Population Assessment

Simulation studies of mark-recapture estimates

William G. Clark and Din G. Chen

Abstract

Simulated mark-recapture data were generated with a detailed model that included migration and variation in harvest rates among statistical areas. An estimation model that allowed for migration and variation in harvest rates only among regulatory areas was fitted to these data.

Introduction

During the summer of 2003 IPHC survey vessels released nearly 44 000 halibut coastwide with PIT tags inserted in the lower jaw. Of these nearly 29 000 were legal-sized and therefore retainable if caught by commercial halibut vessels. The main objective of the project is to estimate the harvest rate of legal-sized halibut by the commercial fishery.

Samplers in the ports scanned a large, known number of fish in the commercial landings in 2003 and 2004. There were only 86 recoveries during the rest of 2003 (Chen 2004), presumably because the fish were not very catchable for some period after being marked and released, so the harvest rate estimate will be based on recoveries during 2004.

The harvest rate to be estimated is the 2004 commercial harvest rate of fully vulnerable fish in each IPHC regulatory area, which according to the 2003 stock assessment (Clark and Hare 2004) will be fish with a fork length of about 120 cm or more. Smaller fish are expected to be less vulnerable to capture, so a length-specific selectivity function will have to be estimated for each area as well. Some fish will be recovered outside the area where they were released, so migration rates will also have to be estimated. The predicted number of recoveries will be the product of the harvest rate, selectivity, and migration rate estimates, so the estimates can be expected to interact strongly. In particular, the estimates may be very poorly determined if a lot migration occurs.

Simulation studies were conducted to investigate the accuracy and precision of the estimates. Simulated recoveries were generated by a simulation model that broke the 8 IPHC regulatory areas and the actual 2003 releases into 63 subareas. Migration was modeled as a diffusion process among subareas, and relative commercial fishing intensity in each subarea was set according to logbook data on the actual distribution of commercial effort. Length-specific selectivity differed among regulatory areas but was the same for all subareas within a given regulatory area.

Harvest rate, selectivity, and migration rate estimates for each regulatory area were computed from the simulated recoveries without using any knowledge of the processes at work at the level of subareas in the simulation model. That is, the estimation model effectively assumed that the parameter values were the same for all fish within a regulatory area whereas in fact the harvest rate, for example, did vary among subareas. This is always the case; we estimate the parameters of a model that is much simpler than the system that produced the observed data, and hope that the parameter estimates are close to the true average.

The combination of a complex simulation model and a simple estimation model made it possbile to investigate the effect of model misspecification on the estimates, along with the effects of migration and true harvest rate (which directly determines the expected number of recoveries).

Simulation and estimation models

The subareas used in the simulation were the original two-digit IPHC statistical areas for coastal waters of the Pacific Ocean. Bering Sea waters were divided into nine subareas. Migration among subareas was represented as a monthly diffusion of a certain proportion of fish from each subarea to contiguous subareas. The twelfth power of the implied migration matrix was used to calculate the subarea distribution in 2004 of the fish released in each subarea in 2003. With this migration process, fish released in subareas near the boundaries of regulatory areas had a higher probability of migration out of the regulatory area than fish released in the middle of a regulatory area, especially a big one like Area 3A. The same migration matrix was used for all legal-sized releases. The resulting true probabilities of migration between regulatory areas were:

	2004 distribution							
	2A	**2B**	**2C**	**3A**	**3B**	**4A**	**4B**	**4CDE**
Release area								
2A	0.84	0.16						
2B	0.01	0.90	0.09					
2C		0.10	0.86	0.04				
3A			0.01	0.92	0.07			
3B				0.07	0.90	0.03		
4A					0.05	0.81	0.07	0.07
4B						0.05	0.92	0.03
4CDE						0.18	0.19	0.63

An index of the relative commercial harvest rate in each subarea was calculated as the ratio of total commercial catch to total legal-sized survey catch (not CPUE) in the subarea. These index values were then rescaled so as to have an average of one, and were used as multipliers of a nominal overall harvest rate to calculate a harvest rate for each subarea.. The multipliers varied from near zero to almost 4. If the nominal overall harvest rate for a regulatory area was 0.2, the average over all subareas would also be 0.2, but the average over all fish would generally be different because there were different numbers of fish in the different subareas. For most regulatory areas, the average over fish was somewhat lower than the average over subareas.

As suggested by the 2003 assessment, length-specific selectivity was represented as a simple hockey-stick function, rising from zero at a length *L0* of 60-80 cm to 100% at a length *L100* of 110-130 cm. The same selectivity parameters were used for all subareas in a regulatory area, but the true average value for fish of a given length could be different because of the uneven distribution of fish (of a given length) and harvest rates among subareas. As a result, the average selectivity schedule for a whole regulatory area did not have exactly the underlying hockey-stick form.

Growth in length between 2003 and 2004 was taken to be 4 cm for females and 2 cm for males. For simulation purposes each fish released in 2003 was assigned a sex based on the sex ratio at length of the 2003 suvey catches. Tagging mortality, tag loss, and natural mortality combined were taken to be 0.2/y. The commercial catch sampling rate was taken to be 40%, and the detection rate 92%.

In the simulations, the probability of recovering a given fish in a given subarea in a given year (2003 or 2004) was computed for each fish according to the parameters of the simulation model. A multinomial sample of size one was then drawn for each fish to determine whether that fish would be recovered and if so, when and where. The recovery data for a single trial then consisted of the sample outcomes for all the released fish.

The recovery probablities in the estimation model were computed in exactly the same way as in the simulation model, but at the level of whole regulatory areas rather than subareas. (A minor difference was that in the estimation model the sex of fish was unknown so a single growth increment was used for all fish.) The parameter values in each trial were chosen so as to maximize the multinomial likelihood of the observed (simulated) recoveries.

The estimation model did use the same hockey-stick function as the simulation model to calculate selectivity at length, but as mentioned above length-specific selectivity does not have exactly this form at the level of regulatory areas anyway. In practice some curve of convenience will be used to approximate length-specific selectivity, and for purposes of investigating the estimates the hockey stick was as good as any other.

Effects of uneven harvest rates and migration on the estimates

The simulation model can be run with uniform harvest rates throughout each regulatory area and no migration. In this case the simulation and estimation models are the same, and the estimates are exact when the estimation model is fitted to the expected numbers of recoveries in each regulatory area. Uneven harvest rates in the absence of migration result in some deviations of harvest rate at length from the parametric hockey stick, but the full-recruitment harvest rate is well estimated when the estimation model is fitted to the expected numbers of recoveries (Fig. 1). The combination of uneven harvest rates and migration results in some cases of positive or negative bias, even when the estimation model is fitted to the expected number of recoveries. For example in Area 3B (upper right panel of Fig. 2), the best fit of the estimation model to the expected recoveries underestimates the true harvest rate because at the same time it underestimates emigration from Area 3B. It thinks more of the 2003 releases in Area 3B are still there in 2004 than truly are, so it can generate the observed recoveries with a lower harvest rate. This sort of bias occurs because migration and harvest rates are not uniform within regulatory areas as the estimation model assumes.

Effect of random variation on the estimates

The actual number of recoveries at a given length in a simulated experiment will generally be more or less than the expected value because of random variation in the outcomes of individual fish. Where the expected number of recoveries is large, these random effects will tend to be swamped, but they can have a large effect where expected recoveries are few. Figures 3a-c show the estimates of harvest rates in Areas 2B, 3A, and 4A respectively obtained in nine random trials. The Area 3A estimates are quite well behaved; the 2B estimates less well-behaved, and the 4A estimates wild in some cases.

Bias and variance of point estimates in simulations

In most areas the bias due to model misspecification (i.e., the difference between the true harvest rate and the value estimated by fitting to the expected number of recoveries) is small—less than 10%—but in a couple of areas it reached 15% (Table 1). The coefficient of variation of the harvest rate estimates is around 10% in Areas 3A and 3B and 15-20% elsewhere except for Area 2A, where it is so large as to make the estimate nearly useless. The coefficient of variation refers to the standard deviation of the estimate about the mean value rather than the true value.

The values shown in Table 1 refer to trials in which *L0* was constrained to lie between 60 and 80 cm and *L100* between 110 and 130 cm. Other estimation schemes were investigated, including:

(i) Estimating the harvest rate from recoveries of only those fish 120 cm or larger at release.

(ii) Constraining *L100* to lie between 110 and 120 cm, or fixing it at 120 cm.

(iii) Fitting a logistic selectivity function rather than the hockey stick.

None of these alternatives provided substantially less variable estimates than the method tabulated, and using only the data on large fish produced substanitially more variable estimates.

Variance estimates

The actual estimates will be computed by fitting the estimation model to the actual recovery data. The variances of the estimates would normally be taken from the inverse of the Hessian matrix at the maximum likelhood estimate. Because of the strong interaction among the parameters, however, the maximum is not always sharply defined and the Hessian-based variances can be quite variable. The simulations also include a few cases where the Hessian is not positive definite at the numerically located maximum, so a Hessian-based variance is not even feasible.

A robust alternative method is to bootstrap the data, i.e. to generate many trials by resampling the actual release/recovery data and calculate the variance of the harvest rate estimates among those trials. Simulations of this procedure show that the bootstrap variance estimates are on average correct, but they too are variable to some extent. Just as a given realization of the mark-recapture experiment can produce a point estimate higher or lower than the true value, it can produce a bootstrap variance estimate higher or lower than the true value. In simulations, the coefficient of variation of the bootstrap estimate of the standard deviation of the harvest rate estimate was 10-25% (Table 2).

Discussion

These simulations were carried out to determine the effects of uneven fishing intensity and migration on mark-recapture estimates of the 2004 harvest rate. The details of the simulation model are not important; one could postulate a variety of alternative detailed models of the migration and recapture processes. The results show that while the estimates may be slightly biased, they should be serviceable.

When the 2004 data are available at the end of the season, the staff will compute the point estimates and variance estimates of the harvest rates, and attempt to construct a confidence interval that is very likely to include the true value. Typically a confidence interval (CI) has the form:

$$CI = (\text{point estimate}) \;\pm\; 2 \times (\text{estimated standard deviation})$$

For the mark-recapture estimates, we will have to recognize that the point estimate may be biased because of model misspecification, and the point estimate of the standard deviation may be low. There will be no way to estimate either of these from the data themselves, but the simulations provide some guidance for making a reasonable allowance for them. In particular, we could allow for a bias of 10% in the point estimate and a CV of 20% in the estimated standard deviation. This would produce a confidence interval of the form:

$$CI = (\text{point estimate}) \;\pm\; (\text{bias}) \;\pm\; 2 \times (1.4) \times (\text{estimated standard deviation})$$

References

Chen, D.G. 2004. Preliminary analysis of the IPHC PIT tagging experiment in 2003. Assessment of the Pacific halibut stock at the end of 2003. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2003:377-390.

Clark, W.G., and S.R. Hare. 2004. Assessment of the Pacific halibut stock at the end of 2003. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2003:171-200.

Table 1. Bias and variance of 2004 commercial harvest rate estimates in simulations. The value estimated is the harvest rate of fully ruecruited fish in each area. In all areas the average harvest rate among subareas was 0.20. The true value shown is the average among fish resulting from the distribution of releases and fishing intensity among subareas within each area. "CV" is coefficient of variation.

	Area 2A	Area 2B	Area 2C	Area 3A	Area 3B	Area 4A	Area 4B
True rate	0.201	0.189	0.152	0.183	0.212	0.184	0.252
Estimated using expected number of recoveries	0.232	0.171	0.166	0.185	0.182	0.213	0.241
CV of estimate	0.522	0.166	0.186	0.085	0.120	0.216	0.190

Table 2. CV (coefficient of variation) of the bootstrap estimate of the standard deviation of the harvest rate estimate.

Area 2A	Area 2B	Area 2C	Area 3A	Area 3B	Area 4A	Area 4B
0.173	0.171	0.179	0.103	0.178	0.251	0.240

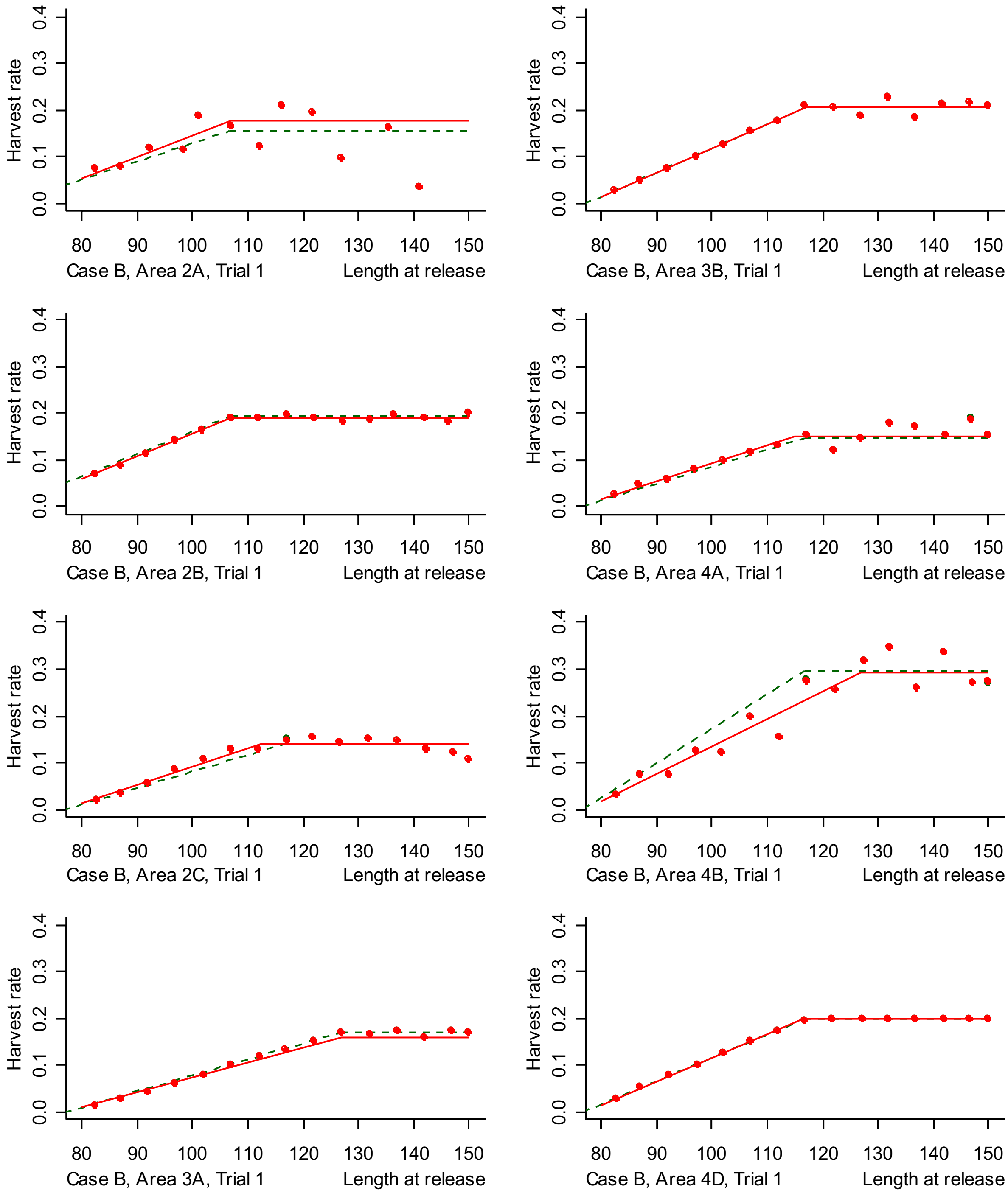


Figure 1. Estimates of harvest rate as a function of length (by 5 cm interval) when harvest rate varies by subarea within regulatory area but there is no migration and no sampling variance (numbers recovered are expected values). True values are dashed line; estimates are solid line.

