and the longline index will be used to monitor the stock and provide interim advice during the period between full assessments. Given the importance of this survey to the management of the halibut fishery, the Atlantic Halibut Council and DFO sponsored a review of the design and models used for the survey. Details on the five terms of reference are provided in Appendix 1. This report presents the results of this review.

TOR 1. EVALUATE HALIBUT CATCH RATE MODELS

CURRENT APPROACH

The Atlantic halibut longline survey was designed to provide an annual index of abundance (numbers or weights) to monitor the status of the stock for management purposes. It was assumed that for a standard soak time, the number of halibut caught in any one set would be proportional to the number of halibut available in the immediate area of the set. That is, as the numerical abundance of halibut increased in an area, the proportion of hooks in a set containing halibut would also increase. This proportion reflects the catch rate of the gear in terms of the expected number of hooks containing halibut. Assuming that all longline sets represent the density of halibut over a similar area, the individual catch rates could be combined in some way to reflect the overall abundance in the management areas. The survey protocol states that sets will consist of 1000 hooks with soak times of 600 minutes, but operational factors lead to variation from these standards (Figure 15, den Heyer et al. 2015). The initial estimates of the index used for this survey adjusted the halibut catch from set i in stratum h, $y_{h,i}$ to these protocols using the observed soak time[1], $S_{h,i}$ and the number of hooks $N_{h,i}$ (Armsworthy et al. 2006).

$$y^*_{h,i} = y_{h,i}\left(\frac{1000}{N_{h,i}}\right)\left(\frac{600}{S_{h,i}}\right) \tag{1}$$

The mean $\bar{y}^*_h$ provides a simple estimate of the catch rate in stratum h in terms of the expected number of halibut caught for the standard set of 1000 hooks and 600 minutes soak time. The stratified mean represents the catch rate for whole survey,

[1] The observed soak time denoted as `DURATION` in the halibut survey database is defined as the time elapsed between the time the gear starts to be deployed and the time the gear has been fully recovered and therefore refers to the maximum soak time. Maximum soak time with adjustments for split sets denoted as `avg_DURATION` was used in this study.

$$\bar{y}^* = \sum_{h=1}^{3} \bar{y}_h^* \frac{A_h}{\sum_{h=1}^{3} A_h} \tag{2}$$

where A_h is the area of catch stratum h. This estimate also assumes that catch will increase in direct proportion to increases in soak time and total numbers of hooks. Catch rates for both numbers and weights using equation 2 were presented in Armsworthy et al. (2006).

Armsworthy et al. (2006) also presented a standardized catch rate index obtained from fitting a Negative Binomial Generalized Linear Model (NB GLM) to survey catch weight with years and stations as fixed effects factor levels. Linear models have been used for a number of fisheries survey and commercial catch rate series to remove the effects of annual changes in the survey or fishery that may not have anything to do with a change in abundance (Maunder and Punt 2004). Changes in season, time of day or depth being fished are examples where catchability may change without a concurrent change in abundance. The usual approach is to model the influential variables as fixed effects (e.g., Month or depth intervals) and choose a standard level (e.g., January or 50–60 m) to predict a standardized catch rate for each year. Interaction terms between main effects are usually avoided because they are difficult to interpret and often random effects terms are used instead (Maunder and Punt 2004).

One of the main assumptions of the fixed effects approach is that on the linear predictor scale, the response for each month or depth interval is higher or lower than other months and depth intervals but parallel over time. Sampling of the fixed stations has been variable over the history of this survey and all stations fished from 1998 to 2006 were included in the model presented in Armsworthy et al. (2006) as a factor term with 261 levels. This model assumes that station and year as factor terms have independent effects on the observed catch rate. That is, stations are assumed to have consistently higher, lower or similar catch rates relative to the other stations over all years with year-to-year differences for all stations captured by the year term. However, of the 261 stations used in the analysis there were only 72 that were reported as being consistently fished each year. Stations that were only fished in a few of the years may have undue influence on the estimates and may also contribute to numerical problems when estimating the parameters of the model.

Few details on the Negative Binomial model were provided in Armsworthy et al. (2006) and it is not clear if the model was fitted to the raw weight, adjusted weight caught per set or to raw weight divided by soak time and total hook number. Depending upon the station chosen as the standard the resulting index may be higher or lower or near the index calculated using the stratified mean above. The GLM index presented in Figure 7 of Armsworthy et al. (2006) is more than two standard errors below the simple mean estimate suggesting that the index was actually predicted on the scale of the linear predictor (logarithmic scale in this case) and not in the original scale of the response (weight/set).

Trzcinski et al. (2009) applied the Negative Binomial model to the weight of halibut caught per set including an offset term for the logarithm of the number of hooks for each set to account for the varying number hooks by set. They also only used stations that had been sampled in the survey four or more years to mitigate against possible numerical issues identified above. This same model was used for the most recent stock assessment and limited to those sets with soak times greater 180 minutes and more than 500 hooks (den Heyer et al. 2015). In both cases the predicted series was calculated for the same station and the offset term set to 1000 hooks, however soak time was assumed to be 10 hours for all sets (C. Den Heyer, pers. comm.).

The analyses conducted in the three stock assessments (Armsworthy et al. 2006; Trzcinski et al. 2009; den Heyer et al. 2015) were compared with results obtained using the most recent version of the data to see if they could be repeated and to evaluate the individual analyses. The analysis of deviance reports from the new analyses were compared with those presented in the stock assessment documents using the same specifications given in the original stock assessments. Revisions and edits to the database probably contribute to the some of the differences seen between the published results and those presented here (Table 1). Overall, the results were quite similar except for the estimate of θ in Armsworthy et al. (2006) which was probably due to a typographical error in the original document. Predicted indices for the same set indicate similar trends for the three series, however the series are not coincident (Figure 1). This can happen due to data edits and the different data specifications used in the assessments with respect to the number of years stations have to be fished to be included in the model as well as limits on soak time and total hook number.

Refitting the models uncovered another problem with the GLM approach used in all three assessments. The model can be written as follows.

$$E[y_i] = \exp(\text{Station} + \text{Year} + \text{Offset(Hooks)}) \tag{3}$$

It should make no difference if the model is set to the above or to Year+Station+ Offset(Hooks) as done in Table 1 and in the original assessments. Year and Station are assumed to be separate effects. However, when the order was reversed from that in the table, Year was a significant effect for the first two stock assessments, the deviance results were quite different and the software reported that the fitting algorithm did not converge. All of these indicators point to a confounding between Year and Station and subsequent numerical problems.

The Negative Binomial model was originally defined for count data (e.g., 0, 1, 2,...) representing the number of specific events that occurs, such as the number of fish caught. Weight data does not correspond to a unique set of events such as numbers of hooks with halibut on them because a catch of 5 kg could represent one fish or five fish at 1 kg each or some other combination. Weights and other continuous random variables are usually recorded as decimal data and any attempt to fit these kind of data for the response variable in the current version of the R software function `GLM.nb` that was used in all three stock assessments would result in warnings that the response variable was non-integer. However, the halibut and total catch weight data supplied from the survey database were rounded to kilograms and would appear as count data to the software.

Generalized linear models are useful for extracting annual trends from catch rate data separately from other factors that may be affecting the catchability of the fish to the survey gear. However, the use of the NB GLM model index for the halibut longline survey should be discontinued. The Negative Binomial model was not designed for continuous response variables and would not be expected to provide a proper description of the probability distribution, variance, confidence intervals or goodness of fit. Even if a distribution appropriate for weight or for numbers per standard set was used, the current model structure of using station in a two-way factorial model (station+year) where not all stations were fished in each year is just asking for numerical problems.

ALTERNATIVE APPROACH

The simple stratified mean adjusted catch rate, the NB GLM discussed above and even the application of either a Negative Binomial or a Binomial distribution to numbers caught all assume that halibut are the only species being caught by the longline hooks with no accounting for other species competing for hooks. In addition, all of these methods implicitly assume that all of the hooks that do not have halibut on them would still be able to catch halibut had there been more of these fish in the area. A number of hooks will be occupied by species other than halibut and other hooks will be empty with bait still attached or missing. Incorporation of the catch of other species as a covariate is a possible way to model the effect of bycatch on the catch of halibut but misses the point that other species compete with halibut for hooks as a function of their relative densities.

Alternative models have been proposed for the underlying process of the interaction between fish and longline gear that account for the competition between target and non-target species for baited hooks as well as the treatment of empty hooks (Rothschild 1967; Ricker 1975; Somerton and Kikkawa 1995; Etienne et al. 2013). These models have demonstrated that the catch of non-target species and the number of empty hooks can affect the behaviour of the catch rate estimate for the target and non-target species. The models are based on the following possible states for any one hook. Assuming that events for any one hook are independent from any other hook for a single longline with N total hooks, the possible states for hook j after soak time S upon retrieval onboard are,

1. Empty hook: n_0 = total number of empty hooks
   (a) baited: $n_{0,b}$ = total number of empty baited hooks
   (b) not baited: $n_{0,u}$ = total number of empty unbaited hooks
2. Halibut: n_1 = total number of hooks with halibut on them
3. Other species of fish: n_2 = total number of hooks with other species on them

Note that $n_0 = n_{0,b} + n_{0,u}$ and the total number of hooks in a set are $N = \sum_{k=0}^{2} n_i$.

The final state for each hook could be set at any time during the time the gear was deployed. The following model was originally developed by Rothschild (1967) but the formulation of Etienne et al. (2013) will be used here because it is more straightforward. Denote T_1 as the time it takes for a target species (e.g., halibut) to be caught on a hook. The probability that a halibut will be caught before time u is obtained from an exponential distribution for T_1 with rate λ_1,

$$P(T_1 < u) = 1 - \exp(-\lambda_1 u) \tag{4}$$

If halibut is the only species caught by the longline survey then the distribution of the n_1 halibut caught for N hooks and soak time S follows a Binomial distribution with the probability of catching a halibut equal to $(1 - \exp(-\lambda_1 S))$.

$$P(n_1; \lambda_1, N) = \binom{N}{n_1} (1 - \exp(-\lambda_1 S))^{n_1} (\exp(-\lambda_1 S))^{n_0} \tag{5}$$

The probability of capturing a halibut is also the catch rate because $n_1 = (1 - \exp(-\lambda_1 S)) \times N$. In turn catch rate is generally assumed to reflect density of the halibut within the area that they can detect the longline gear. However, other species of fish are also caught on the longline and they need to be included. The probability that another species of fish is caught on the hook can also be expressed using to the exponential distribution.

$$P(T_2 < u) = 1 - \exp(-\lambda_2 u) \tag{6}$$

The probability distribution of the time for either a halibut or another species of fish is caught, $T = \min\{T_1, T_2\}$ will also be exponential with rate $\lambda = \lambda_1 + \lambda_2$. If $T_1 < T_2$ then $\lambda_1 > \lambda_2$ indicating that we would expect more halibut to be caught by the longline gear than the other species.

The probability of an empty hook is

$$P(T \geq u) = \exp(-\lambda u) \tag{7}$$

However, this empty hook could be baited or unbaited and lacking direct observation of how a bait was removed without catching a fish, assumptions will need to be made to calculate the probabilities of either event. For now assume that all of the empty hooks are baited ($n_{0,u} = 0$).

The joint likelihood for hooks being empty or not empty and if not empty containing a halibut or another species can be expressed as,

$$\begin{aligned} L(\lambda_1, \lambda_2) &= \binom{N}{n_0}\binom{n_1+n_2}{n_1}(\exp(-\lambda S))^{n_0}(1-\exp(-\lambda S))^{n_1+n_2}\left(\frac{\lambda_1}{\lambda}\right)^{n_1}\left(\frac{\lambda_2}{\lambda}\right)^{n_2} \\ &= \frac{N!}{n_0!n_1!n_2!}(\exp(-\lambda S))^{n_0}\left((1-\exp(-\lambda S))\left(\frac{\lambda_1}{\lambda}\right)\right)^{n_1} \\ &\quad \times\left((1-\exp(-\lambda S))\left(\frac{\lambda_2}{\lambda}\right)\right)^{n_2} \end{aligned} \tag{8}$$

This results in a Multinomial distribution for the number of halibut caught, other species caught and empty (baited) hooks when N hooks are deployed for a soak time of S minutes. The probability for a halibut being caught when there is competition from other species is,

$$(1-\exp(-\lambda S))\left(\frac{\lambda_1}{\lambda}\right). \tag{9}$$

If λ_1 is greater than λ_2 then the probability of having a halibut on a hook will be higher than having another species on the hook.

The expected numbers of the halibut caught by a longline with N hooks is simply N times the probability in equation 9. Therefore the number of the halibut that were observed reflects an index of the abundance of that species modified by the relative abundance of other species competing for the same hooks. The λ_1 and λ_2 parameters reflect the relative abundance of halibut and other species, respectively. The objective here is to estimate each of these parameters to obtain what the probabilities would have been for either halibut (P_1^*) or the other species (P_2^*) had there been no competition. That is,

$$\begin{aligned} P_1^* &= (1-\exp(-\lambda_1 S)) \\ P_2^* &= (1-\exp(-\lambda_2 S)) \end{aligned} \tag{10}$$

The differences between the probabilities for halibut for the cases of competition or no competition will become more pronounced as the proportion of empty baited hooks decreases (Figure 2, see Rothschild (1967)). These differences will be at a maximum for any proportion of empty hooks when the proportions of target and non-target species in the catch are similar — that is, when they are equally likely to be caught on a hook. Estimates of the catch rate for halibut using equation 9 will underestimate the actual relative abundance for halibut due to the competition with the other species for the same hooks. The degree of underestimation will also be a decreasing function of the number of empty baited hooks that are available when the gear are recovered. That is, the impact of other species on the halibut catch rate will be less when there are many empty baited hooks still to choose from but greater when fewer empty baited hooks are available.

The numbers of empty hooks per set were not recorded in the database and were estimated here as the number of hooks deployed minus the estimated total number of fish caught in a set[2]. For this analysis the empty hooks were all assumed to be baited. In 75 percent of the sets, the proportion of empty hooks was 0.86 and higher with this percentage increasing in 2008 and later as the catches of other species in the sets declined (Figure 3). Over the whole time series, the proportion of hooks with halibut was never greater than 0.11. In the context of Figure 2, the degree of underestimation for the probability of catching halibut for most sets may not be serious given the high proportion of empty hooks and the low proportion of hooks with halibut on them. Estimates for λ_1 and λ_2 were obtained for each set in the data base ($>$ 500 hooks) using the maximum likelihood (ML) estimators for the Multinomial model (equation 8) provided in Etienne et al. (2013).

$$\begin{aligned}
\widehat{\lambda}_1 &= \frac{n_1}{N - n_0}\frac{1}{S}\log\left(\frac{N}{n_0}\right) \\
\widehat{\lambda}_2 &= \frac{n_2}{N - n_0}\frac{1}{S}\log\left(\frac{N}{n_0}\right)
\end{aligned} \tag{11}$$

A comparison of the estimated probability of capturing a halibut for each set assuming either competition or no competition shows that there was little difference between the two (Figure 4). Again this was due to the high percentage of empty presumably baited hooks, especially in the latter years (Figure 3). Having more empty hooks available will tend to mitigate against the effects of competition when the relative proportions of halibut and other species being caught are similar. However, this evaluation assumes that all empty hooks still have bait on them and are capable of fishing.