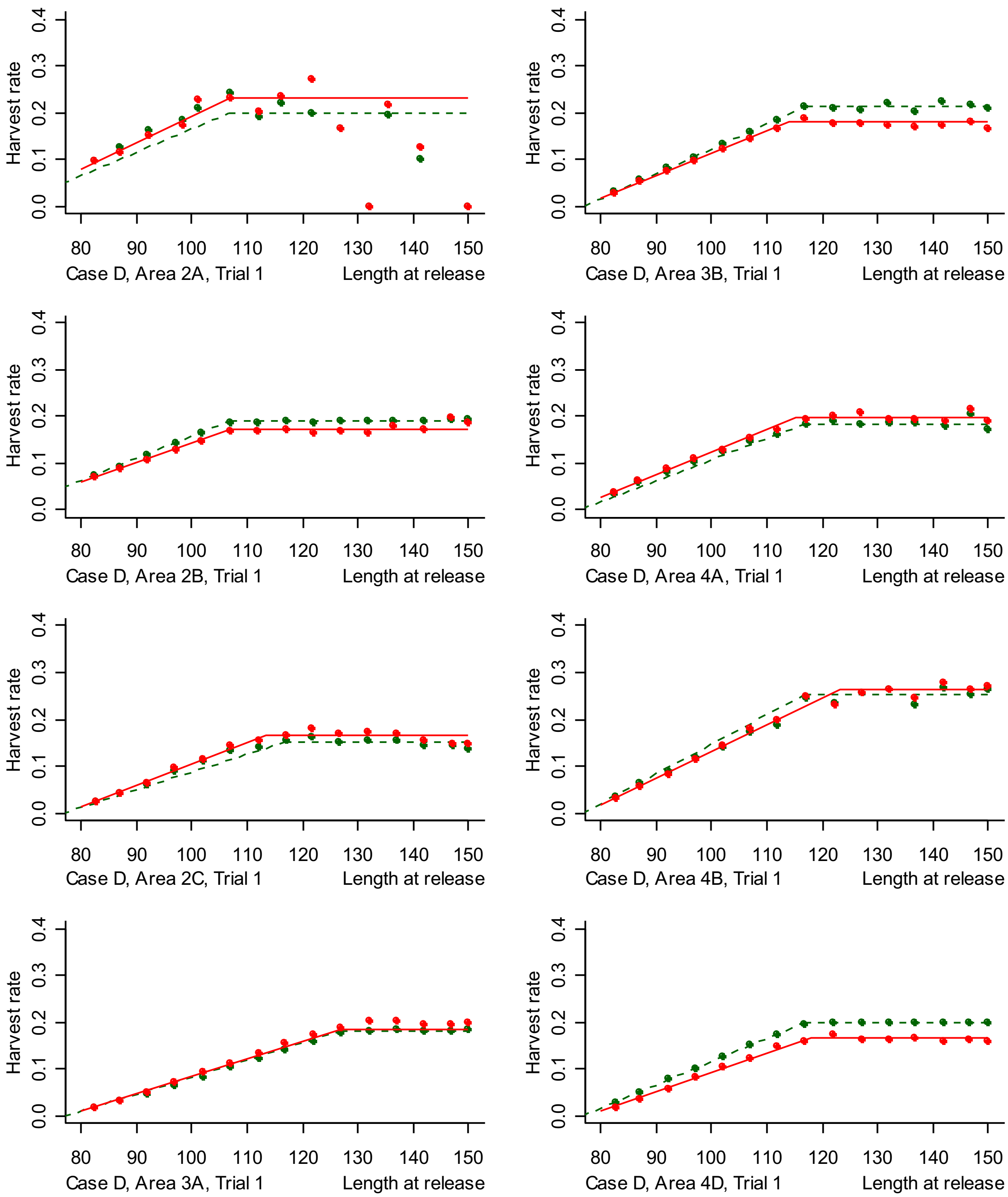


Figure 2. Estimates of harvest rate as a function of length (by 5 cm interval) when harvest rate varies by subarea within regulatory area and there is migration but still no sampling variance (numbers recovered are expected values). True values are dashed line; estimates are solid line.

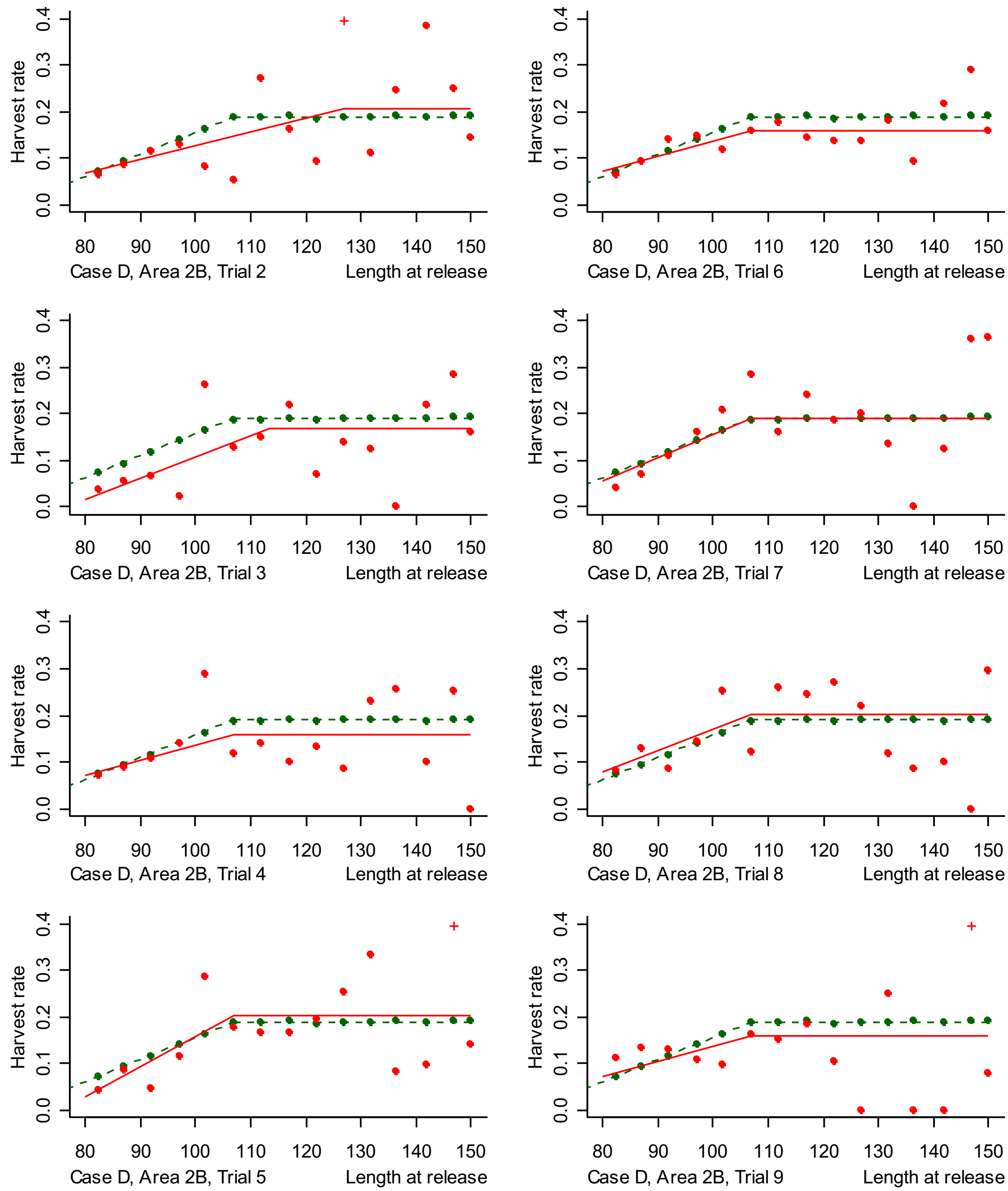


Figure 3a. Estimates of harvest rate in Area 2B as a function of length (by 5 cm interval) when harvest rate varies by subarea within regulatory area and there is migration and sampling variance (numbers recovered are sample outcomes). True values are dashed line; estimates are solid line. Points along dahsed line are expected values; other points are sample values.

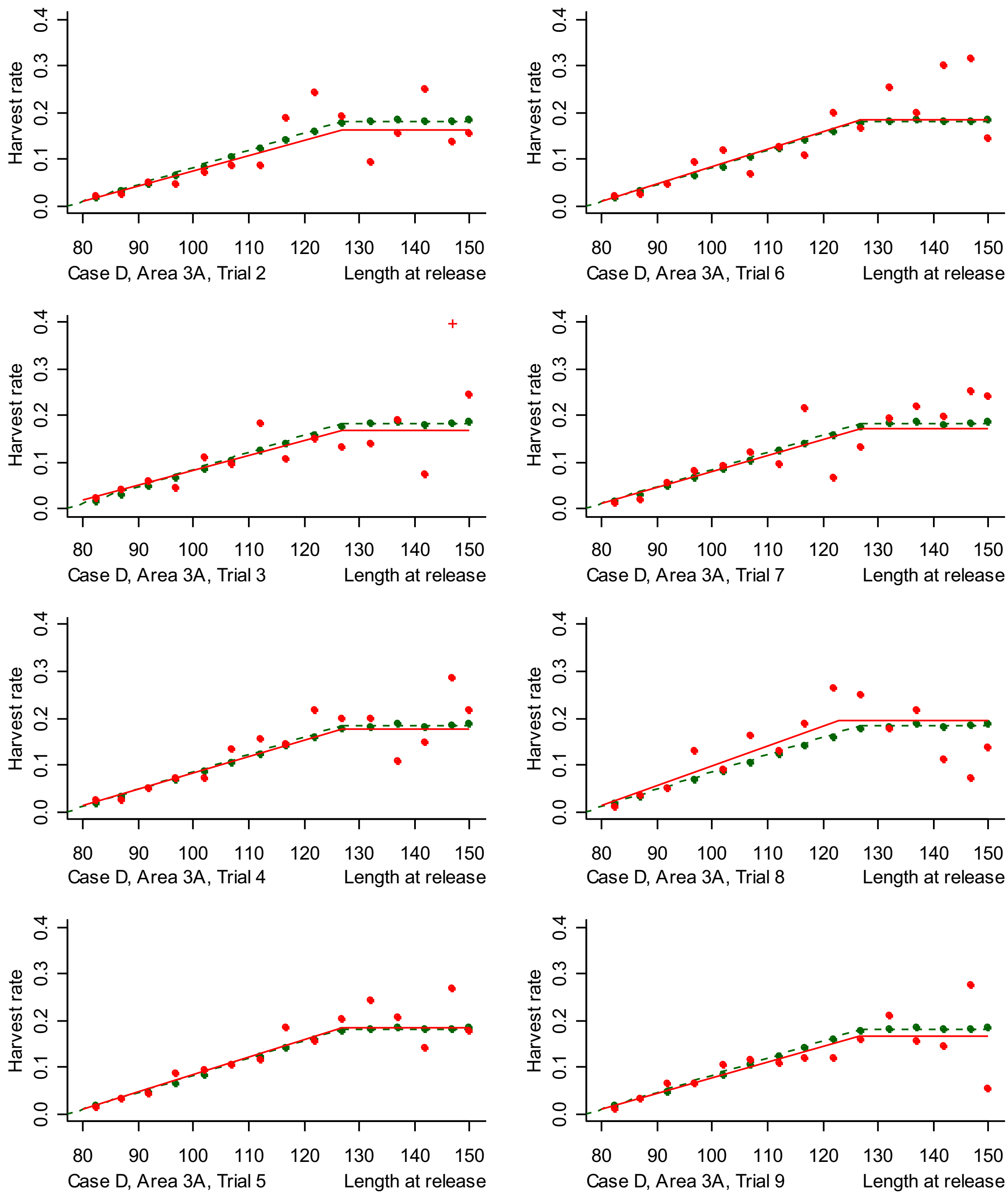


Figure 3b. Estimates of harvest rate in Area 3A as a function of length (by 5 cm interval) when harvest rate varies by subarea within regulatory area and there is migration and sampling variance (numbers recovered are sample outcomes). True values are dashed line; estimates are solid line. Points along dahsed line are expected values; other points are sample values.