The condition of the empty hooks was not recorded in the halibut longline survey but have been recorded in other similar surveys. Webster and Hare (2009) report that bait was missing from 40 to 60 percent of the total number of hooks deployed in the International Pacific Halibut Commission (IPHC) halibut longline surveys conducted from 2007 to 2009. Earlier studies by the IPHC had recorded bait

[2]The total number of fish caught including halibut and the separate count of halibut was provided for each station in the halibut survey database. Stations were removed from the analysis when there was a total weight recorded that was greater than the halibut weight but the total number caught was equal to the number of halibut caught. The total fish catch may have only included counts for some of the bycatch species even though they were all weighed but it was not possible to ascertain this with the data provided.

loss at 63 percent of the total number of hooks deployed (Skud 1978b). Pelagic longline surveys conducted off of Hawaii from 1987–1989 reported an average of 24 percent of the hooks with baits missing (Somerton and Kikkawa 1995) while bottom longline surveys conducted off of Norway in 1983 to evaluate the performance of different kinds of bait indicated that 40 percent of the hooks were returned without bait (Bjordal 1983). The reasons for bait loss include mechanical loss due to deployment, recovery, or currents during fishing as well as removal of bait by target and non-target species that avoided capture. Bait loss is also likely to be affected by the type of bait and hook used (e.g., Bjordal 1983). However, the impact of bait loss is in the reduction in the number of empty hooks capable of catching fish which in turn could increase the degree of underestimation of the probability of capturing halibut on the longline survey.

Assuming that only 50 percent of the empty hooks still have bait on them each year results in increasing differences between the catch rates for halibut assuming either competition or no competition as the proportion of hooks with halibut on them in the survey increases (Figure 5). Estimates for the catch rates were obtained based on the following ML estimates from Etienne et al. (2013). In this case, estimates assume that hooks missing bait were due to species other than halibut that had evaded capture.

$$\begin{aligned}
\widehat{\lambda}_1 &= \frac{n_1}{N - n_{0,b}} \frac{1}{S} \log\left(\frac{N}{n_{0,b}}\right) \\
\widehat{\lambda}_2 &= \frac{n_2 + n_{0,e}}{N - n_{0,b}} \frac{1}{S} \log\left(\frac{N}{n_{0,b}}\right)
\end{aligned} \tag{12}$$

Those authors that have investigated the impact of missing bait on longline estimates have generally assumed that missing baits were due to non-target species (e.g., Bjordal 1983; Hovgård and Lassen 2000; Webster and Hare 2009). The reasoning appears to be that since the longline survey had been designed for the target species in terms of bait, hook type and size, depth, and timing, non-capture would be more likely for other species in the area. On the other hand, Somerton and Kikkawa (1995) chose to allocate catch of target and non-target species to both baited and unbaited hooks in proportion to their relative abundance during the time period. Underwater observations on bait loss from longlines using cameras or submersibles have recorded both target and non-target fish as well as invertebrates removing baits without becoming permanently hooked with some types of baits being more preferred than others (e.g., Pacific halibut High 1980; Atlantic cod, He 1996). Etienne et al. (2013) also provides ML estimates for allocating halibut or other species to empty hooks in proportion to their respective relative abundance. The resulting estimates assuming 50 percent of empty hooks were missing bait indicates a very pronounced underestimation of the halibut catch rate when competition has not been accounted for (Figure 6). Clearly, the assumptions being made about how empty hooks lose their bait can have large effects on estimates of stock status.

Observing and recording the number of empty unbaited hooks retrieved during a longline survey set will add more work on the survey and could result in fewer sets being conducted. The IPHC longline survey does collect this kind of information and estimates of the hook status (e.g., empty, returned bait, species captured, bait type) are based on the first 20 consecutive hooks of each skate retrieved (Henry et al. 2013). The current IPHC survey longline skate consists of an 1,800-foot line with 100-16/0 circle hooks spaced 18 feet apart and baited with 0.25 lbs to 0.33 lbs of chum salmon (*Oncorhynchus keta*). Five skates are sampled at each station in the survey with a minimum soak time of five hours. The choice of a subsample of 20 hooks was evaluated by Webster and Leaman (2013) and found to provide adequate estimates of the proportion of unbaited empty hooks when compared to observing all 100 hooks in a skate. However, the IPHC sampling program cannot determine what species of fish escaped with the bait from the hook and whether or not that fish avoided the gear or was actually caught subsequently. As noted above, even if the number of unbaited hooks is known, allocating these hooks to species can have a large effect on the resulting estimates.

Competition between target and non-target fish for the same hooks may lead to gear saturation where the local population density of the target or non-target species exceeds the number of hooks in an area leading to decreasing efficiency of fishing effort (Beverton and Holt 1957). If this happens catch rate no longer increases in direct proportion to population density. The potential for gear saturation on a set basis will need to be evaluated in the context of how many baited empty hooks remain at the end of the soak time. Even if there are large numbers of baited empty hooks remaining, gear saturation may occur on a local part of the longline where the numbers of fish exceed the number of available hooks (e.g., Somerton and Kikkawa 1995).

EVALUATING COVARIATES

The terms of reference for this project requested an evaluation of the significance of potential explanatory measures such as NAFO division or subdivision, vessel, Captain, depth, temperature, bait, total catch of other species or other measures of gear saturation, and hook size (Appendix 1). The catch of other species has already been incorporated as a response variable in the Multinomial model. All of the remaining measures with the possible exception of depth and temperature could be set up as categorical factors (e.g., captain, hook size) with levels set to individual captains or hook sizes. Depth and temperature could be added to a GLM as continuous covariates. The standard Multinomial GLM (e.g., McCullagh and Nelder 1983) would not apply in general to the model in equation 8 except for the specific case where there was a high proportion of empty hooks and all empty hooks were found to be baited upon recovery. In this case, observed proportions of halibut and other species caught are virtually the same whether competition did or did not occur. The standard Multinomial GLM was used here assuming this specific case holds to provide a preliminary test for the effects of the above factors and covariates. Once more information on bait loss becomes available alternative methods to model covariates will need to be developed.

The ideal situation would be to have all levels of each categorical factor available each year so as to be able to separate out their individual effects independent of annual changes in halibut abundance. This survey was not designed to test the impacts of these covariates or factors in a manner expected for designed experiments and analysed using a GLM. The participation of vessels and Captains has varied over time according to availability and other factors. Bait type, hook size, and hook type have also varied over time. The difficulty that can arise in this and other similar observational studies is that some levels of some of the factors may only occur in a limited number of years leading to difficulties in separating out individual effects. As an extreme example, consider the situation where a certain Captain/vessel only fished in one year within a limited depth range. Without knowing how this Captain fished in other years at other depths, it would be difficult to know what was behind any differences between this Captain's catch rate and those of the other Captains that may or may not have fished in the same year. Therefore it is prudent to screen the data before any models are fitted to it to see what effects may be identifiable or confounded with other effects.

The first step in screening the data was to remove any missing data associated with the number of other species caught in the survey (Table 2). As noted above soak times in the survey have varied over time. Both the stratified simple estimate (equation 2) and the Multinomial model assume that catch increases linearly with soak time while the models used by Trzcinski et al. (2009) and den Heyer et al. (2015) ignored soak time beyond setting a minimum soak time of 180 minutes. The IPHC halibut survey only uses sets with soak times between 300 and 1440 minutes (Henry et al. 2013). A comparison of the proportion of hooks with halibut from the survey with soak time indicates that soak time does not appear to have any effect between around 180 and 1250 minutes as the full range of catches of halibut lies within these bounds (Figure 7). A similar range appears to be appropriate for the proportion of hooks with catches of the other non-target species as well (Figure 8). Therefore, the next step in screening the data was to only use sets that fall within these bounds for soak times as well as the minimum of 500 hooks used in den Heyer et al. (2015) (Table 2). Similar to den Heyer et al. (2015) only sets which have been sampled in four or more years were included. Finally, sets with missing data for any of the above covariates were removed from the data set. Note that hook type and size were not recorded for the first two years of the survey. There were too few sets for 5YZ to be included in any analysis of annual effects.

Temperature data have only been collected for a subset of the longline sets starting in 2006 (Table 2). Therefore, the full data set cannot be used to model all of the covariates and factors listed in the terms of reference.

An additional screening was conducted to determine if all levels of each factor were available each year. Forty individuals have been identified as Captains for the survey with 60 different vessels. While some Captains have used more than one vessel over time more than half were identified with just one vessel over the history of the survey. As a result, it would be difficult to identify separate results associated with vessels and Captains if both were in the model, so only Captain will be considered for any analysis here. The distribution of participation by individual Captains over time is quite patchy for most participants with none of the Captains participating every year in the screened data set (Table 3). In fact only seven have participated for 10 or more years. The overall median participation period was 5 years without those years being necessarily contiguous. Any comparison of Captains in the GLM analysis would be restricted to different groups of individuals for different time periods and unlikely to provide any insight into the annual trends or variation in the survey abundance index.

The different kinds of bait recorded in the survey often refer to mixtures of different species within a set or a group of sets without any information about what was used in any one particular set (Table 4). The majority of sets used herring or mackerel or herring/mackerel mixtures with some including other species. Without knowing the proportions of species within the mixtures and how individual sets were baited with these mixtures it is unlikely that any analysis would be able to identify whether differences were due to herring, mackerel or the mixture of the two plus other species. A subset of the data was extracted consisting of sets from 4VWX for which comparisons could be made between sets that exclusively used herring or mackerel (2000 to 2014, excluding 2003 and 2013 where either mackerel or herring were not used as the only bait for a set).

The major hook type used in the survey was the circle hook and the occasions when other types of hooks were used were so few that any comparison of hook type would be meaningless (Table 5). A comparison of size 14 and 15 circle hooks could be possible using only data from 2000 to 2010 from both 4VWX and 3NOPs however, a comparison of all three sizes (14, 15, 16) could only be made using data from 2010 to 2015 for 4VWX (Table 6).

All of the above screening reduced the candidate covariates and factors for GLM analysis to bait, depth and hook size for subsets of the data, and temperature for a further subset of the data. The response variables for the Multinomial GLM model are defined in terms of the number observed in one of the categories (e.g., halibut) relative to one of the other categories (e.g., other species or empty baited hooks). The ratio of halibut to other species would correspond to a GLM of the ratio of capture rates λ_1/λ_2 as a function of the covariates.

The Multinomial GLM model was fit to the data using the R function `multinom` from the R package `nnet` (Venables and Ripley 2002). The first analysis compared the effects of using only herring or only mackerel bait on the probabilities of catching halibut or other species. The best fitting model included bait type and year as main effects and interactions between the two (Table 7). The main effect of using herring or mackerel bait appears to be an increase in the catch of other species for mackerel bait relative to herring bait in a few of the years (Figure 9). It is possible that there were other factors behind this difference as these results are only based on a limited subset of the data.

There are a few studies on bait types for longlines with many indicating that the main effect is that bait loss rates vary among bait species, with soft-bodied mackerel (*Scomber spp.*) or herring (*Harengus spp.*) more likely to fall off hooks or to be torn from hooks than firmer-bodied squid (Bjordal 1983; He 1996; Sigler 2000; Ward and Myers 2007). However, High (1980) reported from submersible observations that while herring was more likely to be removed from hooks by fish, hooks with herring appeared to be more attractive to fish than squid baited hooks. Woll et al. (2001) reported that grenadier bait resulted in higher catch rates of Greenland halibut and lower catches of other species than squid bait. Given that the Multinomial estimates (e.g., equation 12) can potentially correct for missing baits, studies on the impact of bait type on bait loss may not be necessary if the number of unbaited hooks per set is known.

The next analysis investigated the effects of depth, hook size and year on sets conducted in 4VWX from 2010 to 2015. The initial models included all of these measures as main effects, however summarizing the depth data for each hook size indicated that sets with the size 16 hooks were generally used in deeper water than the other two sizes in all areas except for 4Vn (Table 8). An additional model was fit to the data incorporating a depth/hook size interaction term to accommodate this pattern in the data (Table 9). This latter model explained more of the deviance than the others considered here although it still only explained about one percent of the residual deviance from the null model. The predicted ratios of the catch of halibut catch to the catch of the other species from this model suggest that the size 16 hooks tended to catch more halibut than other species when compared to the other two sizes with this difference decreasing with increasing depth (Figure 10). However, the predicted catch rate for halibut in the Multinomial model for all hook sizes remains similar, it is just the catch rate for the other species that decreases for the size 16 hooks. Therefore it appears that the size 16 hooks may reduce competition between halibut and the other species caught in the survey compared to the two smaller hooks sizes. These results should be considered preliminary due to the limited data set used but given that the maximum likelihood estimates already correct for competition, no other corrections may be required for estimating the annual halibut catch rate.

The hook size effects identified above will need to be included in the analysis for temperature effects. Depth and temperature are often related in the ocean and for this data set there is a strong correlation between the two covariates for depths below about 280 m (Figure 11). For sets deeper than 280 m, it will be difficult to differentiate between the effects of depth or temperature when both are in the model. An interaction term will not help here because there does not seem to be a relationship between the two covariates for depths shallower than 280 m. Restricting the analysis to only those sets shallower than 280 m indicates that there is a significant effect due to temperature and the regression coefficient was less than zero indicating that the ratio of the catch of halibut to the other species will increase as temperatures decrease (Table 10). However, this model only reduced the deviance residuals by four percent relative to the null model. Similar to the hook size analysis, decreasing temperature appears to decrease the catch of other species while not changing the expected catch rate for halibut by an appreciable amount. The effect of this decrease in competition will be accounted for in the ML estimates for λ_1.

OPERATIONAL STANDARDS

The limited data set analysed for the effects of bait suggested little direct effect on the catch of halibut and the effects estimated for the catch of other species only occurred in a few of the years possibly reflecting the effects of other factors that were not included in the model. Differential bait loss could not be assessed here due to the lack of information on baited and unbaited empty hooks. The results for depth, hook size and temperature show that they can affect the catch of other species relative to the catch of halibut, but the relationships were not strong and the impact of these effects can be accounted for as competition in the Multinomial model estimates. All of the above results are based upon subsets of the data and may not represent the results from more detailed experiments on the effects of these and other covariates. There are published studies on the impacts of bait type and hook size plus other operational aspects (e.g., distance between hooks) on the catch of different species including halibut on longlines (e.g., Hamley and Skud 1978; Skud 1978a,b; Nygaard 2014). The studies conducted by the IPHC led to their standardization of the longline gear, hook size and bait type discussed above (Henry et al. 2013).

Adoption of the Multinomial model to estimate the halibut survey index will result in operational changes to the survey. That is, more attention will need to be made to recording the number of other species being caught in the survey and an estimate of the number of unbaited empty hooks will need to be determined for all or a sample of the empty hooks retrieved.

TOR 2. COMPARE WITH THE OTHER INDICES

ANNUAL ESTIMATES

Annual estimates by any of the methods presented here including the Multinomial model assume that locations of longline sets are a random sample of all possible locations that could be fished. However, the survey design for the longline survey specifies fixed locations of longline sets over time. Fixed locations could cause issues with interpreting the annual estimates if these fixed locations do not represent trends over the areas that were not sampled. It would be difficult to evaluate whether or not this is true with the data at hand, however comparison of the longline trends with other data sources may provide some insight. For the analysis at hand, random sampling will be assumed. The survey was also designed with strata reflecting Low, Medium and High catch areas. The stratification will be ignored for the annual estimates that follow to facilitate comparison with NB GLM index as currently used in den Heyer et al. (2015) but will be evaluated in TOR 3 below.