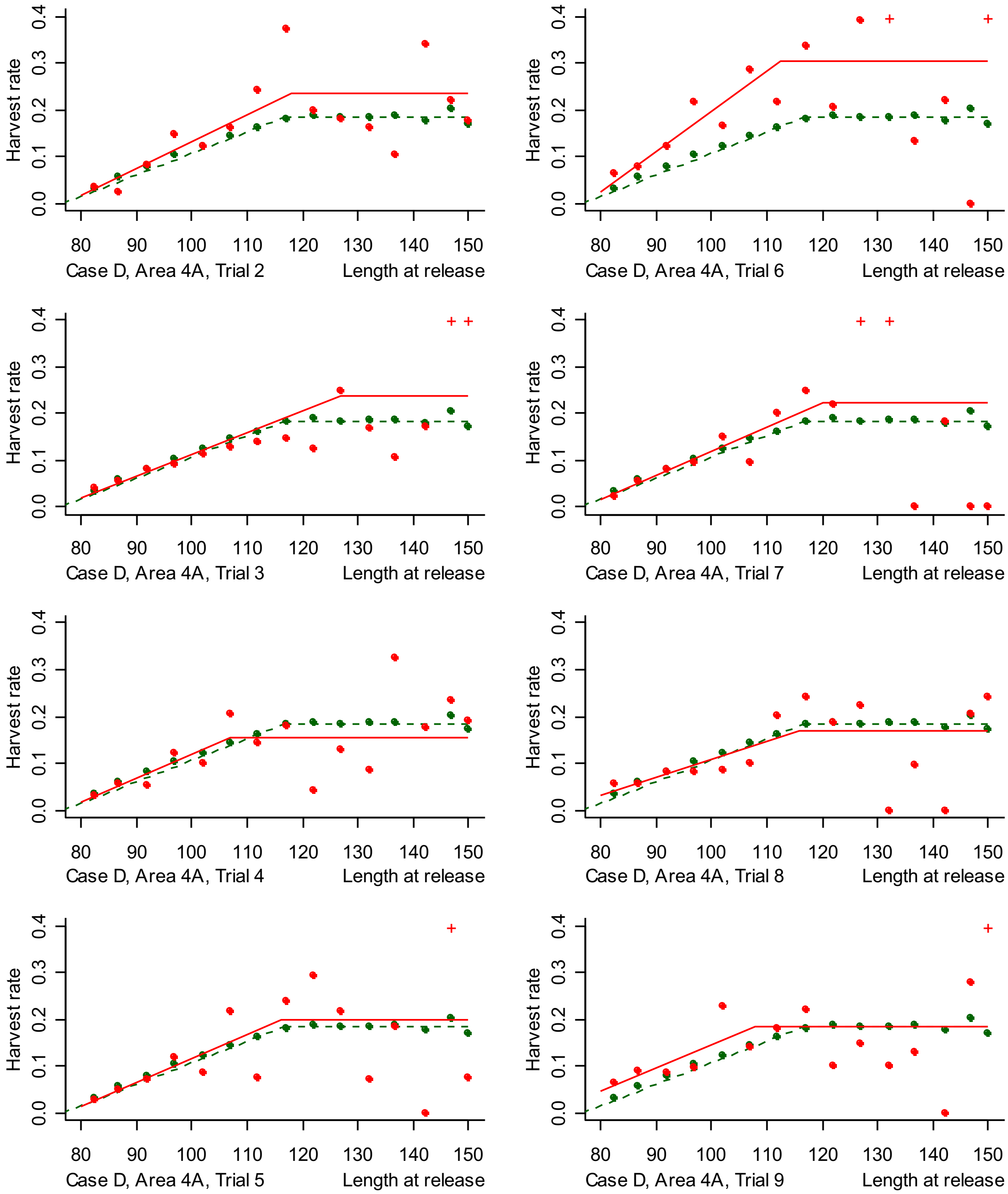


Figure 3c. Estimates of harvest rate in Area 4A as a function of length (by 5 cm interval) when harvest rate varies by subarea within regulatory area and there is migration and sampling variance (numbers recovered are sample outcomes). True values are dashed line; estimates are solid line. Points along dahsed line are expected values; other points are sample values.

Estimates of length-specific commercial selectivity from historical marking experiments

William G. Clark

Abstract

While not reliable for estimating exploitation rates or migration rates, the data from the numerous marking experiments of the 1960s, 1970s, and 1980s are usable for estimating length-specific commercial selectivity, and they provide large datasets for Areas 2B, 2C, and 3A. The data are much fewer for Area 3B and insufficient to make any estimates for Area 4. In Area 2B, estimated length-specific commercial selectivity peaks at about 110 cm and then declines substantially. In Areas 2C, 3A, and 3B it peaks at about 150 cm and then declines slightly.

Background

For a number of years the staff regarded setline selectivity as being a function of age rather than length, but a growing body of evidence reversed that view in 2003 (Clark and Hare 2004). We can estimate length-specific commercial and survey selectivity in the stock assessment, but to do so we need to specify the correct form for the relationship. We also need an estimate of commercial selectivity to interpret the results of the 2003/2004 PIT tag experiment, and for that purpose it now appears likely that apart from Area 3A the recoveries will not be sufficient to determine even the form of the relationship, much less the parameter values.

For both purposes an independent estimate of the form of the relationship would be useful. This paper reports estimates based on the large body of data from historical mark-recapture experiments that employed external tags. A similar analysis was done by Myhre (1969), but he used data from only two experiments, one in Area 2B and the other in Areas 3A and 3B. The present study uses data from more than 100 experiments.

Materials and methods

Kaimmer (2000) describes all IPHC releases of various kinds of external tags from both setline and trawl catches dating back to 1925. The present study uses setline releases of all tag types (except the small strap) dating back to 1960, which is the first year in the computer database. The total number of releases was over 100 000, of which more than 13 000 were recovered. About half of the releases were at systematically placed setline survey stations that covered a large part of a regulatory area. The other half were at spot fishing locations deliberately chosen to produce good catches, either for marking or for gathering data on the performance of different gear types. For this study an experiment was defined as all releases of a given tag type in a given IPHC regulatory area in a given year in either survey or spot fishing operations (not both), which in some cases had the effect of grouping experiments as they appear in the database. Between 1960 and 1990 there were 131 such experiments in which at least 10 fish were released.

These data are not usable for estimating exploitation rates or migration rates because of uncertainty concerning things like recovery effort and reporting rates, but they can be used to estimate commercial selectivity. In the case of a single experiment, a straightforward plot of short-term recovery rate by length at release will show how selectivity changes with length. The absolute recovery rates will depend on usually unknown factors (tagging, fishing, and natural mortality rates; tag loss and reporting rates) but the relative recovery rates should depend mainly on selectivity (barring large variations with length in any of the unknown factors).

Figure 1 shows plots of short-term recovery rate by 10 cm length interval for all single experiments that yielded at least 100 recoveries in the year of release and the two years thereafter. The length shown on the abscissa is the midpoint of the 10 cm release length interval plus a 5 cm allowance for growth, so it is the approximate mean size of the fish in that release length interval during the two years after release. For example, the recovery rate of fish in the release length interval 80-90 cm is plotted at a length of 90 cm with the idea that the fish were 80-90 cm in the year of release, 85-95 cm the first year after, and 90-100 cm the second year after.

Myers and Hoenig (1997) show how data from many experiments can be combined to obtain a single set of selectivity estimates. Summarizing their derivation, let $\pi_{i,l}$ be the recovery rate of fish of length l in experiment i. This rate is treated as the product of a length-specific commercial selectivity s_l, which is the same for all experiments, and an experiment-specific recovery rate r_i that combines all the unknown factors mentioned above. Thus $\pi_{i,l} = r_i \cdot s_l$ and $\log \pi_{i,l} = \log r_i + \log s_l$. This has the form of a generalized linear model with a log link function and a binomial variance, so the point estimates and variance estimates can be obtained in standard fashion.

Some rule has to be chosen for scaling the selectivities to make the model determinate. The most common rule is to require that the maximum be 1.0, but that can involve using a scaling factor that is poorly determined by the data if the maximum occurs in a length group with few releases and recoveries. To avoid that, the rule used here was to define selectivity to be 1.0 at 120 cm. Estimated selectivity could therefore exceed 1.0 at other lengths.

Results

Figure 2a shows the estimates of commercial length-specific selectivity in Areas 2B, 2C, 3A, and 3B obtained by the method of Myers and Hoenig (1997) using all available data in each area. There were insufficient data in Area 4 to calculate useful estimates. The estimates in Figure 2a were calculated using all recoveries from each release, including recoveries from outside the release area and recoveries from unknown locations. Estimates computed using only recoveries from the area of release were no different from those obtained using all of the recoveries.

Similar estimates were calculated with subsets of the data to see if different selectivity patterns occurred before and after the adoption of the 32 in (81 cm) size limit in 1974, or before and after the change from J hooks to C hooks in 1983. In all cases the pattern in each area was similar to that shown by the entire dataset. The largest difference appeared in Area 2B, where the selectivity curve before the increase in the size limit, while still dome-shaped, was a bit flatter than in the years with the higher size limit (Fig. 2b). But even in this case the difference was slight.

Recoveries from releases at spot fishing locations show a selectivity pattern similar to the entire dataset (Fig. 2c). The same is true of survey releases and recoveries except in Area 2B, where the selectivity pattern does not show a decline among larger fish (Fig. 2d). But this impression depends entirely on the last two length intervals where there were only 14 recoveries in total, so it may be a false impression.

In all areas, commercial selectivity in the 1960-1990 period appears to have increased with length up to a maximum and then declined. In Area 2B the peak occurs at about 110 cm and there is a substantial decline thereafter, to around half the peak value. In Alaska (Areas 2C, 3A, 3B) selectivity peaks at a much larger size—150 cm or so. Thereafter the decline is about as steep as in Area 2B but not as large because so little of the length composition remains beyond 150 cm.

Discussion

In modeling length-specific selectivity heretofore, the staff has generally assumed some kind of asymptotic function, with full selection occurring at 110-130 cm. This is consistent with video observations of halibut hooking behavior (Kaimmer 1999), and it has produced satisfactory fits to the observed length composition of survey and commercial setline catches (Clark 1993, Clark and Hare 2004). The large body of mark-recapture data shows a different pattern: selectivity declining after the peak at 110 cm in Area 2B, and not reaching a peak until 150 cm or so in Alaska. These patterns are in fact quite similar to those reported by Myhre (1969).

While an asymptotic function does not appear to be the correct choice, one can see (particularly in Fig. 2b) that the asymptotic functions fitted in the assessment would approximate the dome-shaped selectivity patterns quite well. For Area 2B, the assessment estimates a linear increase in selectivity up to 110 cm and full selectivity thereafter, which can adequately approximate the pattern estimated here for all but the very largest fish (which are very few). Similarly in Area 3A, the assessment estimates a linear increase up to 130 cm and full selectivity thereafter, which again can adequately approximate the observed dome-shaped pattern.

Commercial selectivity reflects ground selection by the fleet as well as size selection by the gear. It is therefore possible that the selectivity of commercial setline gear on a given ground has the expected asymptotic form, but ground selection has the effect of targeting certain size groups and thereby producing a different selectivity schedule. In Area 2B, for example, the best catch rates may be achieved by targeting smaller fish, while in Alaska it pays to target larger fish.

If the decline in selectivity in Area 2B were the result of the commercial fishery targeting areas where the fish are smaller, one would expect to see the decline in data from survey releases (which are done over the whole area) but not in data from releases at spot fishing locations (most of which are customary commercial fishing locations). But the mark-recapture data show the opposite pattern if anything (Figs. 2c and 2d), so ground selection does not appear to be the explanation.

All of the estimates reported here refer to a period (1960-1990) when halibut growth rates were higher and there were substantial numbers of large fish in the catches. Nowadays there are few fish larger than 120 cm in the commercial catches, so commercial selectivity may well have changed. At time of writing (Sept. 2004) that does not appear to have happened. While not numerous, recoveries to date of PIT tags released in 2003 in Areas 2B, 2C, and 3A show selectivity patterns consistent with the estimates reported here.

References

Clark, W. G. 1993. Estimation of halibut growth and selectivity parameters. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 1992:95-112.

Clark, W. G., and S. R. Hare. 2004. Assessment of the Pacific halibut stock at the end of 2003. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2003:171-200.

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

Kaimmer, S. M. 2000. Pacific halibut tag release programs and tag release and recovery data, 1925 through 1998. Int. Pac. Halibut Comm. Tech. Rep. 41.

Myers, R. A., and J. M. Hoenig. 1997. Direct estimates of gear selectivity from multiple tagging experiments. Can. J. Fish. Aquat. Sci. 54:1-9.

Myhre, R. J. 1969. Gear selection and Pacific halibut. Int. Pac. Hal. Comm. Report 51.

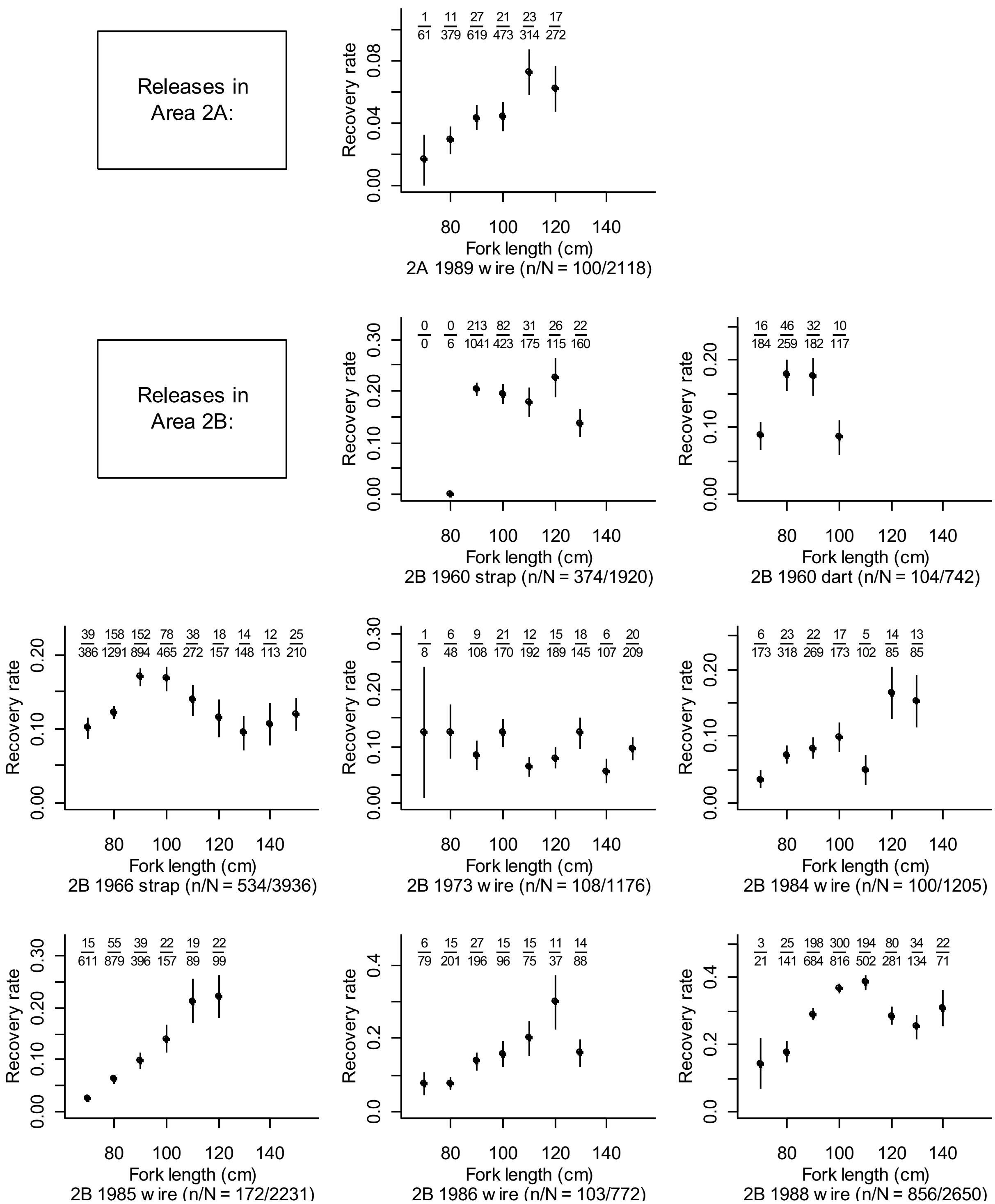


Figure 1. Raw recovery rate during the first two years after release as a function of approximate mean length during those two years, by midpoint of 10 cm length interval. All experiments that yielded at least 100 recoveries are shown. Data are pooled at the upper end so as to have at least 10 recoveries in the last length group plotted. The vertical bars are ± 1 standard deviation. The little numbers at the top are actual releases and recoveries.