The ML estimators for the Multinomial model can provide annual estimates when soak times vary over sets but would require numerical optimization to obtain solutions to the equations. On the other hand, the annual estimates are more straightforward if soak time does not vary over sets or catch rate does not vary with soak time. The catch weight estimates used in den Heyer et al. (2015) assumes a soak time of 600 minutes for all times greater than 180 minutes. The current protocol for the IPHC survey is to assume that catch rate is unaffected by soak times greater than 300 minutes and less than 1440 minutes (Henry et al. 2013). There was no evidence of any relationship between the proportion of halibut or other species caught and soak times between 180 to 1250 minutes for the Atlantic halibut longline survey (Figures 7 and 8).

Assuming that catch rate does not vary with soak time and only using sets that fall within the above range of soak times, annual catch rate estimates were compared for the simple unstratified estimate for numbers, and Multinomial model based estimates assuming no unbaited hooks, 50 percent unbaited hooks which were assigned to other species only and 10 percent unbaited hooks allocated to halibut and other species in proportion to their relative abundance (Figure 12). The estimates were all standardized in terms of a 600 minute soak time and 1000 hooks. There were only minor differences between the simple estimate and Multinomial estimate when all empty hooks were assumed to be baited upon recovery. Overall, proportions of halibut and other species of fish were low and there was only a small correction due to competition because there were still a large number of “baited” hooks available to catch either types of fish. The competition correction was somewhat larger for the 50 percent unbaited case when all of the unbaited hooks were assigned to the other species. However, an assumption of only 10 percent empty hooks with unbaited hooks allocated to both halibut and other species resulted in a more than 5 times increase in the halibut catch rate in the more recent years, very much magnifying the recent increase in abundance. This increase over the other esimates was directly due to the concurrent decrease in the catch of other species so that the relative densities of halibut and other species in the catch in recent years were about equal (Figure 12).

Etienne et al. (2013) defines asymptotic variances for the λ_i and not for the standardized estimate of the number of halibut caught per set for a standard soak time and number of hooks. Development of a variance estimate based on λ_1 for $(1-\exp(-\lambda_1 S)) \times N$ is not straightforward, however bootstrap estimates of the variance and confidence intervals can be defined and used instead (Efron and Tibshirani 1993). The 95 percent bias-corrected accelerated confidence intervals were calculated separately for mean numbers per set for 3NOPs4VWX, 3NOPs, and 4VWX for the soak time limits and total hook limits described above (Figure 13). The wider confidence intervals for 3NOPs mainly reflect the smaller sample sizes there relative to 4VWX.

All of the modelling described above was in terms of numbers caught to correspond to the interactions between individual fish and baited hooks. The NB GLM indices for the halibut longline survey were in terms of weight of halibut caught. Numerical abundance estimates can be transformed to weights by multiplying the mean number of halibut caught by the mean weight of a halibut. To accommodate the variation in mean weight by size over the survey area, abundance and mean weight estimates were calculated within each NAFO (sub)division with the overall estimates of mean weight per standard set calculated as the weighted average over NAFO (sub)divisions with weights equal to the areas. In addition, mean weights were calculated within each bootstrap replicate based on the sets that were chosen to be in the replicate. Note the higher mean weight per set in 3NOPs compared to 4VWX despite having similar mean numbers per set (compare Figures 13 and 14), possibly reflecting the higher frequency of larger fish in the former area as seen in the observer samples of longline sets (Figure 9, den Heyer et al. 2015).

OTHER INDICES

Annual weight per standard set indices using the Multinomial model assuming all empty hooks are baited and the NB GLM standardized estimates given in den Heyer et al. (2015) updated to include data from 2015 (N. den Heyer, pers. com.) were compared against the time series for landings and commercial catch rate. The two survey time series exhibit similar trends with respect to a period of little change from 1998 to 2005/2006 and then a rapid increase to 2011 followed by little change until an increase from 2014 to 2015 (e.g., Figure 15). On the other hand, the landings show an increase from 2000 to 2003, followed by no change until 2009. No data on landings for 2014 and 2015 were provided but landings from 2010 to 2013 increased while the indices indicated no change in abundance.

The commercial halibut index also indicated a period of no change up to 2004, with an increase starting in 2006 and then a rapid increase in 2014 and 2015 (Figure 16).

Research trawl survey indices for mean numbers of halibut per tow in 4VWX and 3NOPs were also compared with the Multinomial model estimates of numbers of halibut per standard longline set. The 4VWX survey index was similar to the longline index with respect to a period of little change in the early part of the series and a general increase until 2011 (Figure 17). However, the trawl survey index showed a rapid decline in 2012 and 2013 followed by an increase. The 3NOPs trawl index exhibited an increase in first few years very similar to that in the landings series (Figure 18). The trawl and longline survey series are similar in showing a general increase in abundance in the latter part of the time series but diverge starting in 2013.

Biomass estimates up to 2013 for legal size halibut from the Statistical Catch-At-Length Model (SCAL) developed by Cox et al. (2016) were provided for comparison as well. This model incorporates the NB GLM index, the landing series as well as the 4VWX trawl survey total numbers index (Figure 19). The estimated biomass reflects trends similar to these other indices and to the Multinomial longline index, however the model predicts continued increases in 2012 and 2013 which appear to be unsupported by the longline or the trawl survey indices.

TOR 3/4. EVALUATING SURVEY DESIGN AND STRATIFICATION

As noted earlier, the current survey was initially set up as a stratified design with fixed stations in each strata. The design has been retained with respect to the allocation of stations by the Low, Medium and High catch strata but the design has not been used in calculating the survey indices beginning with the assessment by Trzcinski et al. (2009). Apparently, the stratified estimates were discontinued because the design did not seem to improve coefficient of variation (CV) of the estimates compared to the case where the design was ignored (C. Den Heyer, pers. com.). Over the recent period (2012 to 2015) the number of stations assigned to each NAFO (sub) division has been stable and CVs for the stratified Multinomial model estimates within NAFO (sub)divisions have ranged from 0.13 to 0.36 in 2015 (Table 11). Generally, the lower CVs would be associated with the higher sample sizes but this was not always the case.

Stratified designs provide more precise estimates than simple random designs when observations are more similar within a stratum than between strata and sample allocation by stratum is either proportional to the size of the stratum or proportional to the size times the standard deviation of the data for each stratum (Cochran 1977; Smith and Gavaris 1993). Standard errors by strata will vary over time and the more robust sampling plan to put into place is to match the sampling rate to the size of the strata. This matching can be done at two stages. A total of

232 sets were originally allocated over all NAFO (sub)divisions in 2015. Assuming that all of these sets actually meet the protocols of recording bycatch by number, keeping soak time within the bounds defined above, and using more than 500 hooks, they would be allocated to the individual NAFO (sub)divisions relative to their areas (Table 12). At the second stage sets are allocated proportional to the area of the catch strata within the NAFO (sub)divisions. This will result in changes to the current allocation where very small areas such as 4W High and 4X Low had received higher sample sizes than the other areas. While those original larger sample sizes would have resulted in very small standard errors, these two strata contributed little to overall stratified estimates of standard error for their respective NAFO (sub)division.

Determining the sample size required to obtain a particular CV is straightforward when dealing with simple means. The relationship between sample size and standard error is more complicated for the Multinomial ML estimates used here and so the bootstrap method was used to evaluate what CV could be expected given the proportional station allocation plan given in Table 12. That is, the original set of observations were resampled with replacement for the new sample sizes in Table 12. This approach assumes that the current set of observations adequately describes the distribution of additional samples even when the proportional allocation plan calls for a larger sample size than was originally used (see for e.g., Manly 1992). In the example below, the observed CV for 2015 is compared with the expected CV for the new sampling plan. In all but one case, the CV was reduced including those NAFO (sub)divisions that had their overall sample size reduced (e.g., 4X5Z).

	3N	3O	3Ps	4Vn	4Vs	4W	4X5Z
CV 2015	0.23	0.24	0.36	0.16	0.13	0.21	0.15
CV Proportional	0.15	0.18	0.24	0.15	0.14	0.13	0.14

The increase in precision from proportional allocation is simply due to how well the strata reflect the distribution of the species being measured. The current strata were based on a limited amount of commercial log data from 1995 to 1997. Improved precision could be obtained with better stratification such as some spatial measure of habitat. Currently measured covariates such as depth will probably be part of the definition of habitat but preferred depths may differ over areas depending upon water current regimes, temperature profiles and bottom type amongst other possible determinants of habitat.

TOR 5. TRANSITIONAL PROCEDURES

The following recommendations are provided for the short-term transition.

1. Adopt the Multinomial model estimate but assume all empty hooks have bait on them.

   The resulting index will be lower than the survey index corrected for unbaited hooks but until the questions of how many hooks are recovered without bait and how these hooks should be allocated to the species caught have been answered, this approach may be considered to be precautionary. However, this approach also requires that the number of hooks occupied by other species is also recorded for reach set. This has not being done for every set in the time series. Additionally, all stations need to meet the protocols of recording bycatch by number, keeping soak time within the bounds defined above and using more than 500 hooks so that all sets can be used for the Multinomial estimate.

2. An alternative approach is to adopt the estimate used for the IPHC survey index of mean weight of halibut caught per skate (WPUE). The number and weight of halibut and other species caught per skate along with the number of hooks with baits remaining at haulback are recorded. Survey indices by management area are adjusted annually for the average number of baits remaining as a proportion of the initial number of baits relative to the overall mean number of baits remaining for the whole stock area (Clark 2007). This factor corrects for both empty unbaited hooks and for hooks that have caught other species. Those areas with a smaller than average proportion of baits returned would have their WPUE index adjusted upwards to account for higher competition for baits in the area while the WPUE would be adjusted downwards in areas with a higher than average proportion of baits returned (Clark 2007; Webster et al. 2010). Within any one year, this correction is very similar to the approach by Etienne et al. (2013) described above, however this latter approach also involves correcting for the proportion of baits recovered over the whole time series. To date, the IPHC has continued to use the within year correction factor only (Webster et al. 2014). Variance estimates or confidence intervals estimates have not been defined for the IPHC WPUE estimate but it may be possible to construct a bootstrap estimate.

3. Re-run the SCAL model to obtain a new estimate of q needed for providing interim stock assessment advice.

   The Multinomial model estimate was always greater than or equal to the Negative Binomial survey index (e.g., Figure 15). This will probably cause issues with the current estimates of q used in the HAL operating model evaluation (Cox et al. 2016).

4. The current data available to evaluate the effects of bait type and hook size can only be analysed for subsets of data that may not adequately represent conditions over all areas, years, and depths. Designed experiments to model

the effects of these different operational parameters may be difficult and expensive to do. There are a number of published studies on the effects of bait type, hook size and other operational parameters such as hook spacing and soak time for longline catchability of a number of species including Pacific halibut. The IPHC studies contributed to the standardization now in place for bait, hooks, hook spacing, and soak time. There are also manuals available that stress the need for standardization for longline surveys (e.g., Hovgård and Lassen 2000). All of this experience would argue for foregoing the field experiments and implementing standards now.

5. The proportional allocation plan was designed for 232 stations which falls within the range of stations allocated to the survey in recent years (Table 12). This new allocation plan, based on the existing stratification scheme, can be implemented for the next survey without the need for additional stations or transition from the old allocation to the new one over a period of time. Where sample sizes have been increased in the new allocation plan, the original stations should be retained and estimates from these stations and the new stations should be compared to evaluate the assumption that the original set of stations adequately describes the distribution of the additional observations in the strata. There is less of a concern for those strata where the sample sizes have been decreased, especially those strata where there were large reductions in sample sizes. In these latter cases, the large reductions were the result of very small strata areas which in turn will result in the means from these strata having very small contributions to the stratified means.

In the medium to long term, the following investigations should be conducted.

1. Consult with survey fishing captains and observers on their experience and best estimate of the number of unbaited (or baited) empty hooks in a set. Starting from these consultations design and implement a pilot study to estimate the number of unbaited hooks recovered per set. A longer term approach should be pursued based on the results of the pilot study.

2. The current stratification reflects the spatial distribution of catch from 1995 to 1997. This spatial pattern may reflect variation in habitat suitability for halibut but will also be affected by fishing patterns of the day. Research on identifying habitat associations for halibut could be used to develop more useful stratification for the survey design. There are methods available to transition to a new stratification scheme and recalculate the survey index over the whole time series according to the new stratification series (Smith et al. 2015).

Acknowledgements

Nell den Heyer, Scott Wilson and Gabrielle Wilson provided patient and timely responses to all of my questions about the database and the survey. Constructive feedback was provided by science, industry and other participants at the project review meeting held at the Bedford Institute of Oceanography, Dartmouth, Nova Scotia on February 18, 2016. Adam Cook and Heath Stone provided helpful comments and editing suggestions on the previous draft of this report.

REFERENCES

Armsworthy, S., Wilson, S., and Mohn, R.K. 2006. Atlantic halibut on the Scotian Shelf and Southern Grand Banks (Div. 3NOPs4VWX5Zc) — Industry/DFO longline survey results to 2005. DFO Can. Sci. Advis. Sec. Res. Doc. **2006/065**: ii + 31 p.

Baum, J.K., and Blanchard, W. 2010. Inferring shark population trends from generalized linear mixed models of pelagic longline catch and effort data. Fish. Res. **102**: 229 – 239.

Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Her Majesty's Stationary Office, London, 533 p.

Bjordal, Å. 1983. Effect of different long-line baits (mackerel, squid) on catch rates and selectivity for tusk and ling. ICES C.M. **1983/B:31**: 9 p. (mimeo).

Clark, W.G. 2007. Effect of hook competition on WPUE. IPHC Rep. Assess. Res. Act. **2007**: 211–215.

Cochran, W.G. 1977. Sampling techniques. John Wiley & Sons, New York, N.Y., 428 p.

Cox, S.P., Benson, A., and den Heyer, C.E. 2016. Framework stock assessment for Atlantic halibut on the Scotian Shelf and Southern Grand Banks. DFO Can. Sci. Advis. Sec. Res. Doc. **2016/001**: v + 54 p.

den Heyer, C., Hubley, B., Themelis, D., Smith, S.C., Wilson, S., and Wilson, G. 2015. Atlantic halibut on the Scotian Shelf and Southern Grand Banks: Data review and assessment model update. DFO Can. Sci. Advis. Sec. Res. Doc. **2015/051**: v + 82 p.

Efron, B., and Tibshirani, R.J. 1993. An introduction to the bootstrap. Chapman & Hall, London, 437 p.

Etienne, M.P., Obradovich, S.G., Yamanaka, K.L., and McAllister, M.K. 2013. Extracting abundance indices from longline surveys: a method to account for hook competition and unbaited hooks. arXiv **1005.0892v3**: 1–35.

Gunderson, D.R. 1993. Surveys of fisheries resources. John Wiley, 248 p.

Hamley, J.M., and Skud, B.E. 1978. Factors affecting longline catch and effort: II. hook-spacing. IPHC Sci. Rep. **64**: 15–24.

He, P. 1996. Bait loss from bottom-set longlines as determined by underwater observations and comparative fishing trials. Fish. Res. **27**: 29–36.