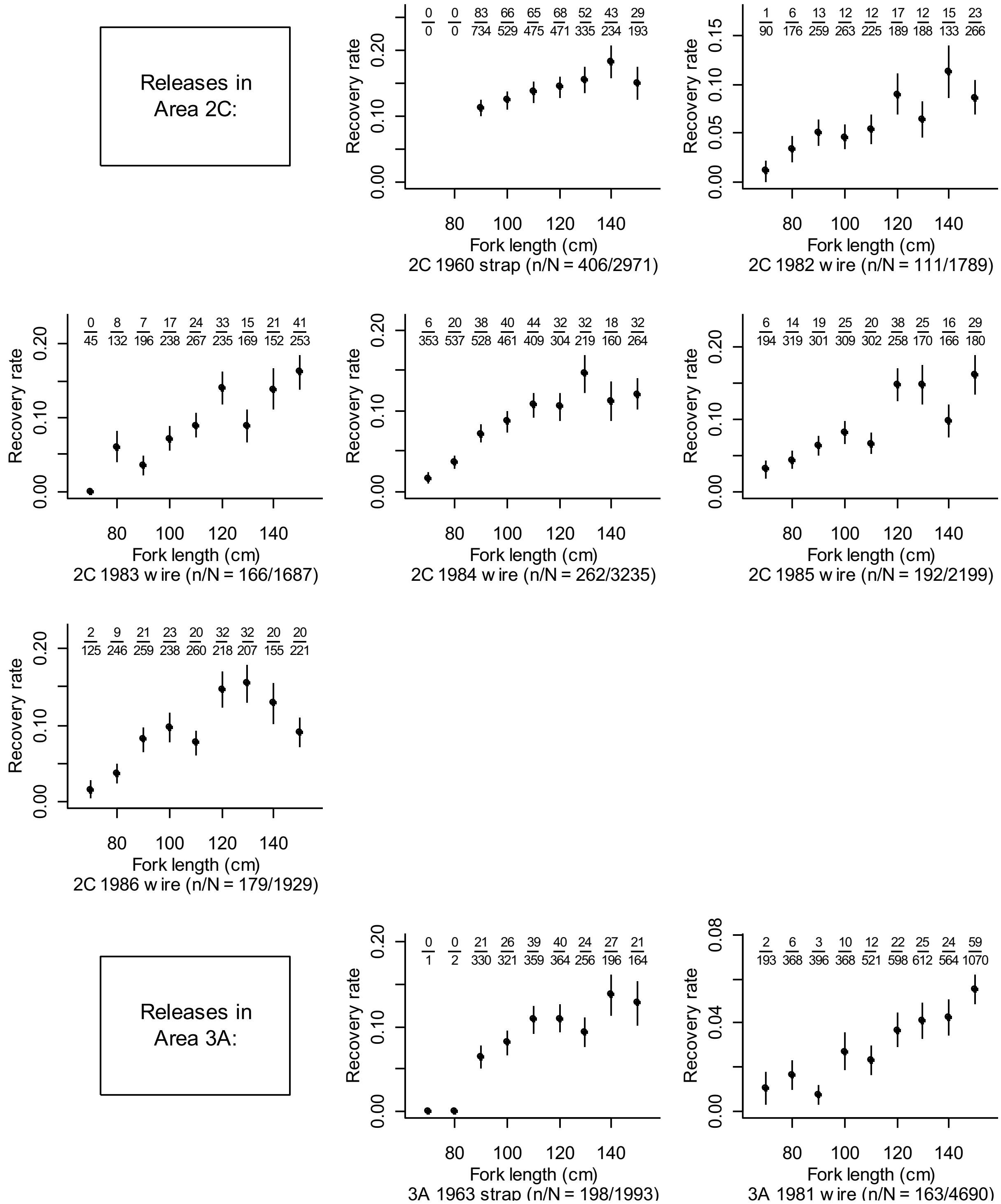


Figure 1 (continued).

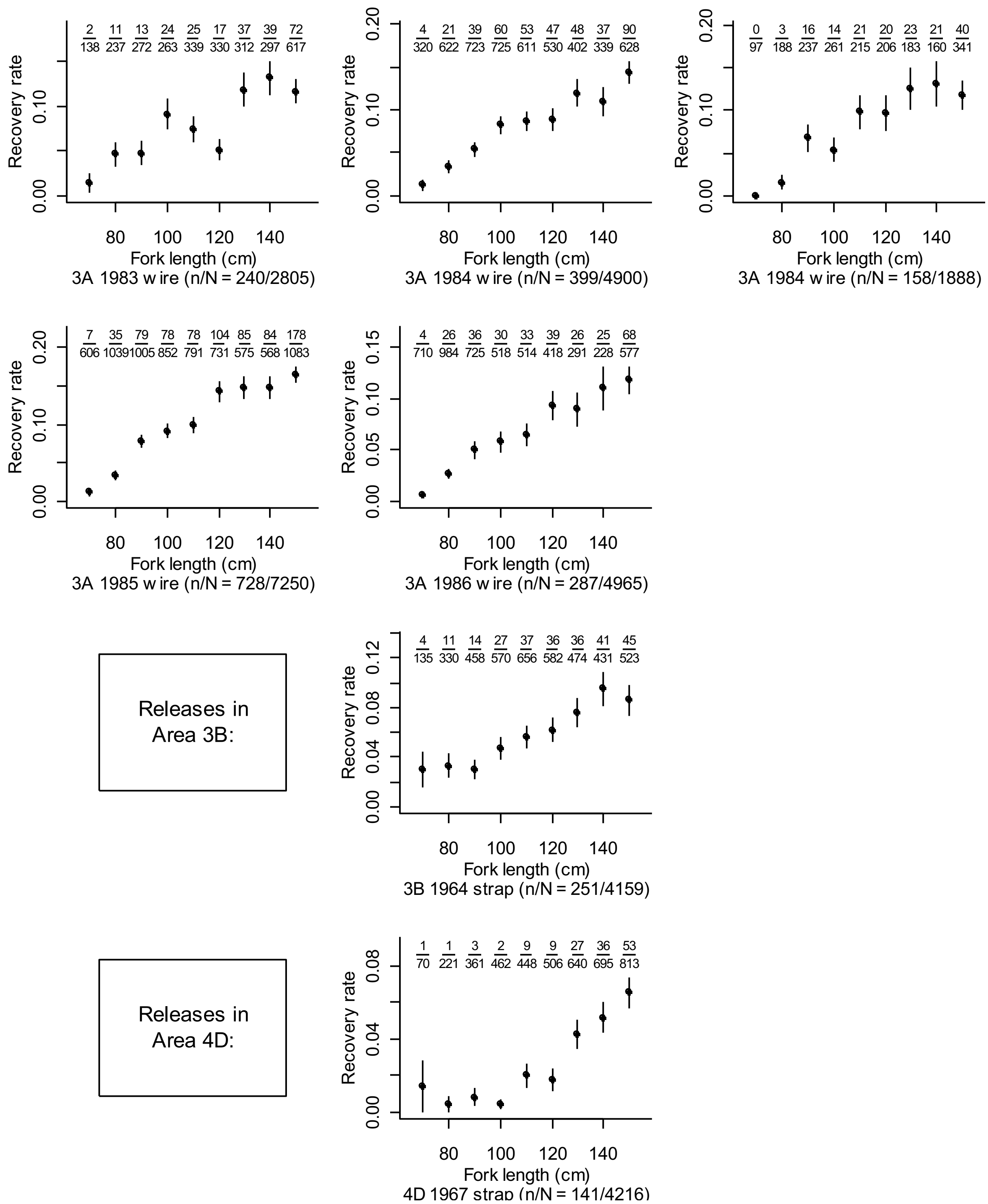


Figure 1 (concluded).

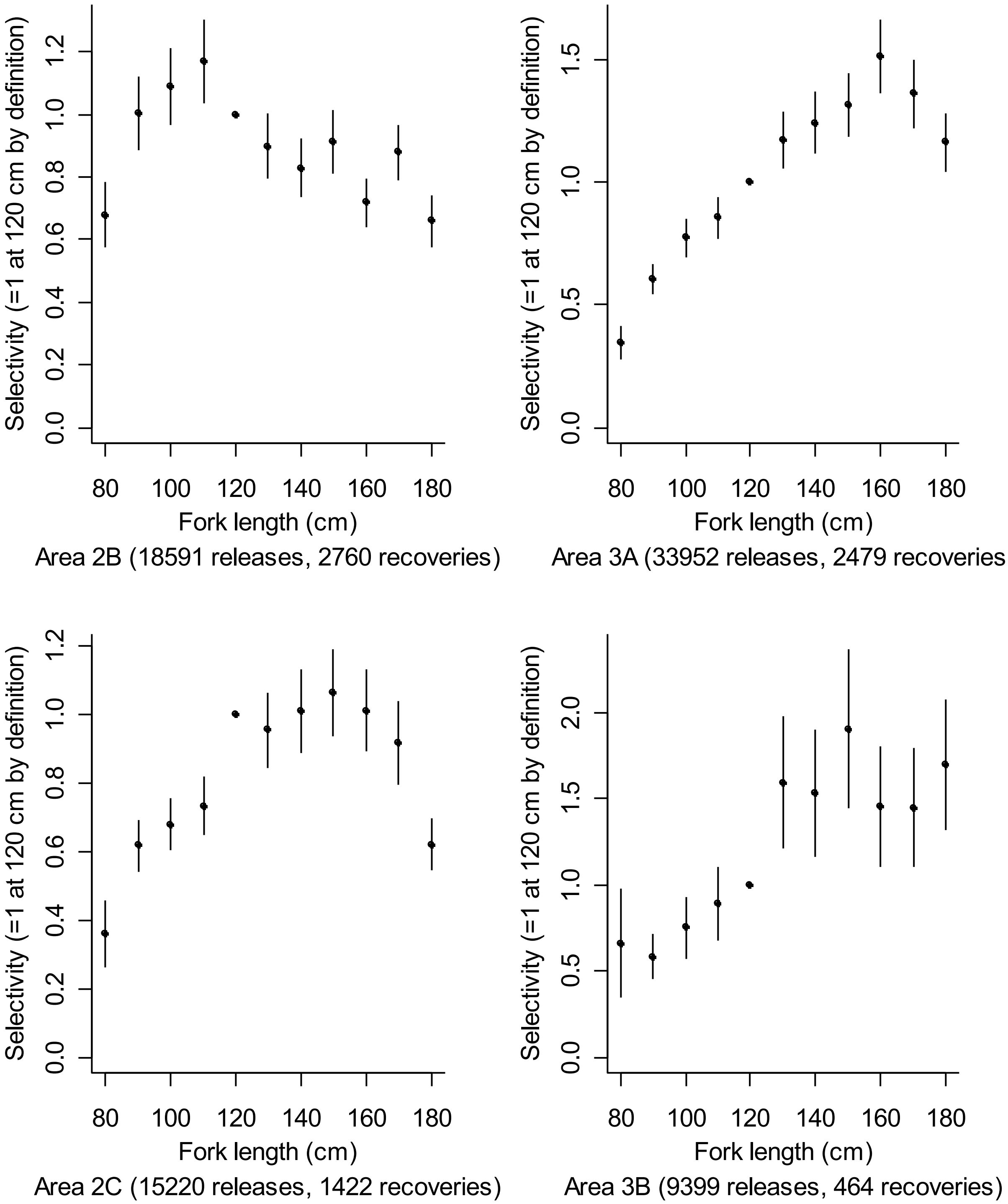


Figure 2a. Estimates of length-specific commercial selectivity (±1 standard deviation) based on all releases 1960-1990, by regulatory area. The scale was set by defining selectivity to be one at 120 cm, so that value has no standard deviation, and other values can and do exceed 1.0.

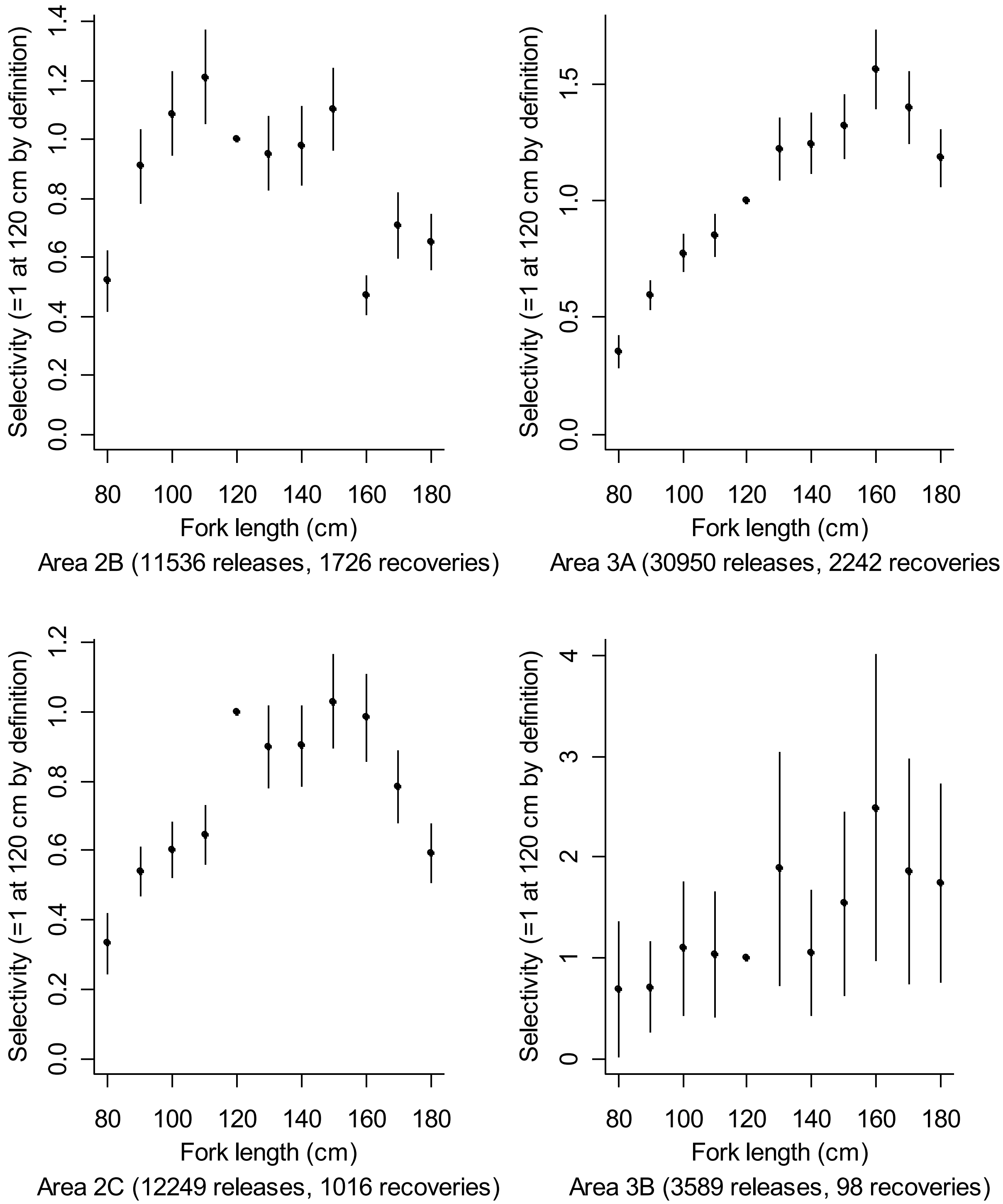


Figure 2b. Estimates of length-specific commercial selectivity (±1 standard deviation) based on releases in 1973-1990 only, by regulatory area. The scale was set by defining selectivity to be one at 120 cm, so that value has no standard deviation, and other values can and do exceed 1.0.

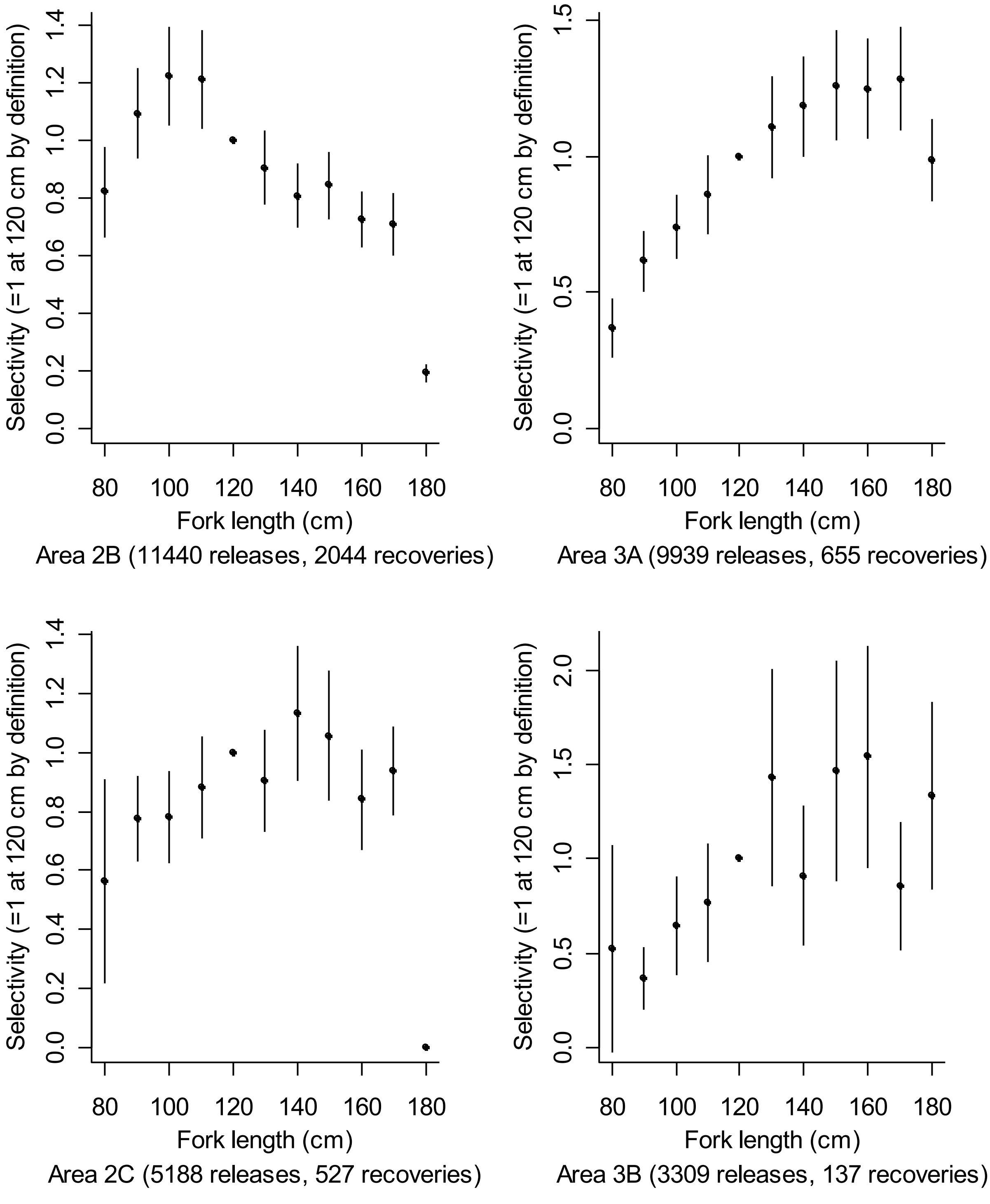


Figure 2c. Estimates of length-specific commercial selectivity (±1 standard deviation) based on releases in 1960-1990 at spot fishing locations only, by regulatory area. The scale was set by defining selectivity to be one at 120 cm, so that value has no standard deviation, and other values can and do exceed 1.0.

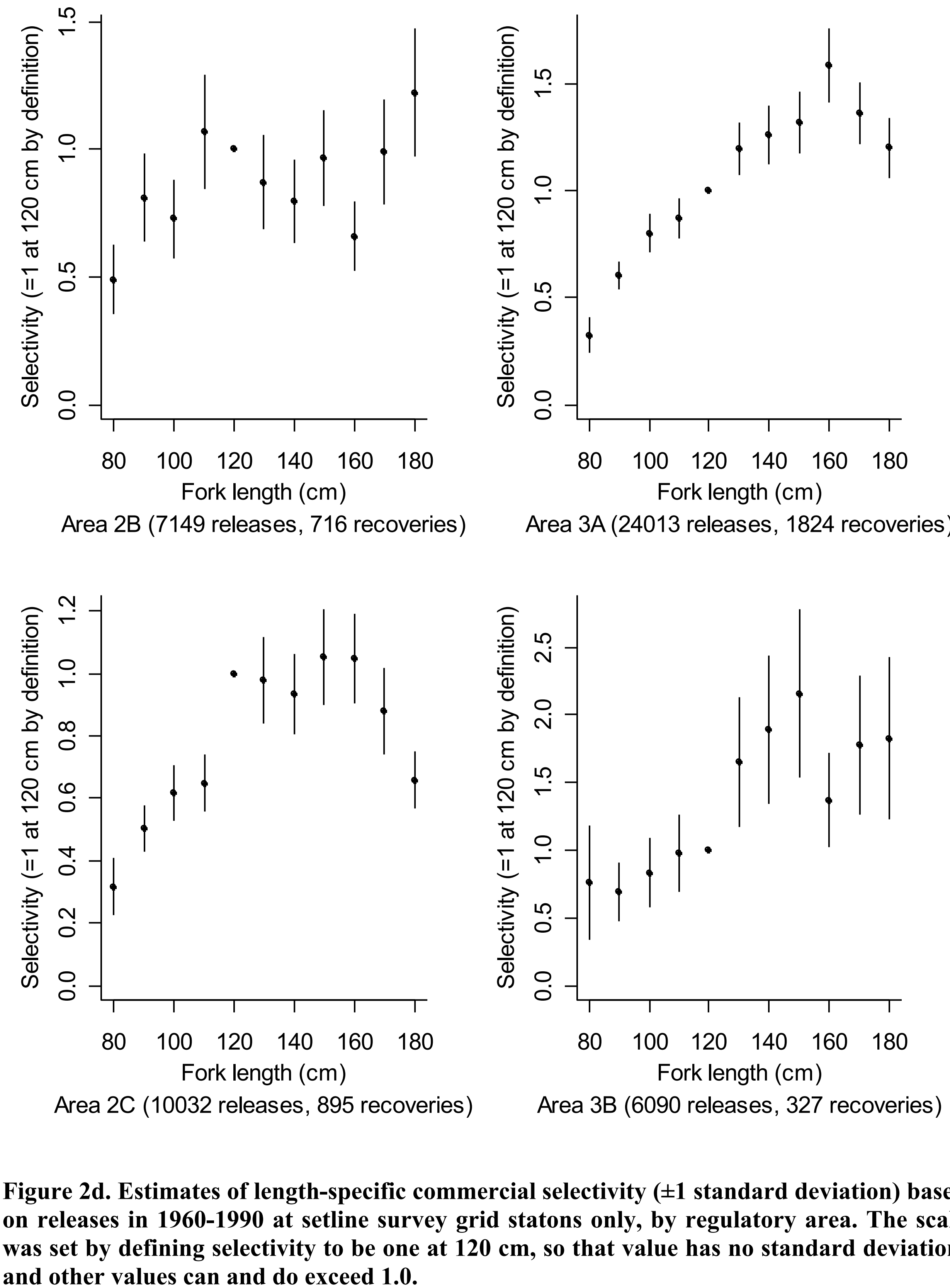


Figure 2d. Estimates of length-specific commercial selectivity (±1 standard deviation) based on releases in 1960-1990 at setline survey grid statons only, by regulatory area. The scale was set by defining selectivity to be one at 120 cm, so that value has no standard deviation, and other values can and do exceed 1.0.

Assessment of the Pacific halibut stock at the end of 2004

William G. Clark and Steven R. Hare

Abstract

This year's assessment uses the same methods as last year's to estimate exploitable biomass. Estimated coastwide exploitable biomass is 395 million pounds compared with 431 million last year, largely due to downward revisions of last year's estimates rather than a real decline in the stocks. A constant harvest rate of 22.5% rather than last year's 25% is used to calculate total CEY in Areas 2 and 3A. In Area 3B the target harvest rate is 20% this year rather than 25% last year. In Area 4 the target harvest rate remains at 20%. Fishery CEY totals 72.17 million pounds.

Introduction

Each year the IPHC staff assesses the abundance and potential yield of Pacific halibut using all available data from the commercial fishery and scientific surveys (Appendix A). Exploitable biomass in each of IPHC regulatory areas 2B, 2C, 3A, 3B, 4A, and 4B is estimated by fitting a detailed population model to the data from that area, going back to 1974 in the eastern areas and to 1996 in Areas 3B and 4. Exploitable biomass in Areas 2A and 4CDE is estimated by applying a survey-based estimate of relative abundance to the analytical estimate of biomass in the adjoining area (2B for 2A, 4A for 4CDE).

A biological target level for total removals is calculated by applying a fixed harvest rate to the estimate of exploitable biomass. This target level is called the "constant exploitation yield" or CEY for that area in the coming year. The corresponding target level for catches in directed fisheries subject to allocation is called the fishery CEY. It comprises the commercial setline catch in all areas plus the sport catch in Areas 2A and 2B. It is calculated by subtracting from the total CEY an estimate of all unallocated removals—bycatch of legal-sized fish, wastage of legal-sized fish in the halibut fishery, fish taken for personal use, and sport catch except in Areas 2A and 2B.

Staff recommendations for catch limits in each area are based on the estimates of fishery CEY but may be higher or lower depending on a number of statistical, biological, and policy considerations. Similarly, the Commission's final quota decisions are based on the staff's recommendations but may be higher or lower.

Evolution of assessment methods through 2003

From 1982 through 1994, the halibut stock assessment relied on CAGEAN, a simple age-structured model fitted to commercial catch-at-age and catch-per-effort data. The constant age-specific commercial selectivities used in the model were fundamental model parameters, estimated directly.

Beginning in the late 1980s, halibut growth rates in Alaska declined dramatically. As a result, age-specific selectivity decreased. CAGEAN did not allow for that, and by the mid-1990s was seriously underestimating abundance. In effect, it interpreted lower catches as an indication of lower abundance, whereas the real cause was lower selectivity. Incoming year classes were initially estimated to be small, but in subsequent years' assessments those estimates would increase when unexpectedly large numbers of fish from those year classes appeared in the catches. The year-to-year changes in the stock trajectory shown by the assessment therefore developed a strong retrospective pattern. Each year's fit showed a steep decline toward the end, but each year the whole trajectory shifted upward.

The staff sought to remedy that problem by making selectivity a function of length in a successor model developed in 1995. It accounted not only for the age structure of the population, but also for the size distribution of each age group and the variations in growth schedule that had been observed. The fundamental selectivity parameters in this model were the two parameters of a function (the left limb of a normal density) by which the selectivity of an individual fish was determined from its length. The age-specific selectivity of an entire age group was calculated by integrating length-specific selectivity over the estimated length distribution of the age group, and that age-specific selectivity was used to calculate predicted catches. The new model was fitted to both commercial data and IPHC setline survey data, with separate length-specific selectivity functions. Commercial catchability and selectivity were allowed to drift slowly over time, while survey catchability and selectivity were held constant (Sullivan et al. 1999).

When this model was fitted to data from Area 2B and Area 3A, quite different length-specific selectivities were estimated, which suggested that fishery selectivity was not wholly determined by the properties of the gear and the size of the fish but also depended on fish behavior (e.g., migration). These behavioral elements are likely to be more related to age than size. The age of sexual maturity, for example, remained virtually the same in Alaska despite the tremendous decrease in growth, so the size at maturity is now much smaller than it was. While size must affect selectivity, it was thought that age was also influential.

To allow for that, the model was fitted in two ways. The original form was called the "length-specific" fit, because a single set of estimates of the two parameters of the length-based survey selectivity function was used in all years. In a second form, called the "age-specific" fit, the parameters were allowed to drift over time (like the commercial selectivity parameters), but they were required (by a heavy penalty) to vary in such a way that the integrated age-specific selectivities calculated in each year remained constant over time.

The usual diagnostics gave little reason to prefer one fit over the other. Goodness of fit was similar: good for both in 2B, not so good for either in 3A. The retrospective behavior of both fits was dramatically better than that of CAGEAN and quite satisfactory in all cases, although the length-specific fit was more consistent from year to year in 3A and the age-specific fit was more consistent in 2B (Clark and Parma 1999). The two fits produced very similar estimates of abundance in Areas 2B and 2C, but in 3A the length-specific estimates were substantially higher, so out of caution the staff catch limit recommendations were based on the age-specific fit through 1999.

The assessment model was simplified and recoded as a purely age-structured model in 2000 to eliminate some problems associated with the modeling of growth and the distribution of length at age (Clark and Hare 2001). It retained the option of modeling survey selectivity as a function of mean length at age (observed not predicted), but the production fits continued to be based on constant age-specific survey selectivity, estimated directly as a vector of age-specific values rather than as a parametric function of age.