Henry, E., Soderlund, E., Dykstra, C.L., Geernaert, T., and Ranta, A.M. 2013. Standardized stock assessment survey. IPHC Rep. Assess. Res. Act. **2013**: 379–414.

High, W.L. 1980. Bait loss from halibut longline gear observed from a submersible. Mar. Fish. Rev. **42**: 26–29.

Hovgård, H., and Lassen, H. 2000. Manual on estimation of selectivity for gillnet and longline gears in abundance surveys. FAO Fish. Tech. Paper. FAO, Rome. **397**: 84 p.

Kaimmer, S.M. 1999. Direct observations on the hooking behavior of Pacific halibut, *Hippogolossus stenolepis*. Fish. Bull. **97**: 873–883.

Kimura, D.K., and Zenger, Jr, H.H. 1997. Standardizing sablefish (*Anoplopoma fimbria*) long-line survey abundance indices by modeling the log-ratio of paired comparative fishing cpues. ICES J. Mar. Sci. **54**: 48–59.

Manly, B.F.J. 1992. Bootstrapping for determining sample sizes in biological studies. J. Exp. Mar. Bio. Ecol. **158**: 189–196.

Maunder, M.N., and Punt, A.E. 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res. **70**: 141–159.

McCullagh, P., and Nelder, J. 1983. Generalized linear models. Chapman and Hall, New York, NY, 580 p. p.

Nygaard, R. 2014. Trawl, gillnet and longline survey results from surveys conducted by the Greenland Institute of Natural Resources in NAFO Division 1A Inshore. NAFO SCR Doc. **14**/**038**(Serial No. N6334): 13 p.

Pelletier, B., and Parma, A.M. 1994. Spatial distribution of Pacific halibut (*Hippoglossus stenolepis*): an application of geostatistics to longline survey data. Can. J. Fish. Aquat. Sci. **51**: 1506–1518.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Bd. Can **191**: 382 p.

Rodgveller, C.J., Lunsford, C.R., and Fujioka, J.T. 2008. Evidence of hook competition in longline surveys. Fish. Bull. **106**: 364–374.

Rothschild, B.J. 1967. Competition for gear in a multiple-species fishery. J. Cons. perm. int. Explor. Mer **31**: 102–110.

Sigler, M.F. 2000. Abundance estimation and capture of sablefish (*Anoplopoma fimbria*) by longline gear. Can. J. Fish. Aquat. Sci. **57**: 1270–1283.

Skud, B.E. 1978a. Factors affecting longline catch and effort: I. general review. IPHC Sci. Rep. **64**: 6–10.

Skud, B.E. 1978b. Factors affecting longline catch and effort: III. Bait loss and competition. IPHC Sci. Rep. **64**: 25–42.

Smith, S.J., and Gavaris, S. 1993. Improving the precision of fish abundance estimates of Eastern Scotian Shelf cod from bottom trawl surveys. N. Am. J. Fish. Manage. **13**: 35–47.

Smith, S.J., Nasmith, L., Glass, A., Hubley, B., and Sameoto, J.A. 2015. Framework assessment for SFA 29 West scallop fishery. DFO Can. Sci. Advis. Sec. Res. Doc. **2014/110**: v + 69 p.

Soderlund, E., Randolph, D.L., and Dykstra, C. 2012. IPHC Setline Charters 1963 through 2003. IPHC Tech. Rep. **58**: 264 p.

Somerton, D.A., and Kikkawa, B.S. 1995. A stock survey technique using the time to capture individual fish on longlines. Can. J. Fish. Aquat. Sci. **52**: 260–267.

Somerton, D.A., Kikkawa, B.S., and Wilson, C.D. 1988. Hook timers to measure the capture time of individual fish. Mar. Fish. Rev. **50**: 1–5.

Trzcinski, M.K., Armsworthy, S.L., Wilson, S., Mohn, R.K., Fowler, M., and Campana, S.E. 2009. Atlantic halibut on the Scotian Shelf and Southern Grand Banks (NAFO Divisions 3NOPs4VWX5Zc) — Industry/DFO longline survey and tagging results to 2008. DFO Can. Sci. Advis. Sec. Res. Doc. **2009/026**: vi + 43 p.

Venables, W.N., and Ripley, B.D. 2002. Modern applied statistics with S. Springer, New York, fourth edition, 495 p.

Ward, P., and Myers, R.A. 2007. Bait loss and its potential effects on fishing power in pelagic longline fisheries. Fish. Res. **86**: 69–76.

Webster, R.A., Dykstra, C.L., Henry, E., Soderlund, E., and Kong, T. 2014. Setline survey expansions in 2014 and use of sablefish longline survey data for a deep-water density index. IPHC Rep. Assess. Res. Act. **2014**: 603–617.

Webster, R.A., and Hare, S.R. 2009. Adjusting IPHC setline survey WPUE for survey timing, hook competition and station depth. IPHC Rep. Assess. Res. Act. **2009**: 187–208.

Webster, R.A., Hare, S.R., Valero, J.L., and Leaman, B.M. 2010. Notes on the IPHC setline survey design, alternatives for estimating biomass distribution, and the hook competition adjustment. IPHC Rep. Assess. Res. Act. **2010**: 229–240.

Webster, R.A., and Leaman, B.M. 2013. Setline survey hook counts: are the "first 20 hooks" sufficient? IPHC Rep. Assess. Res. Act. **2013**: 421–432.

Woll, A.K., Boje, J., Holst, R., and Gundersen, A.C. 2001. Catch rates and hook and bait selectivity in longline fishery for Greenland halibut (*Reinhardtius hippoglossoides*, Walbaum) at East Greenland. Fish. Res. **51**: 237–246.

Zwanenburg, K.C.T., and Wilson, S. 2000. The Scotian Shelf and Southern Grand Banks Atlantic halibut (*Hippoglossus hippoglossus*) survey — Collaboration between the fishing and fisheries science communities. ICES CM **2000/W:20**: 25 p.

APPENDIX 1

DELIVERABLES

1. Evaluate the current GLM model used to standardize halibut catch rates and size composition. Investigate the significance of co-factors such as NAFO division or subdivision, vessel, captain, depth, temperature, bait, total catch of other species or other measures of gear saturation and hook size. Propose operational standard for the survey based on outcomes of the GLM investigations (gear, soak time, bait).

2. Compare improved index with the other indices of abundance of both undersized and exploitable halibut (e.g. RV surveys in 4VWX and 3NOPs, NAFO division or subdivision-specific landings, Commercial Index CPUE) provided by DFO staff to evaluate ability of survey to track changes over time.

3. Evaluate the current survey design in terms of the stratification and station allocation scheme used. This evaluation should include:
   - Spatial distribution of survey stations
   - Role and standardization (if possible) of significant cofactors identified in step 1.

4. Propose alternate survey design and include a criteria for evaluating the impact of any changes relative to the outcomes of 1–3. Criteria should include an evaluation of changes in: sampling rate within strata
   - sampling rate within a management area
   - total sampling effort as indicator of cost of the survey, using the sampling rate in last 5 years (roughly 220 stations/year) as baseline.

5. Recommend approach to improving survey index of abundance, while maintaining ability to provide science advice on halibut, based on q-adjusted index.

6. Two meetings with DFO:
   - to initiate contract and review data
   - mid-way review of analysis

7. Present criteria and proposed survey design at Halibut Survey Review meeting at BIO (February 9).

8. Report in DFO technical report series (final March 1, 2016)

Table 1. A comparison of the analysis of deviance tables for the three Negative Binomial generalized linear models used to standardize halibut mean weight per longline set. Published results were compared against the same model applied to data supplied for this review.

Source	Terms	Df	Deviance Resid.	Df	Resid. Dev	Pr($> \chi^2$)
1998–2005						
Armsworthy et al. (2006)	Null			1443	2675	
$\hat{\theta}$ = 0.0373	Year	7	1	1436	2674	0.944
	Station	260	1282	1176	1392	$<$ 0.001
Current database	Null			1516	3091.3	
$\hat{\theta}$ = 0.3977	Year	7	4.61	1509	3086.7	0.708
	Station	277	1668.83	1232	1417.9	$<$ 0.001
1998–2008						
Trzcinski et al. (2009)	Null			1949	3274	
$\hat{\theta}$ = 0.3525	Year	10	10.5	1939	3264	0.4
	Station	230	1289.2	1709	1975	$<$ 0.001
Current database	Null			1936	3386.7	
$\hat{\theta}$ = 0.3635	Year	10	7.29	1926	3379.4	0.698
	Station	230	1433.22	1696	1946.2	$<$ 0.001
1998–2013						
den Heyer et al. (2015)	Null			3088	5627.6	
$\hat{\theta}$ = 0.4323	Year	15	150.63	3073	5476.9	$<$ 0.001
	Station	265	2084.83	2808	3392.1	$<$ 0.001
Current database	Null			3061	5385.1	
$\hat{\theta}$ = 0.3635	Year	15	141.72	3046	5243.4	$<$ 0.001
	Station	265	1912.01	2781	3331.4	$<$ 0.001

Table 2. The number of stations available each year for analysis in generalized linear models in 3NOPs4VWX5YZ Halibut longline survey. All stations refers to all sets completed without damage. The next column only includes stations where the catch in terms of the number of other species was recorded. Soak and Hook limits refer to all stations where soak time was between 180 and 1250 minutes and greater than 500 hooks were used. The numbers of stations are further reduced by only using stations that were sampled over four or more years as per the current model in den Heyer et al. (2015). The next two columns reduces the number of stations for those where data on depth and bait, hook type, and hook size, respectively, were available. The final column lists the number of stations that meet all of the previous conditions and have temperature data available.

Year	All stations	Other species	Soak & Hook	≥ 4 years	Depth	Bait, hook, size	Temperature
1998	175	171	156	140	140	0	0
1999	167	166	165	159	159	0	0
2000	217	204	204	196	196	196	0
2001	190	184	184	182	182	182	0
2002	200	196	187	186	186	186	0
2003	189	177	175	171	162	162	0
2004	215	195	194	184	184	184	0
2005	164	155	155	152	152	124	0
2006	163	155	155	152	152	152	93
2007	241	199	190	167	167	163	80
2008	283	231	223	192	192	186	129
2009	213	187	177	173	173	162	91
2010	215	187	186	184	184	184	78
2011	217	204	203	202	202	195	83
2012	217	189	188	187	187	187	67
2013	233	210	204	201	201	126	9
2014	233	204	199	194	194	194	102
2015	232	211	201	200	200	200	131
3NOPs	594	533	533	513	513	452	130
4VWX	3152	2879	2801	2704	2695	2326	731
5YZ	18	13	12	5	5	5	2
Total	3764	3425	3346	3222	3213	2783	863

Table 3. The number of sets available each year for each Captain (anonymous ID number) in the halibut longline survey.

ID	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
1													11	8		
2							21									
3								6								
4										4						
5												4				
6								4				9				
7						8										
8								7	7							
9												8				
10													12			
11								17	11							
12															9	10
13															14	11
14								4	2	6	11	9	11	5	5	4
15						13	11	14	7		9	8	11	10	4	5
16											13	13		9		
17							7	5	10	14	13	12				
18									4		11		9	9	12	10
19								5	15	15	13	11	16	16	13	16
20								18	15	26	25	32	20	34	31	34
21							11					14	14	12	12	13
22							11	9	7	6	9	11	11	13	14	12
23						17	17	10	9	1	10		9	9	10	10
24	9						9									
25	9															
26	12															
27	16															
28	9	9	10			6	20								0	
29			22	16												
30			15	14	23	12	19		9	22						
31	13	13	9	12	17								1			
32	14	16	4	11	15	15	12									
33	13	12	12	18	12										0	
34	16	26	25	17	18	16										
35	15	14	15	13	13	16	14	18	20							
36	15	13	15	11	6			6	12	12	5					
37	16	19	19	15	17				8	12	8	12	11	11	8	11
38	9	14			17			13	14	19	15	16	16	17	17	17
39	11	25	23	17	30	27		18	28	25	32	34	27	38	34	35
40	19	20	16	18	16	9		9	8	10	10	9	8	10	11	10

Table 4. The number of stations available each year for the different bait types used in the halibut longline survey. Note that bait type was not recorded in 1998 or 1999.

				Herring		Herring/Mackerel			Mackerel		
Year	NS	Gaspereau	Herring	Squid	Mackerel	White Hake	Squid	Mackerel	Squid	Red Hake	Squid
2000	106		25		32		15	18			
2001	81		29		24		26	22			
2002	46		51	8	24		23	34			
2003	36				25		4	97			
2004	59		69		22			32		2	
2005	8		56		52			8			
2006			63		34		17	38			
2007			95	13	24			25	6		
2008		10	54	20	29		24	26			23
2009	22	4	72	13	19		5	12			15
2010		15	91	18	25		13	17			5
2011			99	18	59		13	6			
2012		11	83		38		27	9	16		3
2013			55	16	46						9
2014	4	12	105	30	34			9			
2015	3	23	90	18	34	19	3				10

Table 5. The number of stations available each year for analysis of hook type. Note SS-Circle refers to Short-Shank Circle hooks.

Year	Circle	J-Hook	SS-Circle
2000	196		
2001	182		
2002	186		
2003	162		
2004	176	8	
2005	124		
2006	152		
2007	154		9
2008	176		10
2009	149		13
2010	184		
2011	177		18
2012	187		
2013	126		
2014	194		
2015	200		

Table 6. The distribution of the number of stations by hook size by NAFO divisions by year in the halibut longline survey

	3NOPs			4VWX		
Year	14	15	16	14	15	16
2000	43			153		
2001	37			131	13	
2002	24			133	13	15
2003	9	13		107	33	
2004	21	13		125	25	
2005		19		55	69	
2006		6		106	40	
2007	18	9		112	24	
2008	19		18	116	23	10
2009		30		76	66	
2010		16	17	93	43	15
2011			32	89	33	48
2012			23	94	31	39
2013			37	79	57	28
2014			32	69	46	47
2015	0	0	35	57	72	34

Table 7. Analysis of deviance for the effects of bait (herring or mackerel) on the relative catch of halibut to the catch of other species in halibut longline survey (4VWX).

Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr($> \chi^2$)
1	2294	489260.1				
as.factor(YEAR)	2270	448565.0	1 vs 2	24	40695.15	< 0.001
BAIT + as.factor(YEAR)	2268	443624.1	2 vs 3	2	4940.93	< 0.001
BAIT * as.factor(YEAR)	2244	440265.6	3 vs 4	24	3358.48	< 0.001

Table 8. Summary of the depths (m) fished by different sizes of hooks for 4VWX and broken out by 4Vn, 4Vs, 4W, and 4X.