The fit of this model to Area 3A data in 2002 showed a dramatic retrospective pattern, similar to the pattern of successive CAGEAN fits in the mid-1990s. Treating setline survey selectivity as length-specific rather than age-specific largely eliminated the pattern. Accumulated data showing very similar trends in catch at length in IHPC setline surveys and NMFS trawl surveys provided further evidence that setline selectivity is, after all, determined mainly by size rather than by age (Clark and Hare 2002).

Another anomaly of the 3A model fit in 2002 was the unexpectedly large number of old fish (age 20+) in the last few years' catches. This was found to be the result of an increase in the proportion of otoliths read by the break-and-burn rather than surface method. Surface readings tend to understate the age of older fish, and IPHC age readers had been gradually doing more and more break-and-burn readings as the number of older fish in the catches increased. The poor model fit at these ages indicated a need to deal explicitly with the bias and variance of both kinds of age readings.

An entirely new model was written for the 2003 assessment (Clark and Hare 2004). Both commercial and survey selectivity were parameterized as piecewise linear functions of mean length at age in survey catches, and were required to reach an asymptote of one at or before a length of 130 cm. Because females are larger than males, all of the population accounting and predictions were done separately for each sex. (The age/sex/size composition of the commercial landings was estimated external to the assessment for this purpose.) The observed age compositions (surface or break-and-burn) were predicted by applying estimated misclassification matrices to the age distributions. Even in its most parsimonious form—with just one survey and one commercial selectivity schedule for both sexes in all years—this model achieved very good fits to the sex-specific observations and good retrospective performance. It also produced somewhat higher estimates of average recruitment and recruitment variability. With this simple model it was feasible do standalone analytical assessments of abundance in Areas 3B, 4A, and 4B for the first time, using data from 1996-2003.

Features of the 2004 assessment

Only two minor changes were made for the 2004 assessment, and neither had a significant effect on the estimates of abundance. First, both the 2004 PIT tag recoveries (Clark and Chen 2005) and a reanalysis of earlier wire tag data (Clark 2005) indicated that commercial selectivity is not always asymptotic; it appeared to be more dome-shaped in Area 2B and more ramp-shaped in Area 3A. Fitting the assessment model with free-form selectivity schedules showed much the same thing for commercial selectivity (Fig. 1), namely an assortment of shapes beyond 120 cm. Nevertheless a schedule that reaches an asymptote of one at 120 cm is a good approximation to and compromise among the free estimates, and using an asymptotic commercial schedule is desirable for computing exploitable biomass and reporting harvest rates, so that it what was used in the assessment. All of the freely estimated survey selectivities either level out or increase after 120 cm. Freely estimated survey selectivities present no practical difficulties, so they were estimated that way in the assessment, and most of the estimates were ramp-shaped.

The second minor change was allowing sex-specific values for survey and commercial catchabilities. This was done only in the Area 3A assessment, where the standard model fitted the age composition of male catches well but the numbers of males in the catches were generally in excess of the predictions. In Area 3A males were estimated to be about twice as catchable as females of the same size; in other areas there was no difference. Even with higher catchability, males in Area 3A were estimated have quite low fishing mortality rates because they are so small.

Analytical estimates of abundance and CEY

Like last year's, this year's model fits are generally good (Fig. 2) and recent retrospective performance is satisfactory. Changes in stock biomass from the beginning of 2004 to the beginning of 2005 as estimated within this year's assessment are all 5% or less except in Area 4B, where there was an estimated 20% decrease. Some of the estimates of stock biomass have changed much more than 5% from last year's assessment because the addition of the 2004 data to this year's model fit has revised last year's estimate of biomass at the beginning of 2004, in most cases downward.

	2004 biomass 2003 assessment	**2004 biomass 2004 assessment**	**2005 biomass 2004 assessment**
Area 2A	8.5	7.9	7.0
Area 2B	65	61	58
Area 2C	80	65	66
Area 3A	146	154	146
Area 3B	65	54	56
Area 4A	21	20	20
Area 4B	15	12	10
Area 4CDE	30	28	32
Total	431	402	395

It is these downward revisions of last year's estimates that mainly account for the reduction of estimated coastwide exploitable biomass from 431 million pounds to 395. Female spawning biomass remains far above the minimum that occurred in the mid 1970s.

Table 1 shows estimates of exploitable biomass, total CEY, and fishery CEY. Exploitable biomass in Alaska is calculated with a fixed set of length-specific commercial selectivities that increase linearly from zero at 80 cm to one at 120 cm. In Area 2B the locally estimated selectivities are used because they are substantially higher than the values estimated for the Alaska areas.

Exploitable biomass in Area 2A is calculated as a proportion of the Area 2B analytical estimate. The proportion used is the ratio of survey CPUE's (three-year running mean) weighted by bottom areas:

$$\text{proportion} = \frac{(\text{2A CPUE}) \times (\text{2A bottom area})}{(\text{2B CPUE}) \times (\text{2B bottom area})}$$

The idea here is that survey CPUE is an index of density and multiplying it by the total bottom area gives an index of total biomass. The value of the scaling proportion this year is 12%, down from 13% last year as a result of updating the CPUE values. In the same way, exploitable biomass in Area 4CDE is calculated as 160% of the Area 4A biomass (up from 142% last year).

Total CEY is calculated by applying a harvest rate of 22.5% in Areas 2A, 2B, 2C, and 3A, and 20% in Areas 3B and 4 (Hare and Clark 2005). Last year the target harvest rate for Areas 2 and 3 was 25% pending a reanalysis of harvest policy using the new estimates of length-specific commercial selectivity.

Reliability of model fits to short data series

In Areas 2B, 2C, and 3A the model is fitted to 31 years of data (1974-2004), but in Areas 3B, 4A, and 4B to only 9 years (1996-2004). The performance of fits to short data series can be examined by comparing fits to the full series and fits to shorter subsets in areas with long data series.

Figure 3a shows the Area 3A fit to the first 9 years of data (1974-1982), and on the same graph the estimates from this year's fit to the full series. The 1982 fit agrees quite well with the 2004 fit in respect of selectivities, numbers at age in 1974, and fishing mortality rates. Most of the 1982 estimates of recruitment are a bit high and as a result the estimate of exploitable biomass in 1983 is high by 15% or so, but as explained below that is well within the normal error range of fits based on many years of data.

Similarly Figure 3b compares this year's full Area 2C fit (1974-2004 data) with one that uses data from 1996-2004 only. Again the estimates from the short data series compare quite well with those from the full data set. In this case the exploitable biomass estimate from the short data series is about 10% lower than the full estimate.

Figure 3c shows the same comparison for Area 2B. In this case the 1996-2004 fit is pathological. Estimates of selectivities and numbers at age agree very well with the full assessment, but fishing mortality is greatly overestimated and as a result exploitable biomass is underestimated by more than 30%. This occurs because the survey CPUE series happens to begin with the very high 1996 and 1997 values. The later, lower survey catch rates suggest a substantial total mortality rate, which produces the high estimated fishing mortality rates and low abundance estimates. The commercial CPUE values show quite a different trend, so it is doubtful that this assessment would have been taken at face value even if no more data were available, but this example does show that an assessment based on a short data series can be strongly influenced by a few stray data points.

We do not believe that this year's 3B, 4A, and 4B assessments are suspect for this reason because in every case the survey and commercial CPUE series are very consistent and coherent in showing steady declines over the last 5-6 years. It is conceivable that these downward trends are an artifact of a widespread decline in setline catchability, and the low PIT tag recoveries certainly put the fishing mortality estimates in doubt, but the assessment data by themselves do not raise any suspicions about the fits or the estimates.

Variance estimates

Our estimates are maximum-likelihood estimates, and their variances can be estimated by any of a number of standard methods. In practice all of the methods produce very similar estimates, and the estimates are much too low when the model is misspecified (Punt and Butterworth 1993), which is almost always true of stock assessment models. The usual estimates of standard deviation for our model fits are less than 5%, but this year's estimates of abundance at the beginning of 2004 differ from last year's estimates by up to 20%. Changes of this size do not result from statistical variability but from trends in the stock and the fishery that are not reflected in the necessarily parsimonious parameterization of the model (Clark et al. 2004). These trends cause the model fits to make large excursions and abrupt corrections that appear as year-to-year changes much larger than what would be expected from sampling errors.

Figure 4 illustrates this characteristic with the retrospective behavior of the Area 3A exploitable biomass estimates. The estimate shoots way up in the mid-1980s, drops way down in the early 1990s, and since 1993 tracks reasonably well.

By now we know the actual abundance of the stock in the 1980s and early 1990s because all of the year classes then present have by now passed through, and their abundance in the 2004 model fit is therefore entirely determined by the catch at age. We can configure the 2004 assessment model to calculate the trends in catchability and selectivity that were occurring at that time and see why the current estimates make the observed excursions and corrections. Doing that shows that selectivity has changed little over the years, but that both commercial and especially survey catchability were increasing quite rapidly in the late 1970s and early 1980s. This was a period when people in the industry spoke of a sudden increase in abundance that could only be explained by fish coming out of a "black hole." In retrospect it is clear that catchability was increasing, perhaps because halibut took up a more demersal habit after the 1977 regime shift. The assessment model fits for that period allow for a slow drift in commercial catchability but no change in survey selectivity, so the increasing catch rates can only be fitted in successive assessments by increasing the abundance estimates. The estimated value of (constant) survey selectivity also increases in successive assessments, but even so the model fits cannot match the observed increase in survey catch rates (Fig. 3a).

In the latter 1980s and early 1990s commercial catchability declined owing to the nature of the derby fishery. There were no surveys between 1986 and 1993, so the assessments of that period rely on commercial CPUE as an index of abundance; hence the downward excursion of the biomass estimate in the early 1990s and the abrupt correction in 1993 when the next survey index of abundance became available. At its worst in 1991 the biomass estimate was low by about 50%, but the model fit to the data was good and the nominal variance was small (Fig. 5). In the 2004 assessment (Fig. 2c) the 1985-1995 period appears as a time when commercial catchability fell (before increasing again after the adoption of individual quotas in 1995) but there was little change in exploitable biomass despite the wide swings in the assessment estimate.

Because the actual abundance is effectively known for the early years of the data series (say eight or more years before the last data year), it is possible to calculate an empirical error variance from the observed deviations between the first assessment estimate for a given year and the converged value that becomes known eight or more years later. The 1982 assessment error, for example, can be calculated as the difference between the estimate made in that year's assessment and the historical value for 1982 computed in this year's assessment. Doing that for all possible years and two different models in Areas 2B, 2C, and 3A produces a generic value of the standard deviation of the biomass estimate of about 20%. Most (not all) of the really large deviations occur in the earlier years. More recently the assessment has been tracking somewhat better, so a 10-15% standard deviation seems more reasonable as a working value.

References

Clark, W.G. 2004a. A method of estimating the sex composition of commercial landings from setline survey data. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2003:111-162.

Clark, W.G. 2004b. Statistical distribution of IPHC age readings. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2003:99-110.

Clark, W.G. 2005. Estimates of length-specific commercial selectivity from historical marking experiments. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2004: this volume.

Clark, W.G., and Chen, D.G. 2005. Preliminary estimates based on 2004 PIT tag recoveries. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2004: this volume.

Clark, W.G., and Hare, S. R.. 2003. Assessment of the Pacific halibut stock at the end of 2002. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2002:95-120.

Clark, W.G., and Hare, S. R.. 2004. Assessment of the Pacific halibut stock at the end of 2003. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2003:171-200.

Clark, W. G., and Parma, A. M.. 1999. Assessment of the Pacific halibut stock in 1998. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 1998: 89-112.

Clark, W.G., Hare, S.R., and Leaman, B.M. 2004. Sources of uncertainty in annual CEY estimates. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2003:163-170.

Hare, S.R., and Clark, W.G. 2005. Analysis of the constant harvest rate policy for 2005. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2004: this volume.

Punt, A.E., and Butterworth, D.S. 1993. Variance estimates for fisheries assessment: their importance and how best to evaluate them. p. 145-162. *In* S.J. Smith, J.J. Hunt, and D. Rivard [ed.] Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120.

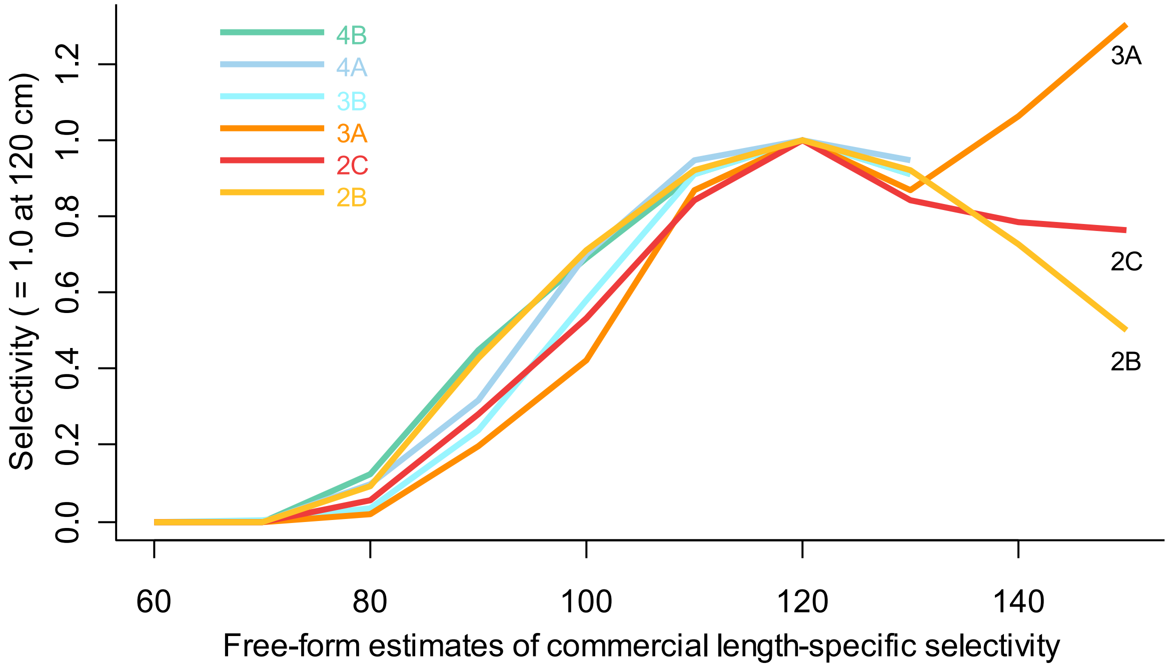
Sullivan, P. J., Parma, A. M., and Clark, W. G. 1999. The Pacific halibut stock assessment of 1997. Int. Pac. Halibut Comm. Sci. Rep. 79.