	14	15	16
	4VWX		
Minimum	26.0	27.0	53.0
1st Quartile	95.0	97.0	119.5
Median	128.0	134.0	322.0
Mean	186.2	196.2	289.6
3rd Quartile	221.0	229.0	415.0
Maximum	929.0	931.0	640.0
	4Vn		
Minimum	53.0		57.0
1st Quartile	135.8		64.8
Median	185.0		110.5
Mean	201.7		164.4
3rd Quartile	229.8		230.5
Maximum	432.0		408.0
	4Vs		
Minimum	26.0	27.0	53.0
1st Quartile	97.0	93.0	136.5
Median	137.0	192.0	331.0
Mean	206.9	247.3	293.2
3rd Quartile	255.0	387.0	415.0
Maximum	929.0	931.0	640.0
	4W		
Minimum	26.0	27.0	57.0
1st Quartile	95.0	80.5	87.5
Median	123.0	116.0	272.5
Mean	192.6	191.0	319.2
3rd Quartile	221.0	196.0	548.2
Maximum	929.0	823.0	640.0
	4X		
Minimum	36.0	31.0	112.0
1st Quartile	93.0	97.0	115.0
Median	115.0	123.0	320.0
Mean	158.1	157.1	319.7
3rd Quartile	170.8	170.0	585.0
Maximum	824.0	750.0	640.0

Table 9. Analysis of deviance for the effects of hook size on the relative catch of halibut to the catch of other species in halibut longline survey (4VWX, 2010–2015).

	Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr($>\chi^2$)
1	1	1946	260787.3				
2	YEAR	1936	260356.4	1 vs 2	10	430.89	<0.001
3	HOOKSIZE + YEAR	1932	258401.9	2 vs 3	4	1954.56	<0.001
4	SDEPTH + HOOKSIZE + YEAR	1930	258364.5	3 vs 4	2	37.38	<0.001
5	HOOKSIZE*SDEPTH + YEAR	1926	258033.8	4 vs 5	4	330.71	<0.001

Table 10. Analysis of deviance for the effects of temperature on the relative catch of halibut to the catch of other species in the halibut longline survey (4VWX, 2010–2015).

	Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr($> \chi^2$)
1	1	848	132778.7				
2	+ Year	838	131006.4	1 vs 2	10	1772.33	$<$ 0.001
3	SDEPTH*HOOKSIZE + Year	828	128387.8	2 vs 3	10	2618.60	$<$ 0.001
4	MedT+SDEPTH*HOOKSIZE + Year	826	127279.2	3 vs 4	2	1108.64	$<$ 0.001

Table 11. Bootstrap results for the coefficient of variation (CV) of stratified Multinomial model mean weight of halibut (kg) caught per standard longline set within catch strata in each NAFO division or subdivision. Mean weights were calculated separately for each strata and within each bootstrap replication. NA indicates that data were not available for all strata after data were filtered.

NAFO	Area	CV					Sample size				
(sub)division	km^2	2012	2013	2014	2015	Mean	2012	2013	2014	2015	Mean
3N	16919.36	0.24	0.26	0.32	0.23	0.26	5	6	6	6	5.75
3O	38795.11	0.28	0.26	0.20	0.24	0.24	11	19	17	18	16.25
3Ps	24003.70	0.46	0.16	0.19	0.36	0.29	7	12	9	11	9.75
4Vn	17047.88	0.37	NA	0.18	0.16	0.24	11	NA	10	10	10.33
4Vs	36107.79	0.25	0.17	0.17	0.13	0.18	25	33	32	34	31.00
4W	83159.01	0.38	0.21	0.16	0.21	0.24	50	50	50	46	49.00
4X	75640.81	0.15	0.12	NA	0.15	0.14	78	75	NA	74	75.67

Table 12. Allocation of the original number of stations by NAFO (sub)divisions for 2012–2105 and by catch strata for 2015. Mean CV and mean sample size are from Table 11. Proportional stations are allocated according the relative area for each NAFO (sub)division. Proportional strata allocations (Low, Medium and High) are based on the area within each strata within each NAFO (sub)division. Note that the 3 stations allocated to 5Z were included in the 4X allocation by strata.

NAFO	Unfiltered Stations				Planned Allocation (2105)			Mean, Table 11		Proportional allocation (2015)			
(sub)divisions	2012	2013	2014	2015	Low	Medium	High	Sample size	CV	Stations	Low	Medium	High
3N	6	6	6	6		2	4	5.75	0.26	13		5	8
3O	20	20	20	20	3	4	13	16.25	0.24	30	2	10	18
3Ps	10	13	9	12	2	3	7	9.75	0.29	19	2	6	13
4Vn	12	13	14	14	4	8	2	10.33	0.24	15	4	9	2
4Vs	35	33	37	35	5	6	24	31.00	0.18	28	2	12	14
4W	51	58	52	55	8	23	24	49.00	0.24	65	20	43	2
4X	83	89	94	88	42	43	3	75.67	0.14	59	4	35	23
5Z	0	0	0	2	0	0	0			3			
Total	217	232	232	232				197.75		232			

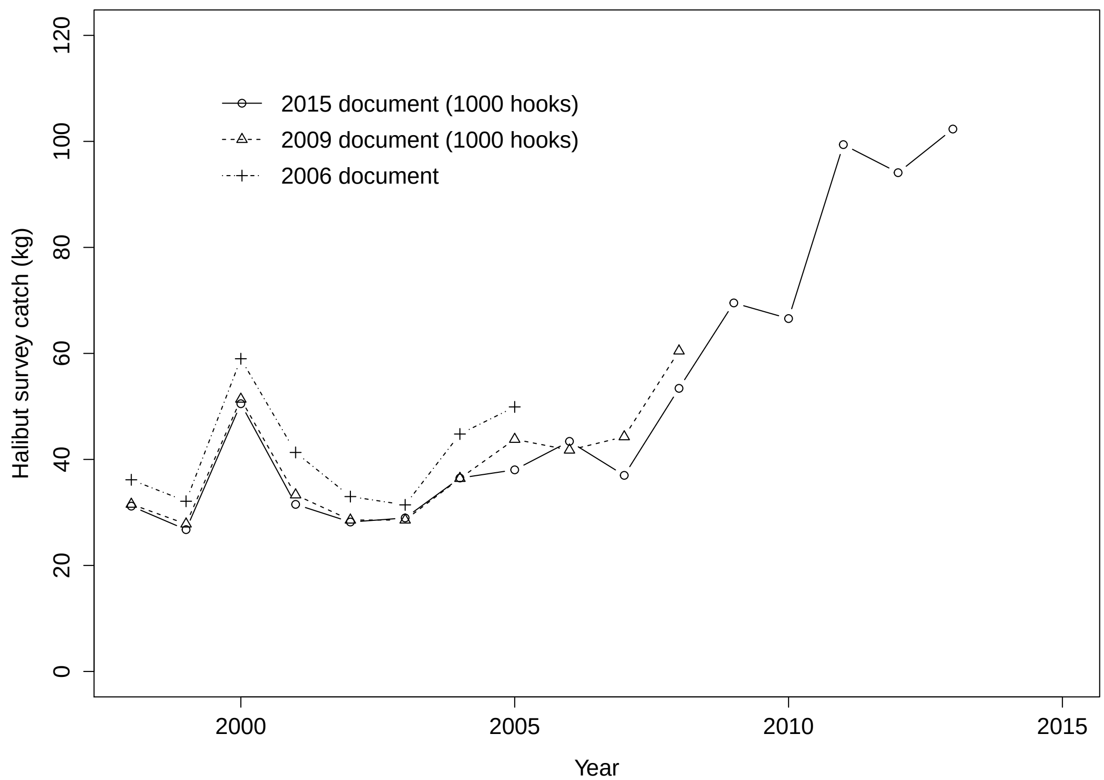

Figure 1. Predicted halibut catch rate (kg/1000 hooks) survey indices for the same set based on current fits of the Negative Binomial generalized linear models used in the 2006 (Armsworthy et al. 2006), 2009 (Trzcinski et al. 2009), and 2013 assessments (den Heyer et al. 2015). Note that the total number of hooks per set was not included in the 2006 model.

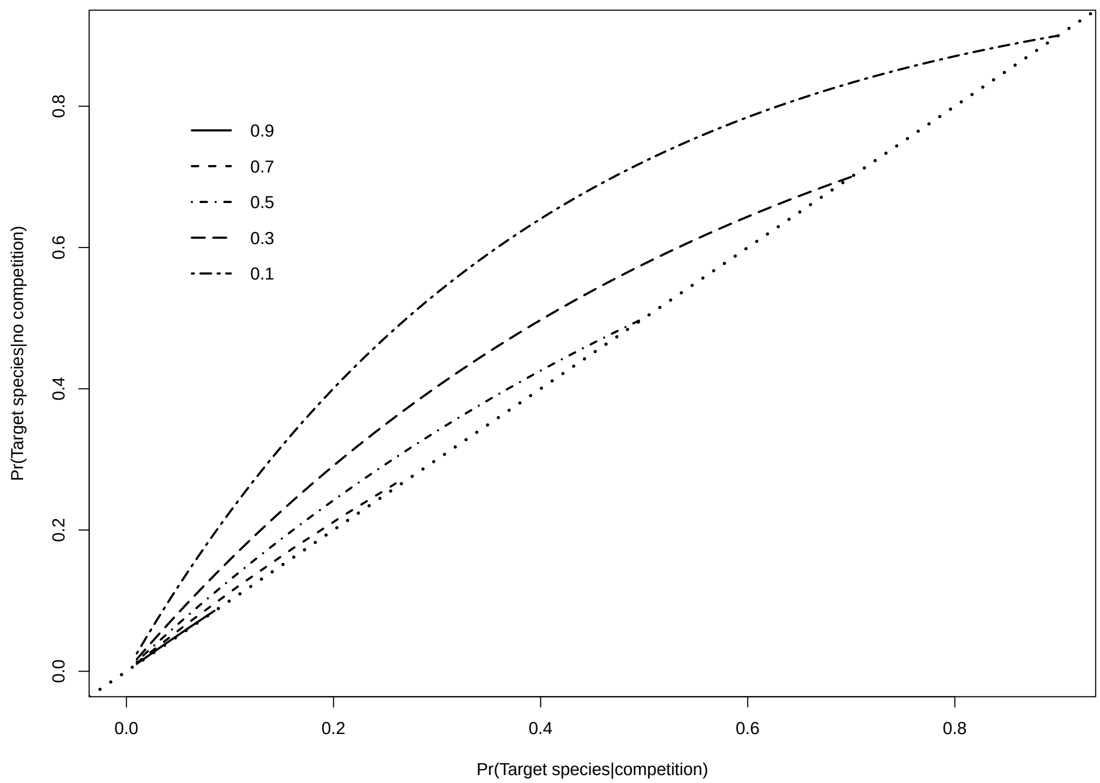

Figure 2. A comparison of the probabilities of capturing the target species for a survey on any one hook for a range of probabilities of empty hooks (0.1,0.3,0.5,0.7,0.9) for cases of no competition (no bycatch) versus competition (bycatch). The 1:1 line is included for reference.

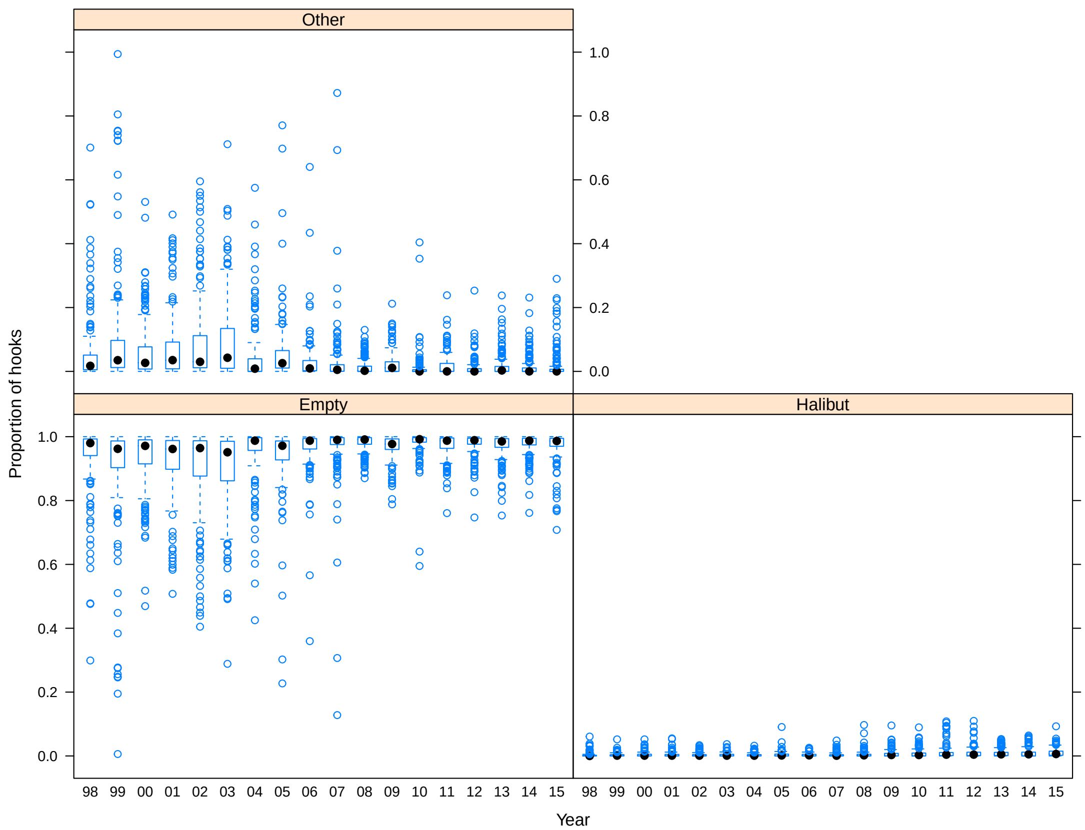

Figure 3. The proportion of hooks that were empty, with halibut or with other species each year for the halibut longline survey in 3NOPs4VWX5Z.

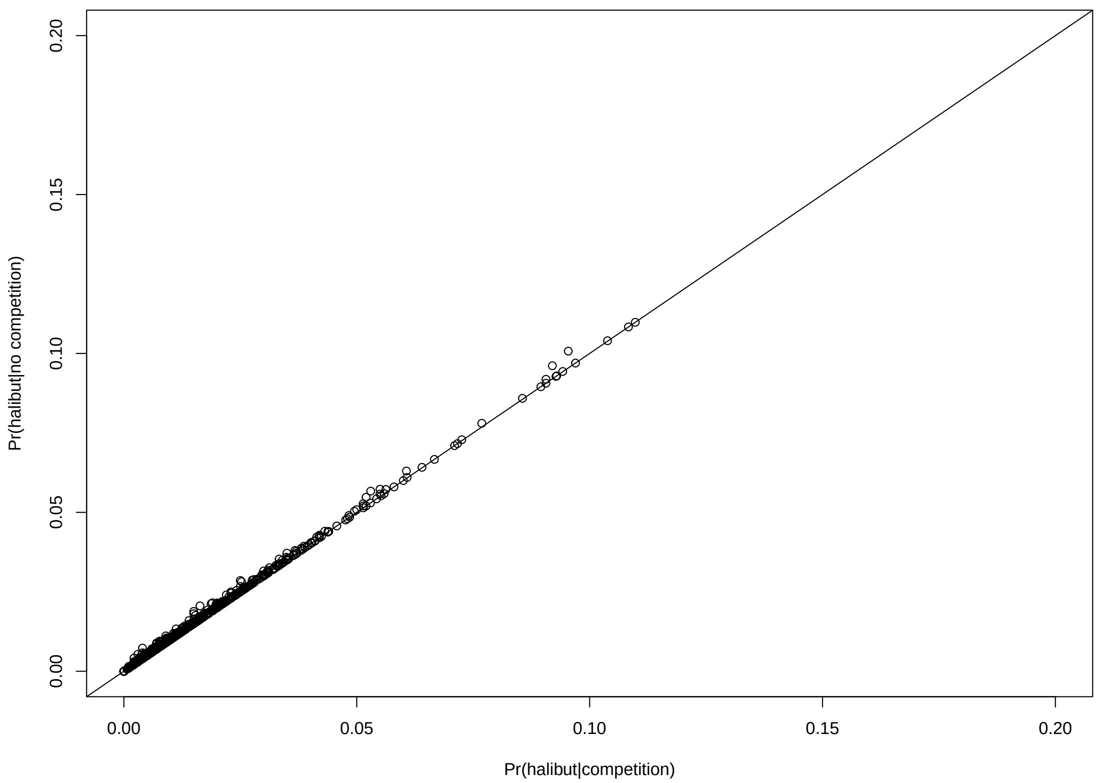

Figure 4. Probabilities of hooks with halibut by set over all years estimated for competition and no competition cases. Estimates were obtained assuming that all empty hooks were still baited upon recovery of gear. The 1:1 line is included for reference.