Table 1. Estimates of exploitable biomass and CEY.

	Area 2A	Area 2B	Area 2C	Area 3A	Area 3B	Area 4A	Area 4B	Area 4CDE	Total
2004 catch limit	1.48	13.80	10.50	25.06	15.60	3.47	2.81	3.79	76.51
2004 exploitable biomass (2003 assessment)	8.5[1]	66	80	146	65	21	15	30[2]	431.5
2005 exploitable biomass (2004 assessment)	7.0[1]	58	66	146	56	20	10	32[2]	395.0
Other removals									
Sport catch	0.51	1.37	2.31	4.74	0.01	0.02	0.00	0.00	8.96
Legal-sized bycatch	0.37	0.14	0.15	1.52	0.39	0.52	0.29	1.92	5.3
Personal use	0.02	0.30	0.63	0.28	0.03	0.02	0.00	0.08	1.36
Legal-sized wastage	0.00	0.02	0.03	0.07	0.03	0.02	0.00	0.01	0.18
Total	0.90	1.83	3.12	6.61	0.46	0.58	0.29	2.01	15.8
…excluding sport catch	0.39[3]	0.44[4]	---	---	---	---	---	---	
Total CEY at 20%	1.40	11.6	13.2	29.2	11.2	4.0	2.0	6.4	79
Fishery CEY at 20%	1.01	11.2	10.1	22.6	**10.7**	**3.4**	**1.7**	**4.4**	**65.11**
Total CEY at 22.5%	1.56	13.1	14.9	32.9	12.6	4.5	2.3	7.2	89.06
Fishery CEY at 22.5%	**1.17**[3]	**12.7**[3,4]	**11.8**	**26.3**	12.2	3.9	2.0	5.2	75.27

Notes:

1. Area 2A exploitable biomass estimated as 13% of Area 2B for 2004 and as 12% for 2005 due to changes in survey catch rates.
2. Area 4CDE exploitable biomass calculated as 142% of Area 4A biomass for 2004 and as 160% for 2005.
3. Fishery CEY includes sport catch in Areas 2A and 2B.
4. Combined sport and commercial CEY for Area 2B includes Area 2B sport catch landed in the U.S. (0.200 million lb) and legal sized wastage (0.02 million lb) to conform with the Canadian allocation program. If sport landings in the U.S. are left out of fishery CEY, they become a subtraction from total CEY and fishery CEY is reduced to 12.5 million lb.



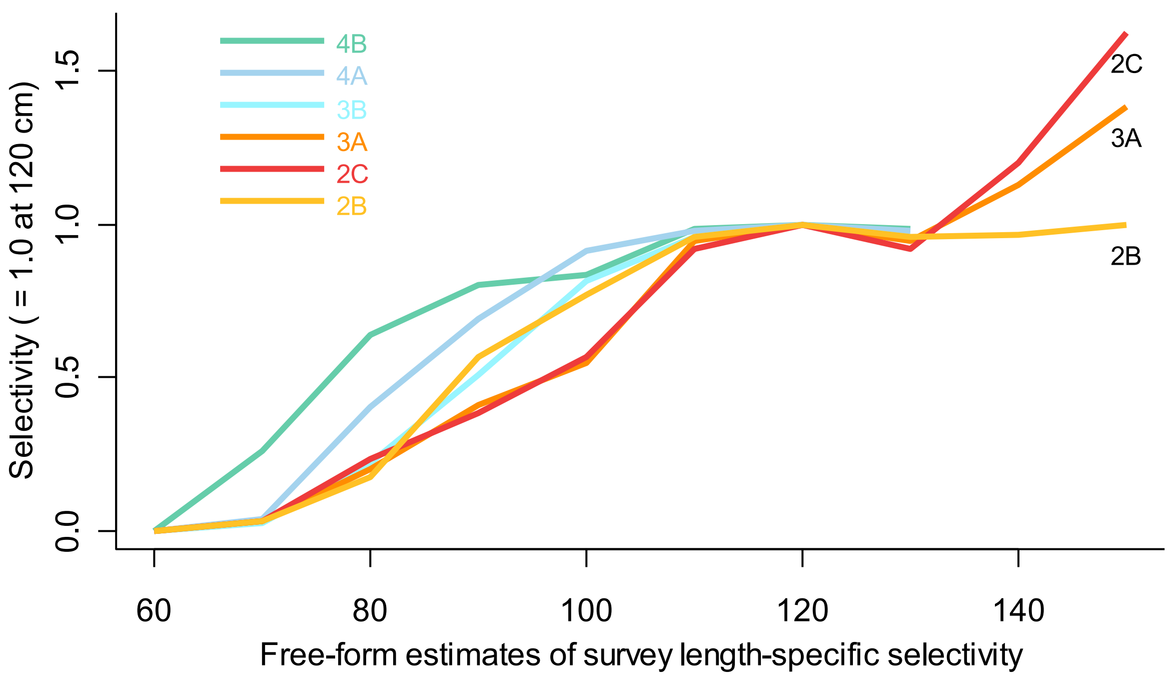


Figure 1. Free-form estimates of length-specific commercial (above) and survey (below) selectivity from assessment model fits. In the production fits commercial selectivity was asymptotic, but not survey selectivity.

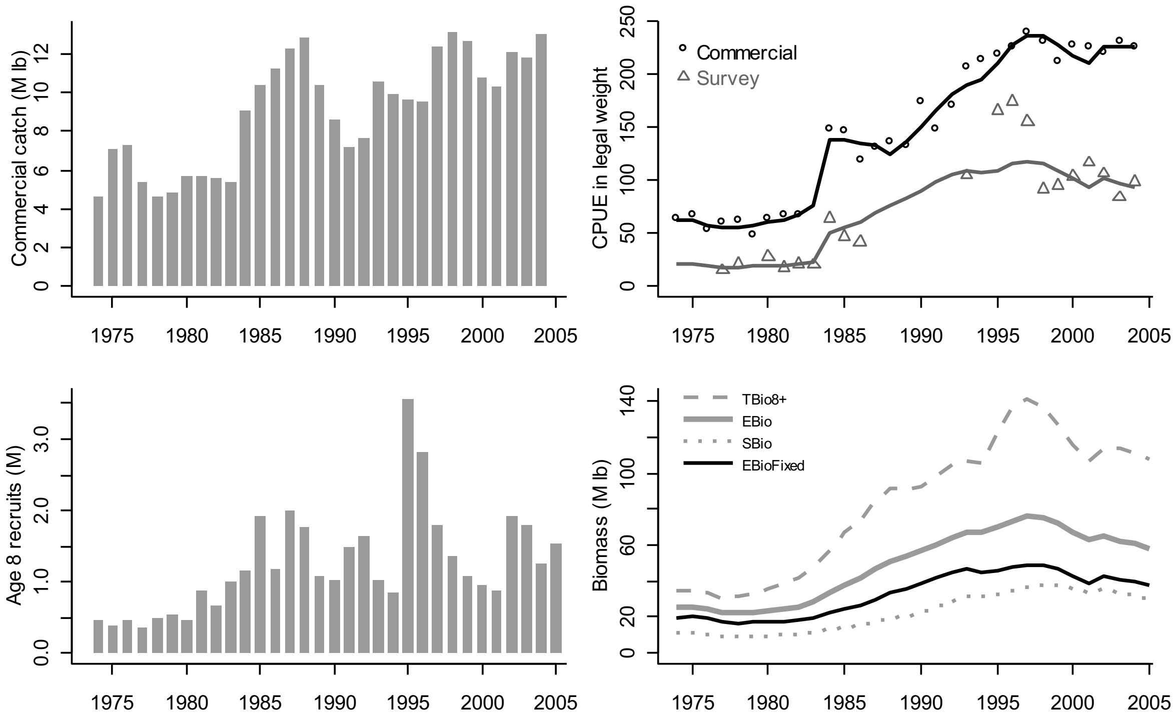


Figure 2a. Features of the 2004 assessment in Area 2B.

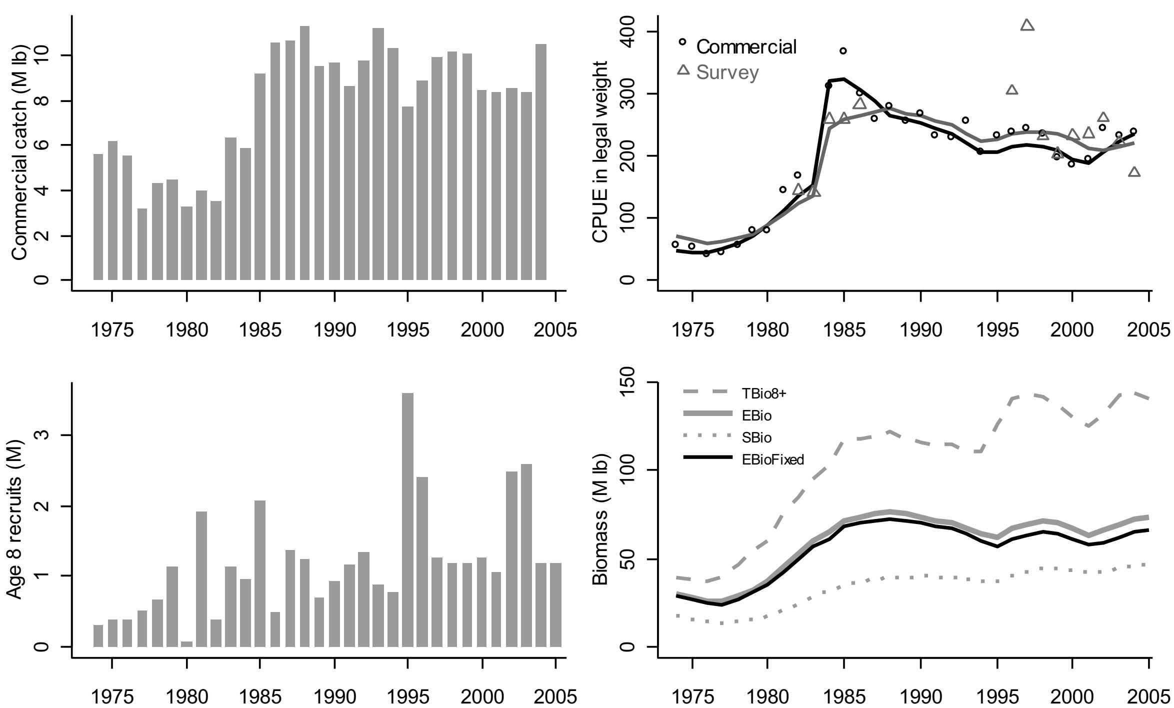


Figure 2b. Features of the 2004 assessment in Area 2C.

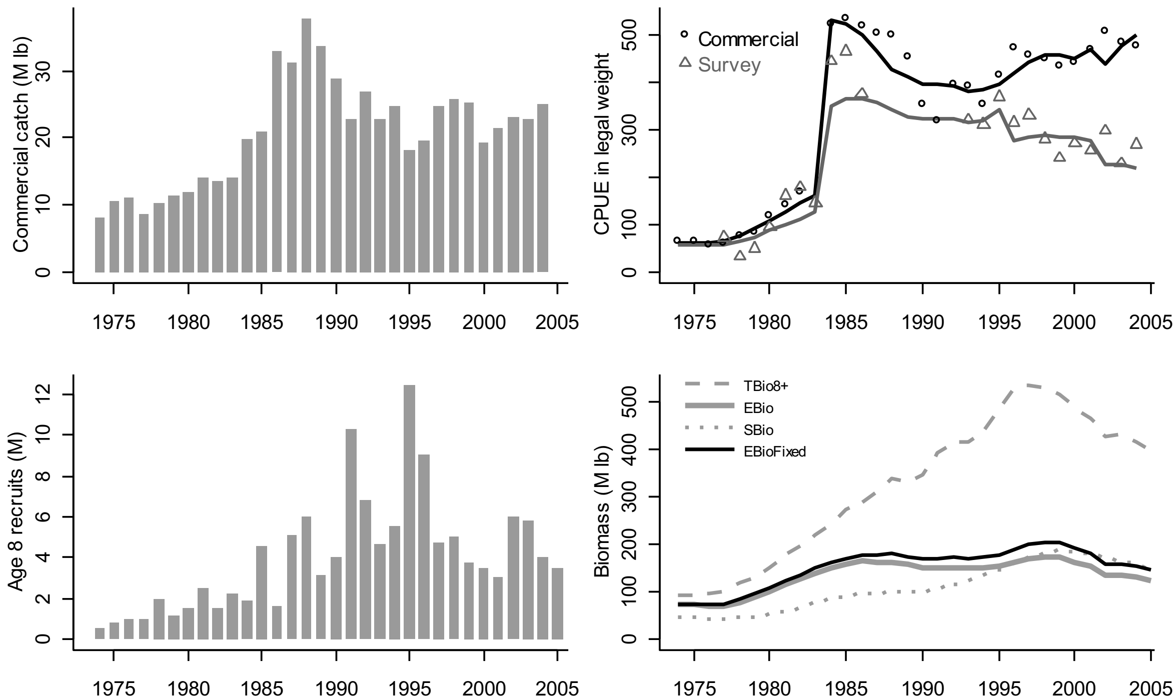


Figure 2c. Features of the 2004 assessment in Area 3A.

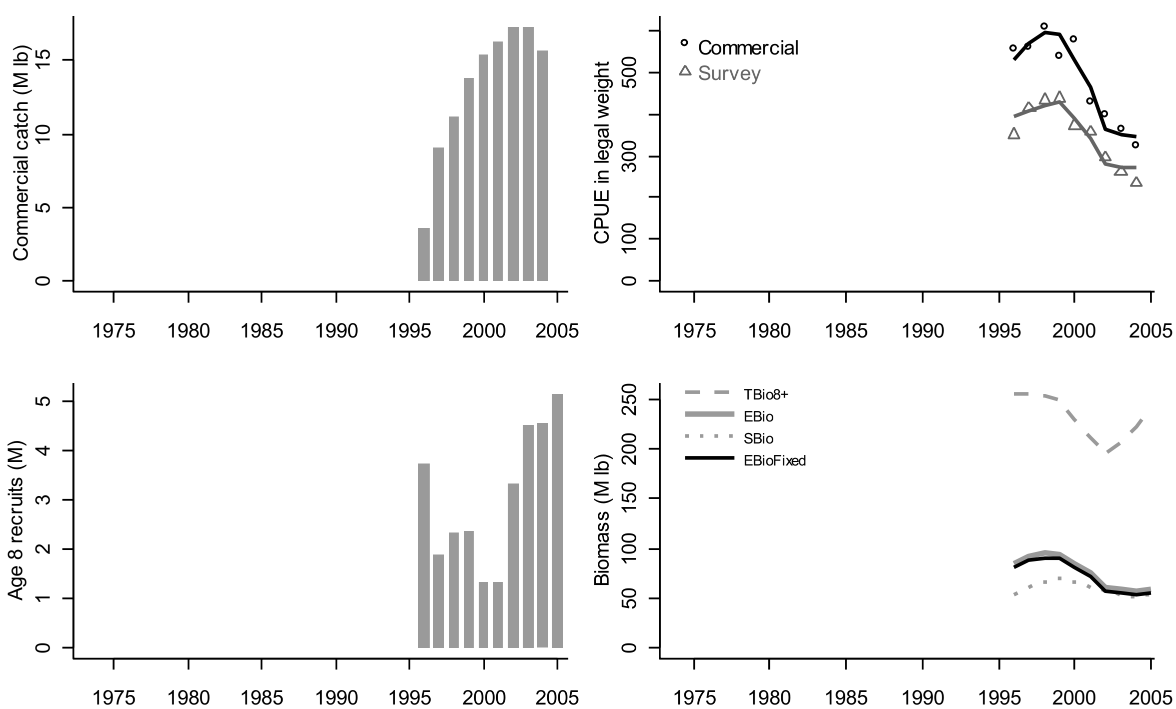


Figure 2d. Features of the 2004 assessment in Area 3B.