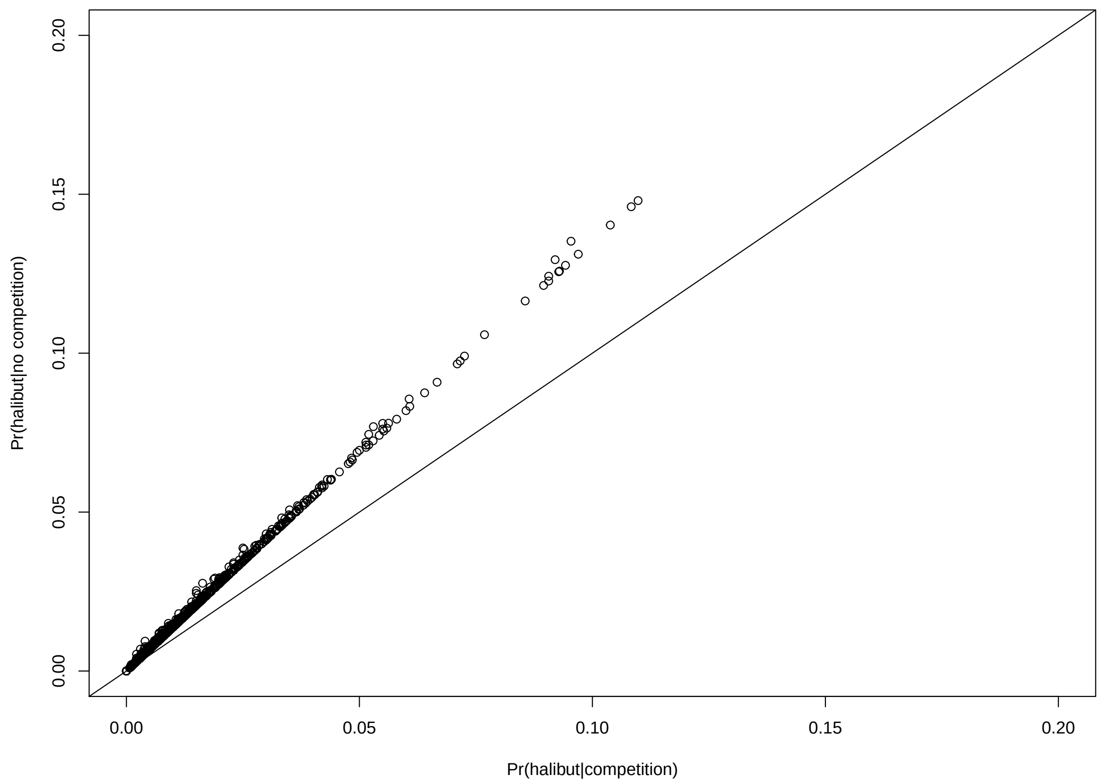

Figure 5. Probabilities of hooks with halibut by set over all years estimated for competition and no competition cases. Estimates were obtained assuming that 50 percent empty hooks were still baited upon recovery of gear with the unbaited hooks assumed to due to other species only. The 1:1 line is included for reference.

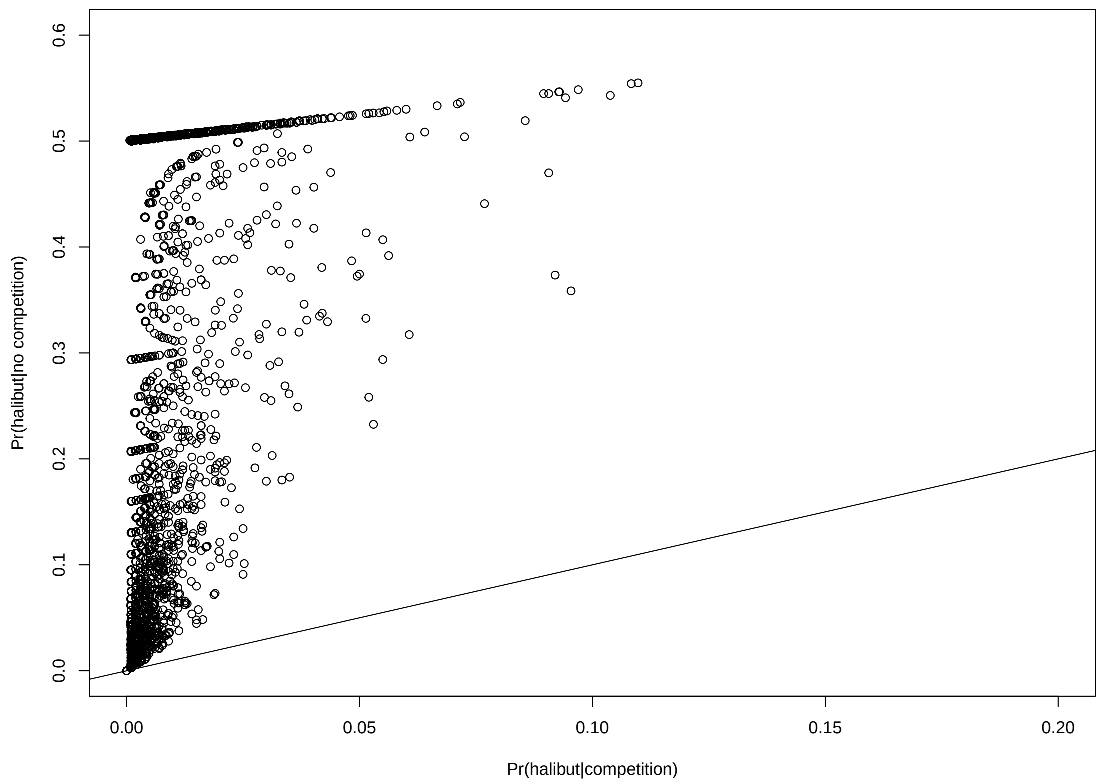

Figure 6. Probabilities of hooks with halibut by set over all years estimated for competition and no competition cases. Estimates were obtained assuming assume that 50 percent empty hooks were still baited upon recovery of gear with the unbaited hooks allocated to halibut and the other species by relative abundance. The 1:1 line is included for reference.

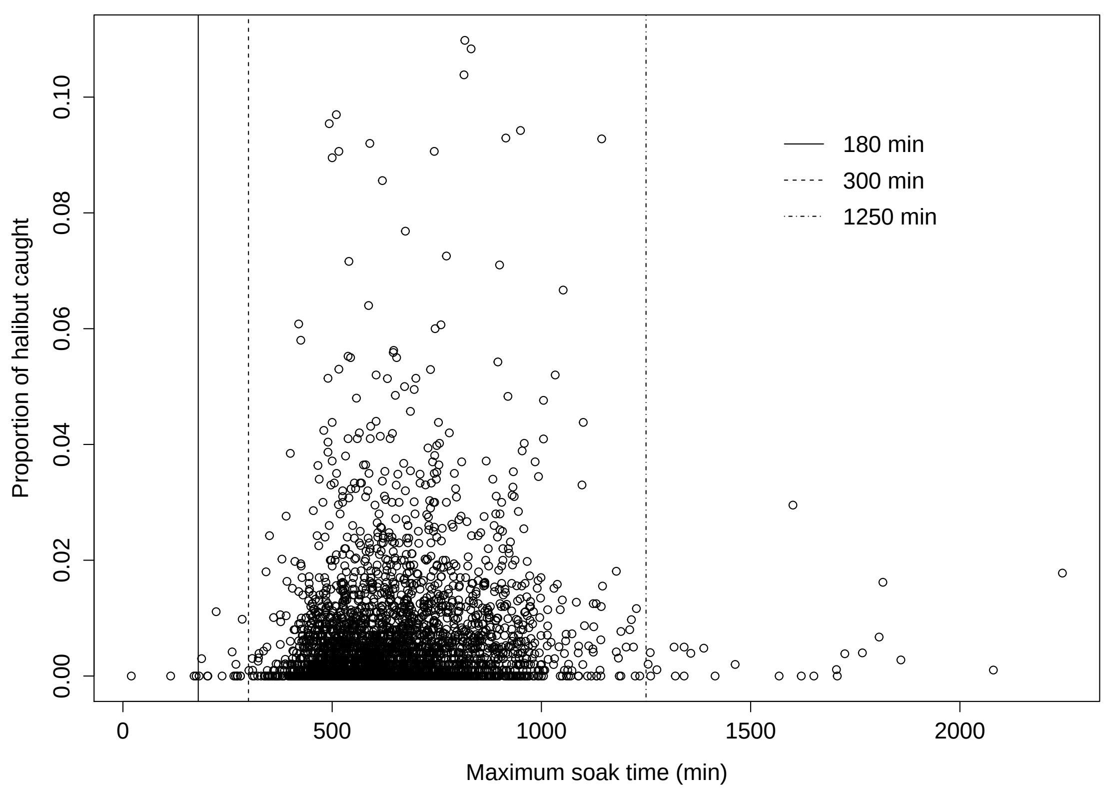

Figure 7. A comparison of the proportion of halibut caught for different soak times over all years from longline survey. The lower limit of 180 minutes was used by den Heyer et al. (2015), while 300 minutes is the lower limit used for the International Pacific Halibut Commission survey. The upper limit of 1250 minutes was used for the Multinomial model.

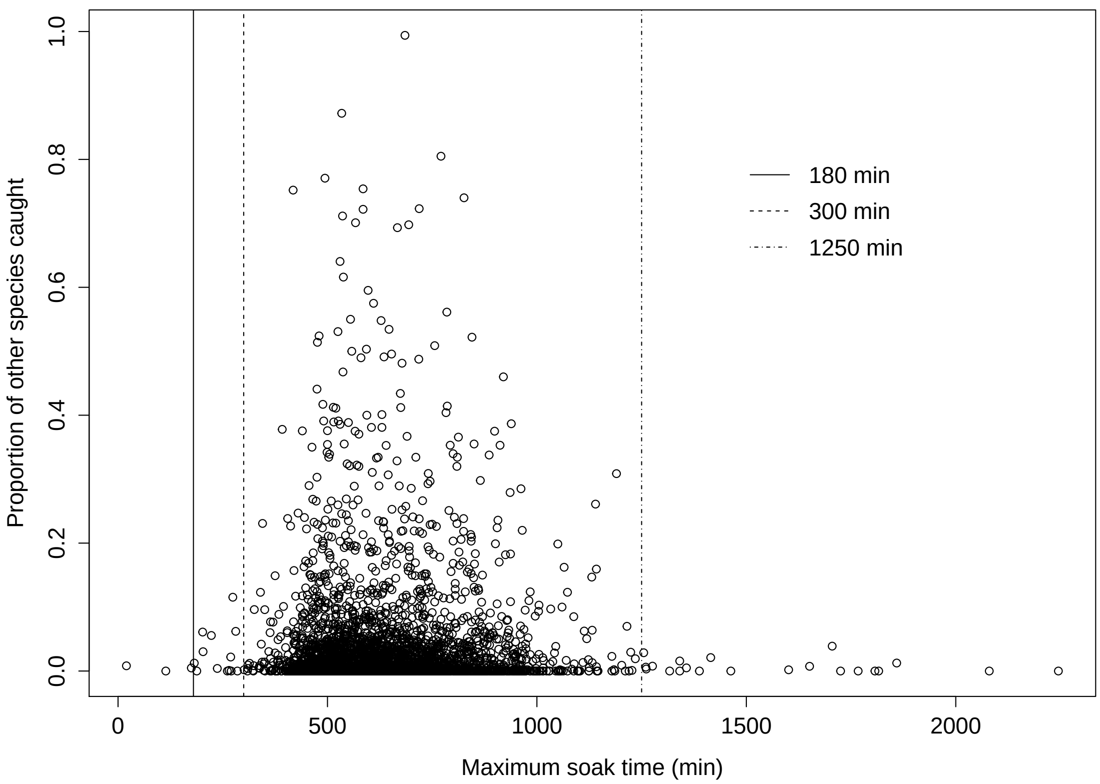

Figure 8. A comparison of the proportion of other species caught with soak times over all years from longline survey. The lower limit of 180 minutes was used by den Heyer et al. (2015), while 300 minutes is the lower limit used for the International Pacific Halibut Commission survey. The upper limit of 1250 minutes was used for the Multinomial model.

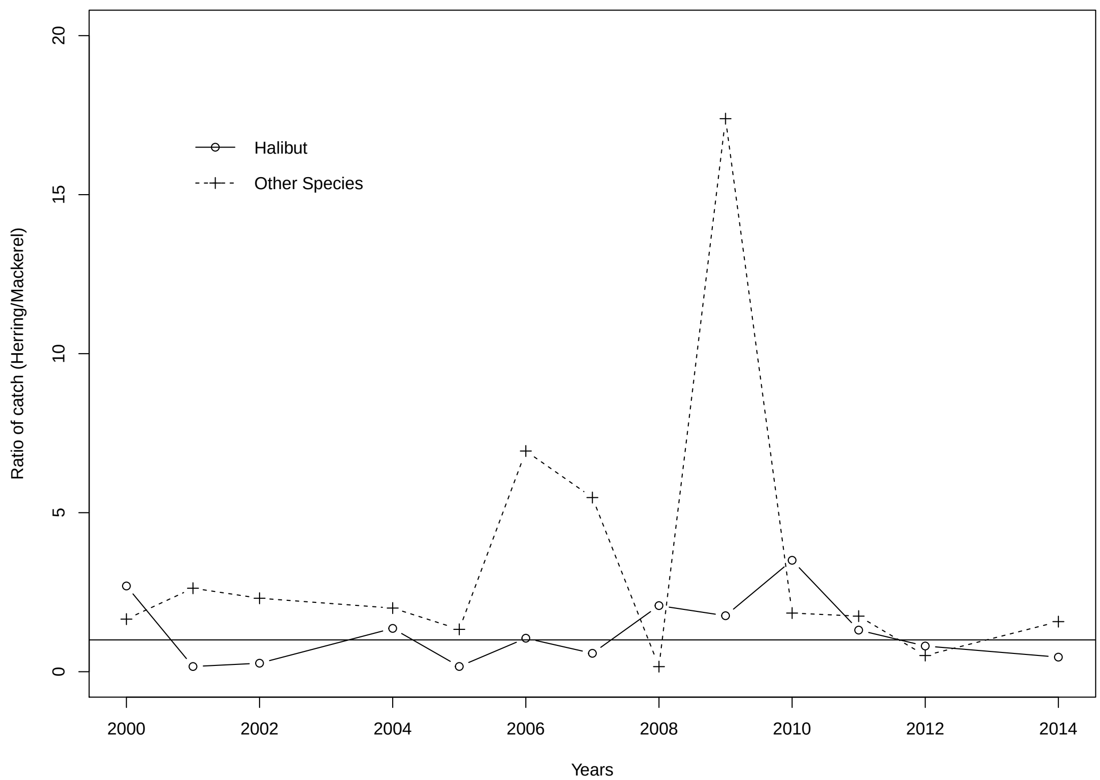

Figure 9. Predicted ratios of probabilities of catching halibut or other species using herring bait versus mackerel bait from Multinomial generalized linear model with main effects bait type and year, and an interaction term for bait type and depth. Data limited to sets in 4VWX.

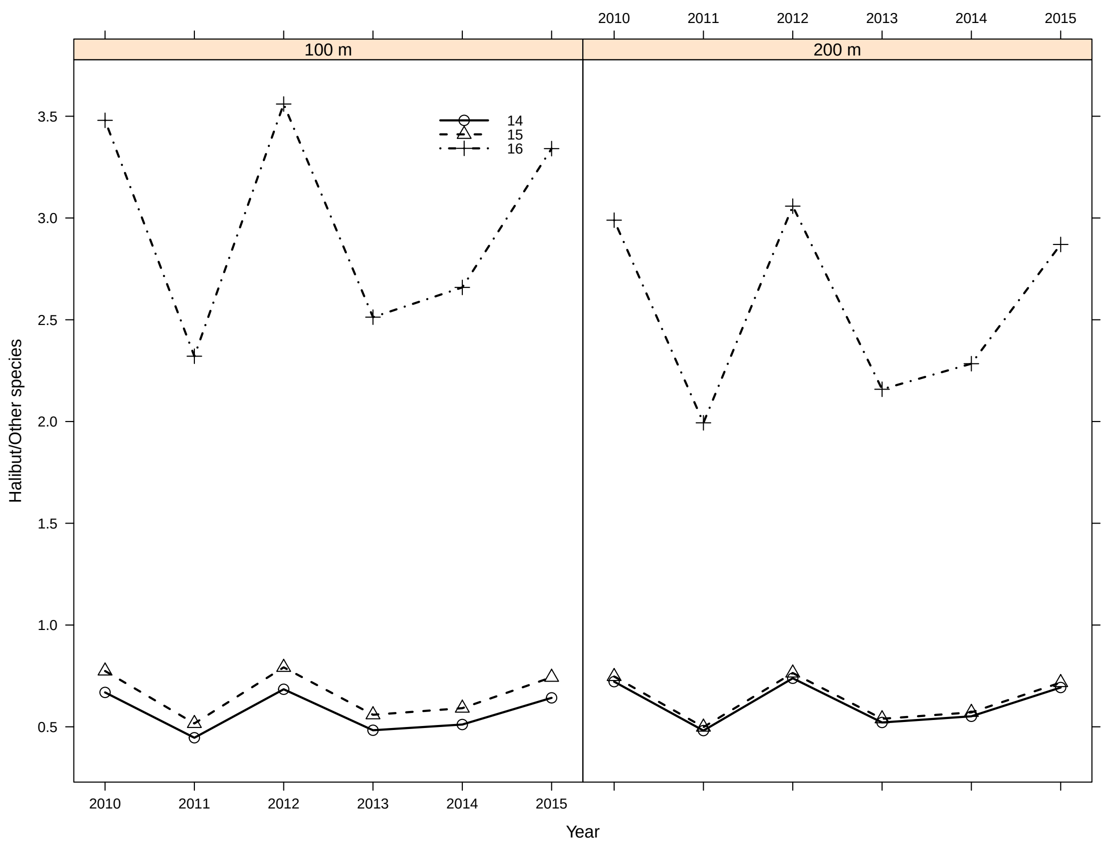

Figure 10. Predicted ratio of probabilities of halibut caught to other species caught for each hook size and two different depths from Multinomial generalized linear model with main effects hook size, depth and year, and an interaction term for hook size and depth. Data limited to sets in 4VWX from 2010 to 2015.

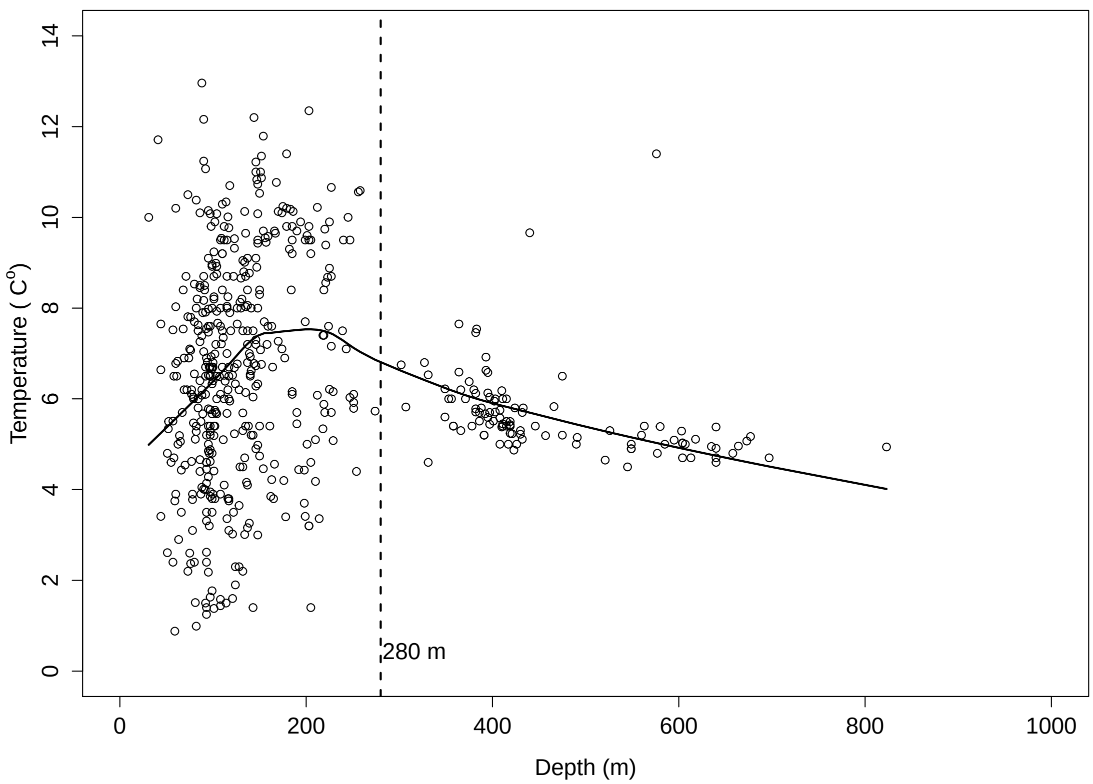

Figure 11. Distribution of temperature with depth for data limited to sets in in 4VWX from 2010 to 2015. Lowess smooth line added for reference. The generalized linear model was limited to sets less than 280 m in depth.

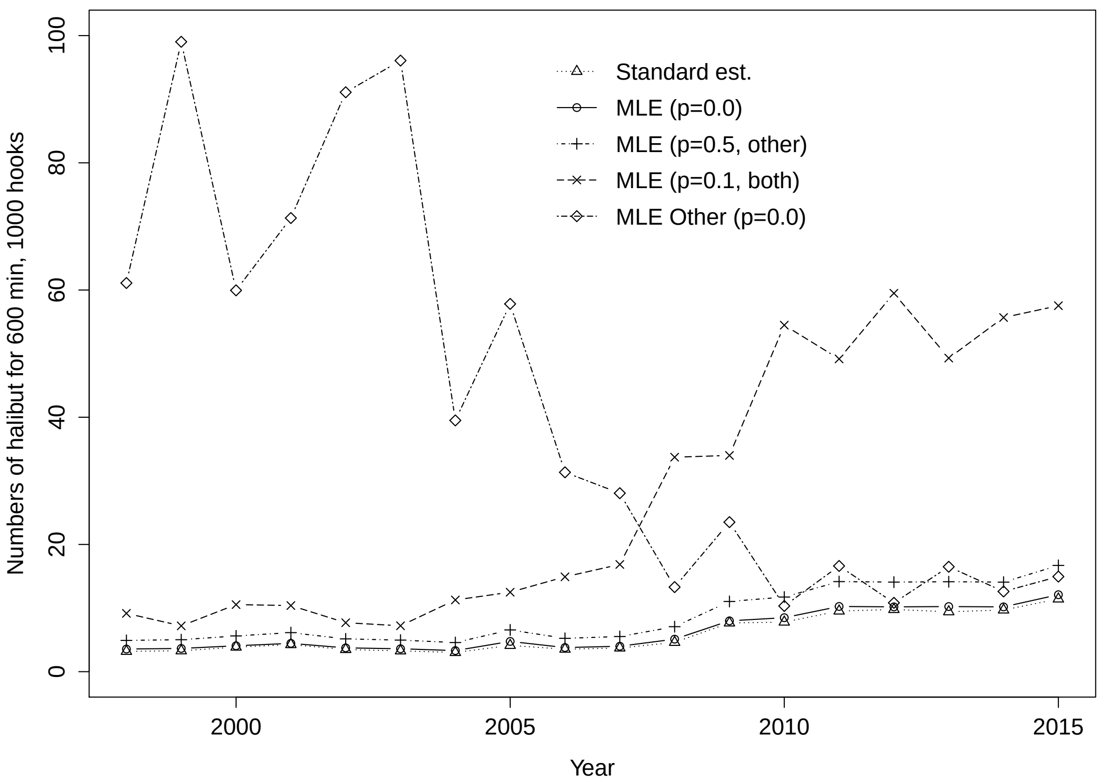

Figure 12. Annual estimates of the number of halibut caught per standard longline set using the standard estimate given in Armsworthy et al. (2006), the Multinomial Maximum Likelihood Estimate (MLE) for all empty hooks with bait (p = 0.0), 50 percent empty hooks unbaited and allocated to other species only (p = 0.50, other), and 10 percent empty hooks unbaited and allocated to either halibut or other species (p = 0.10, both). The MLE for the number of other species caught when all empty hooks were assumed to be baited (MLE other, p = 0.0) is also included.

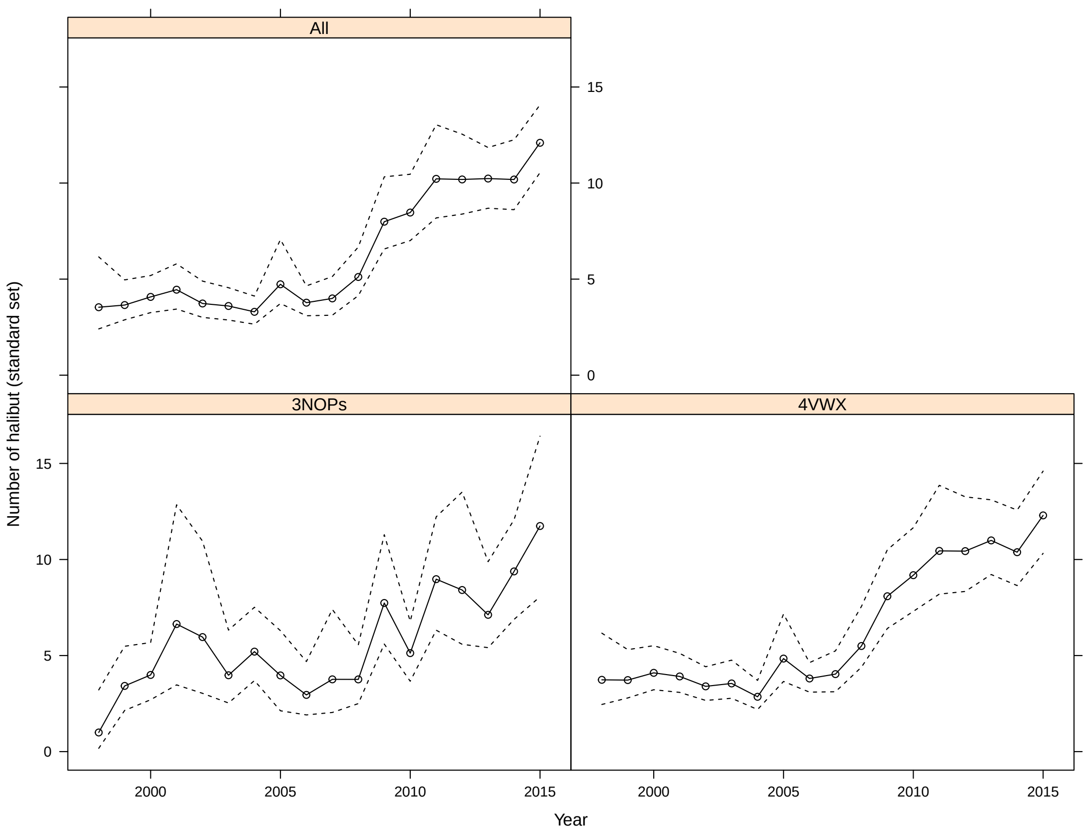

Figure 13. Annual estimates of mean number of halibut caught per standard set with 95 percent bootstrap confidence intervals.

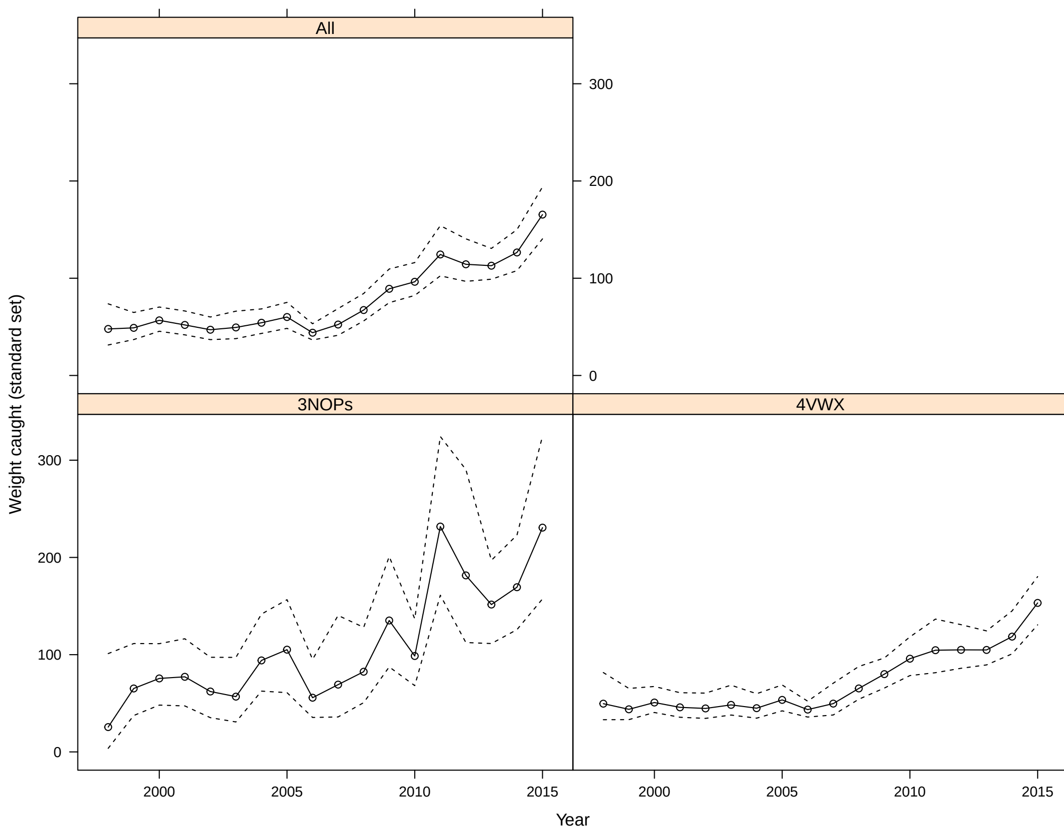

Figure 14. Annual estimates of mean weight of halibut per standard set with 95 percent bootstrap confidence intervals.

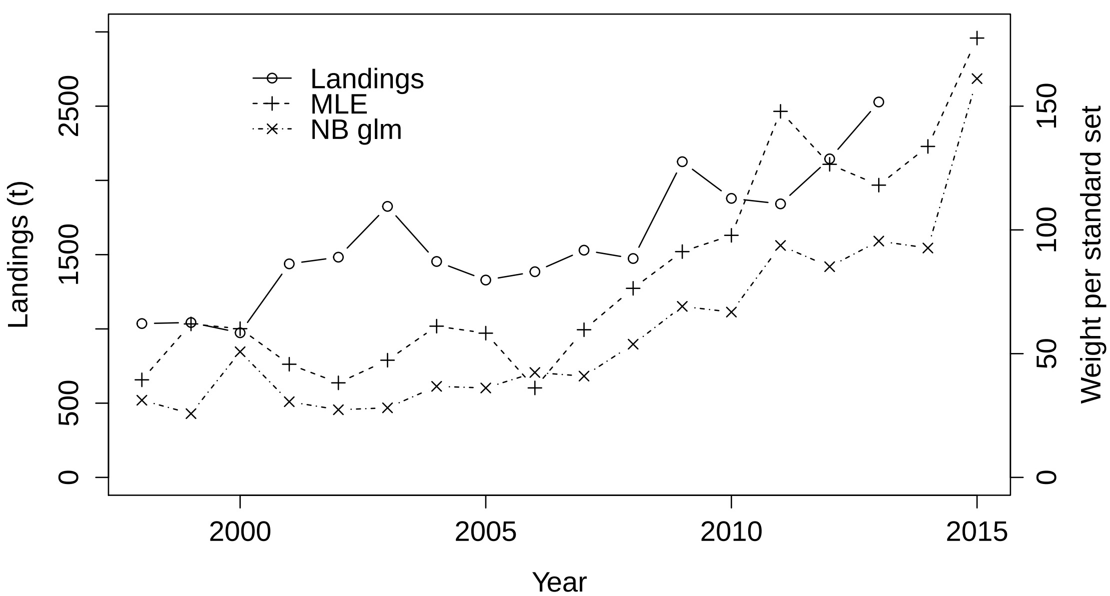

Figure 15. Comparison of annual landings of halibut (t) from 3NOPs4VWX5YZ with the annual Multinomial Maximum Likelihood Estimates (MLE) of mean weight (kg) per set assuming all empty hooks were baited and the Negative Binomial generalized linear model (NB GLM) standardized estimates of mean weight per set.

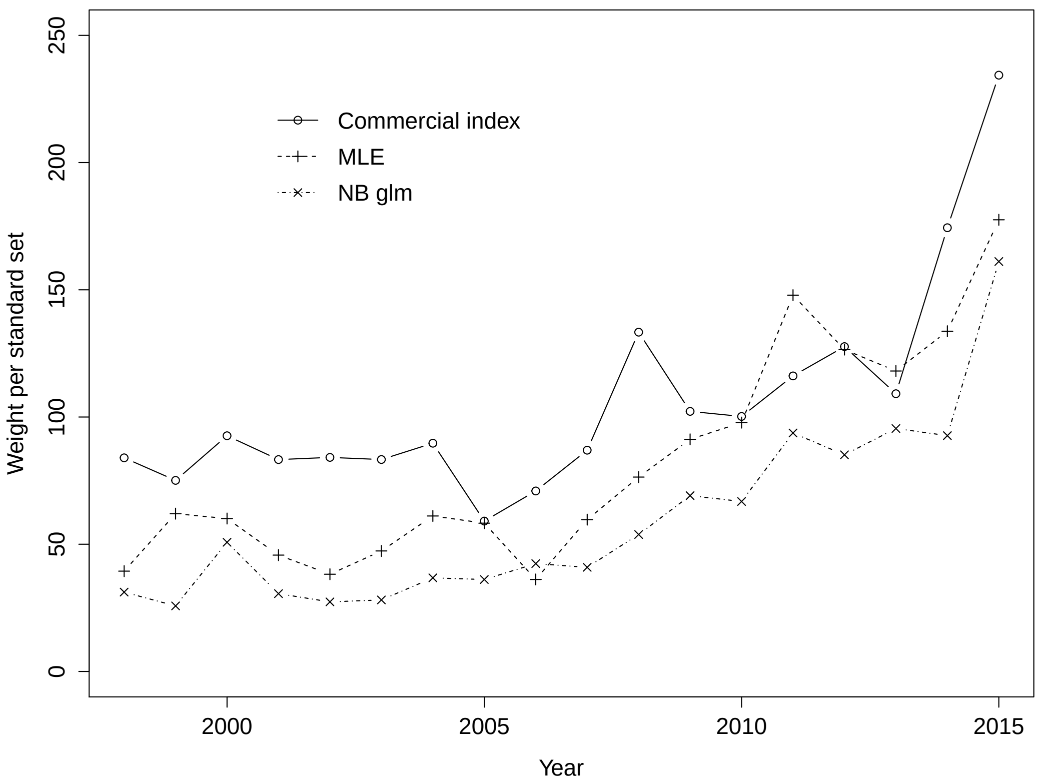

Figure 16. Comparison of the annual commercial index of weight (kg) per standard set from 3NOPs4VWX5YZ with the annual Multinomial Maximum Likelihood Estimates (MLE) of mean weight (kg) per set assuming all empty hooks were baited and the Negative Binomial generalized linear model (NB GLM) standardized estimates of mean weight per set.

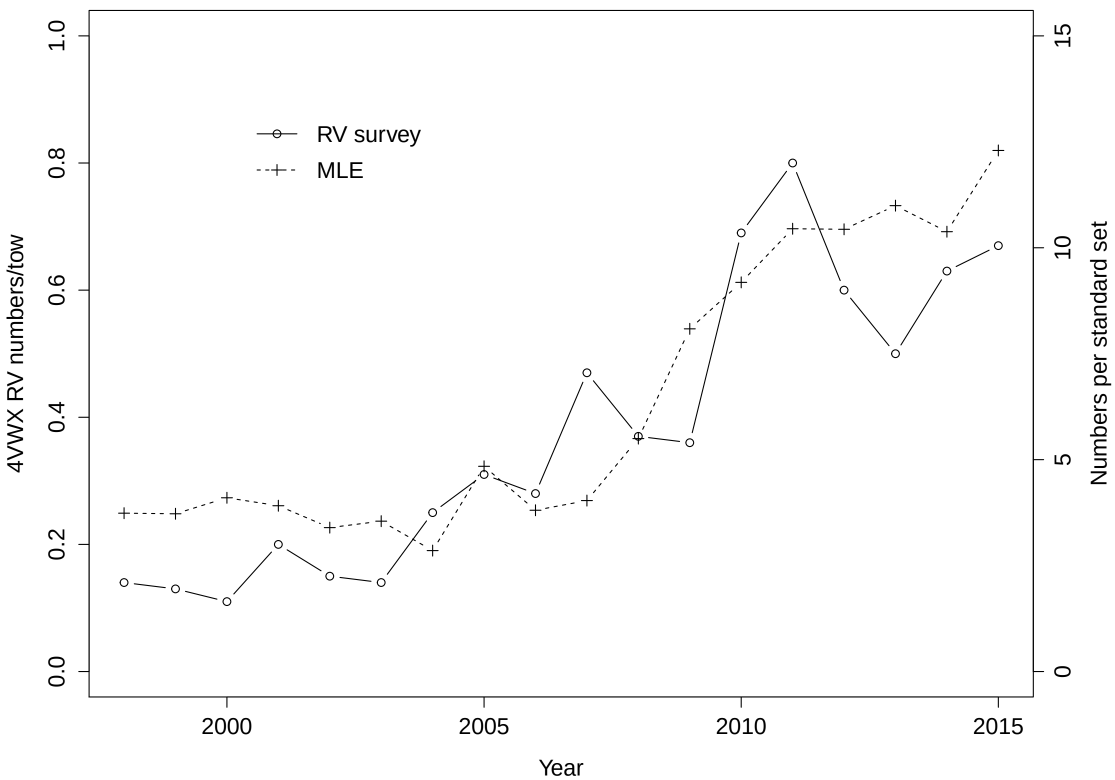

Figure 17. Comparison of the mean numbers of halibut per tow from the 4VWX research survey with the annual Multinomial Maximum Likelihood Estimates (MLE) of mean number per set assuming all empty hooks were baited.

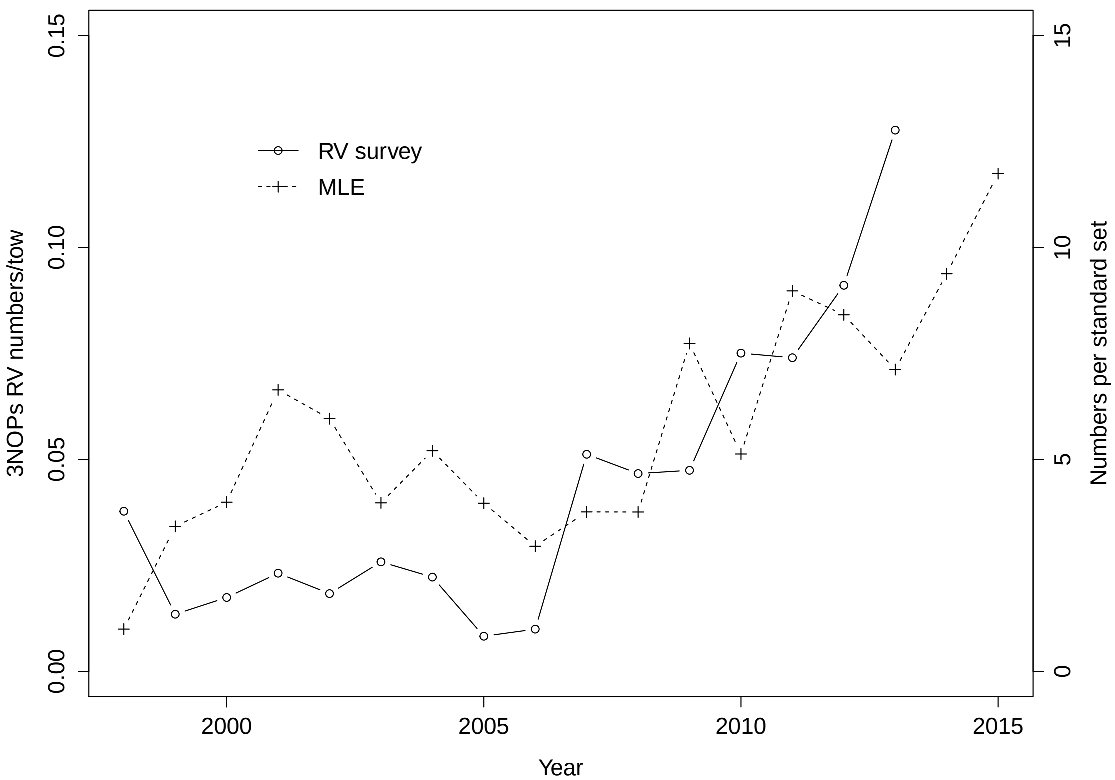

Figure 18. Comparison of the mean numbers of halibut per tow from the 3NOPs research survey with the annual Multinomial Maximum Likelihood Estimates (MLE) of mean number per set assuming all empty hooks were baited.

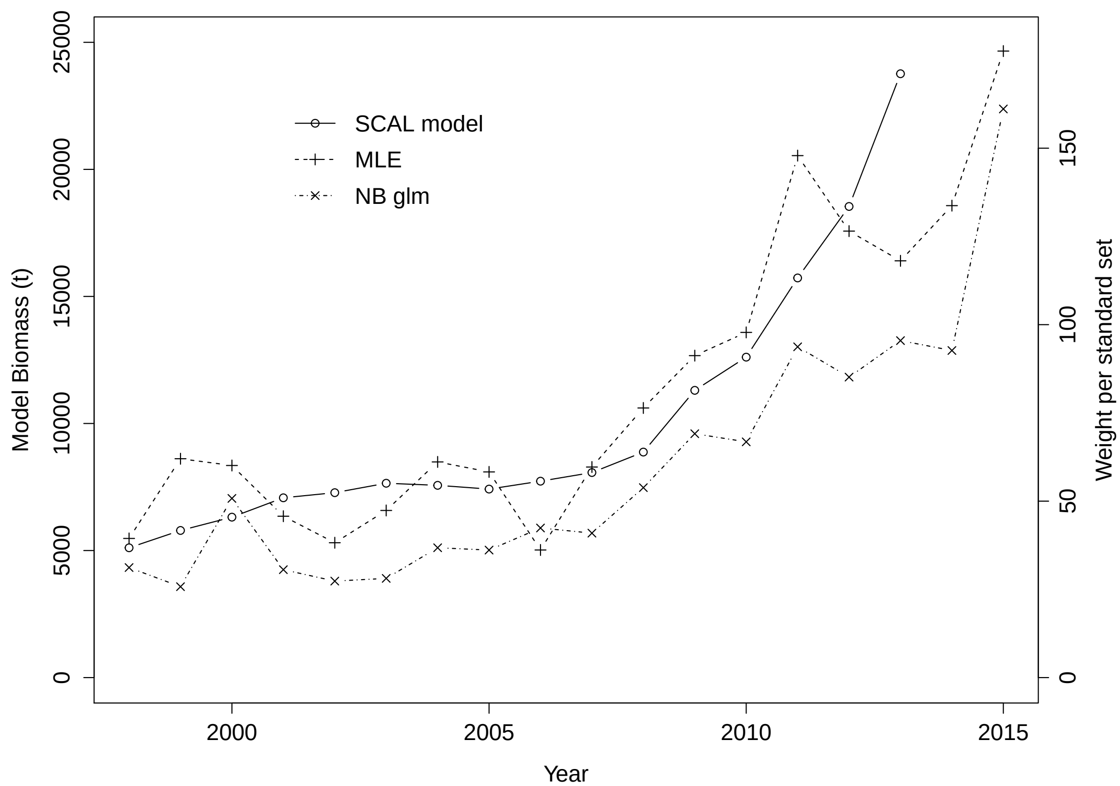

Figure 19. Comparison of the estimated population biomass (t) of legal size halibut from the Statistical Catch at Length (SCAL) model form Cox et al. (2016) with the annual Multinomial Maximum Likelihood Estimates (MLE) of mean weight (kg) per set assuming all empty hooks were baited and the Negative Binomial generalized linear model (NB GLM) standardized estimates of mean weight per set.