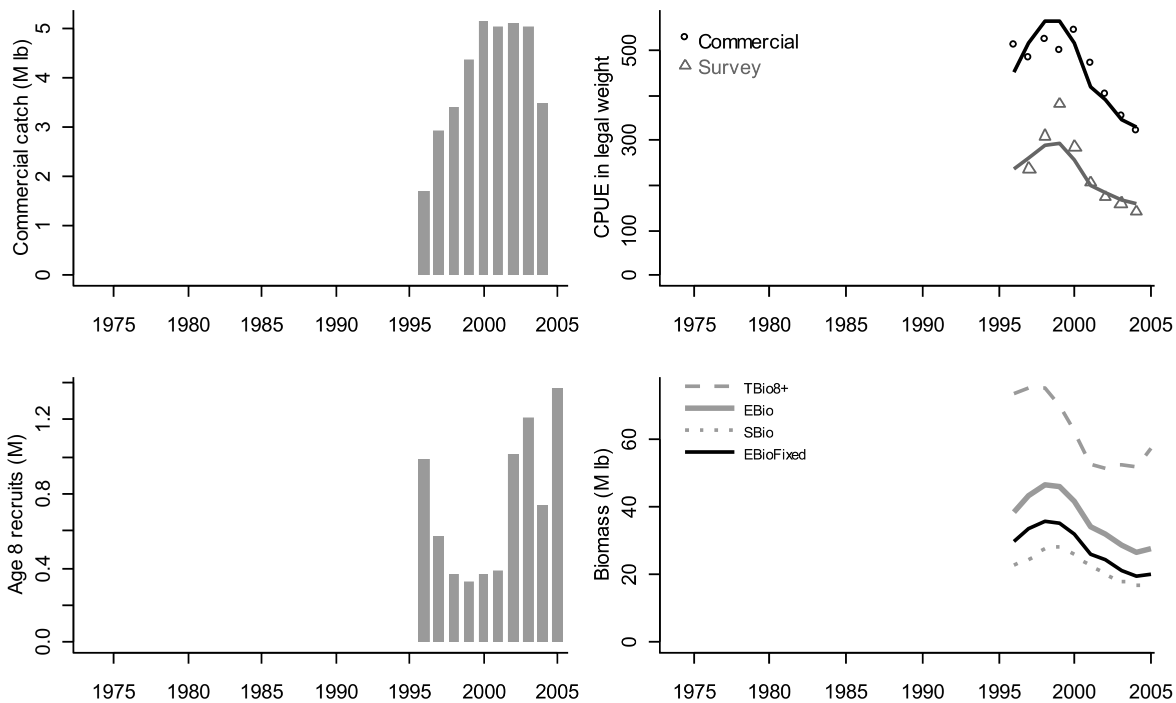


Figure 2e. Features of the 2004 assessment in Area 4A.

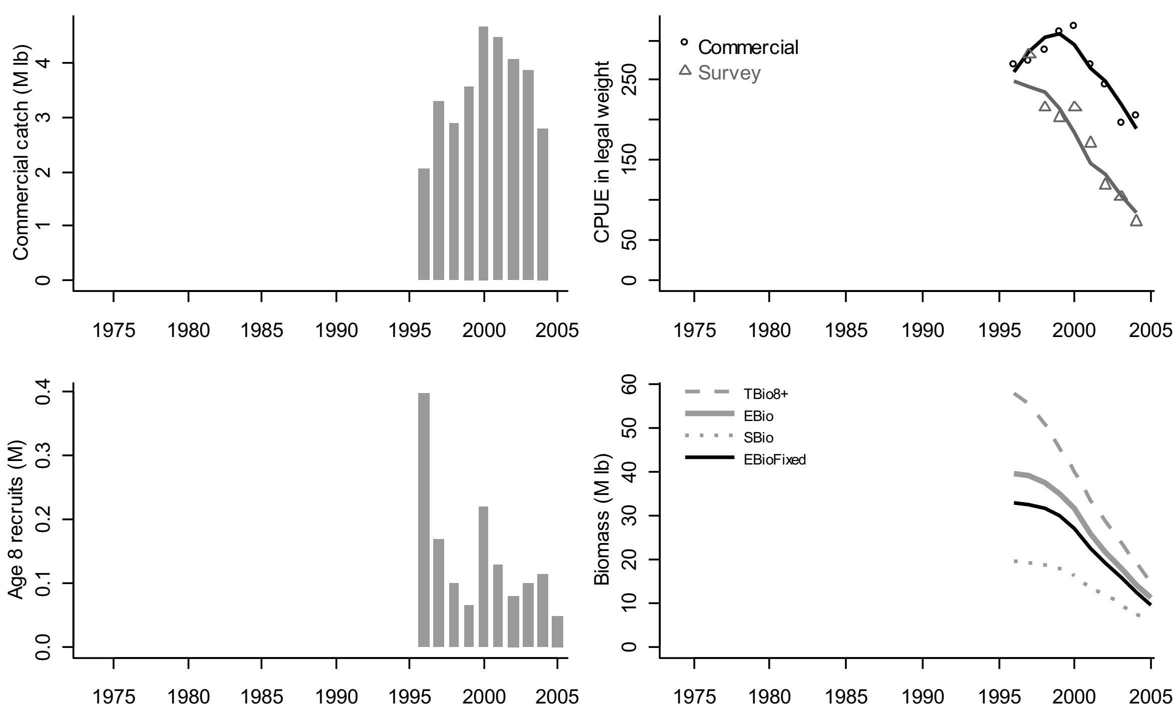


Figure 2f. Features of the 2004 assessment in Area 4B.

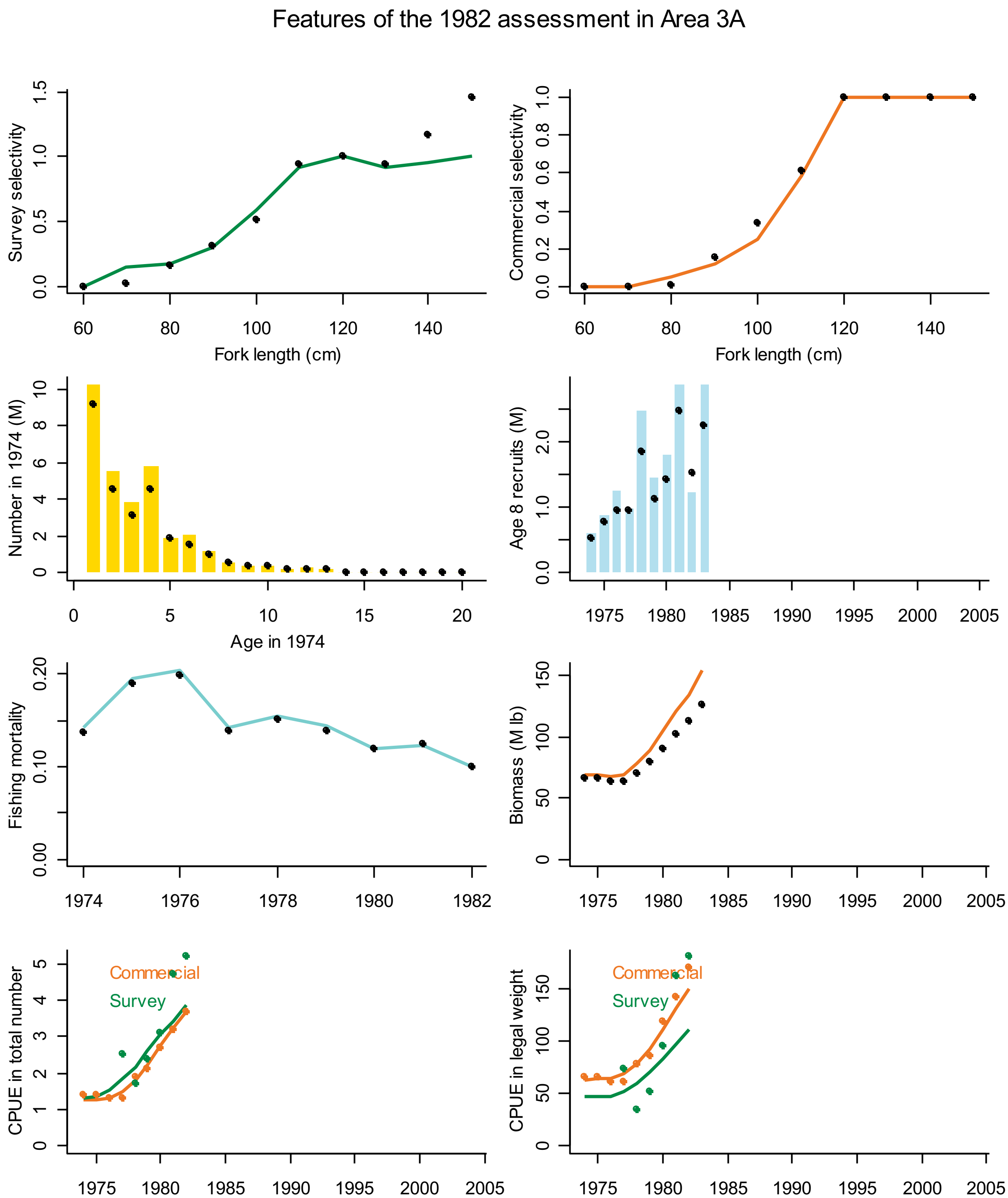


Figure 3a. Estimates of the stock in Area 3A in 1982 from a 1974-1982 fit (lines in upper six panels) and this year's 1974-2004 fit (points in upper six panels). In the bottom two panels the points are observed values and the lines are predictions from the 1974-1982 fit.

Features of the 2004 assessment in Area 2C

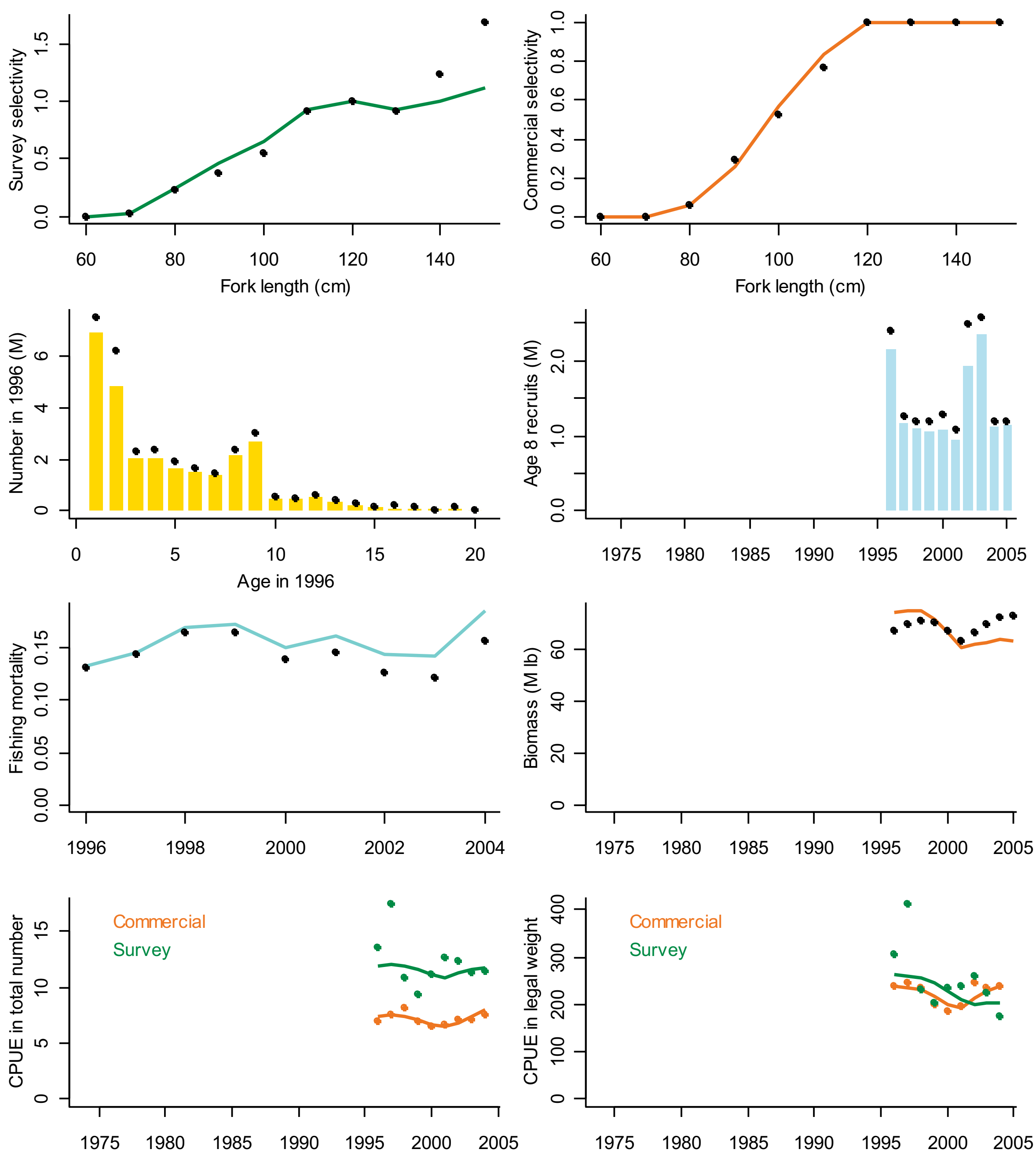


Figure 3b. Estimates of the stock in Area 2C in 2005 from a 1996-2004 fit (lines in upper six panels) and this year's 1974-2004 fit (points in upper six panels). In the bottom two panels the points are observed values and the lines are predictions from the 1996-2004 fit.

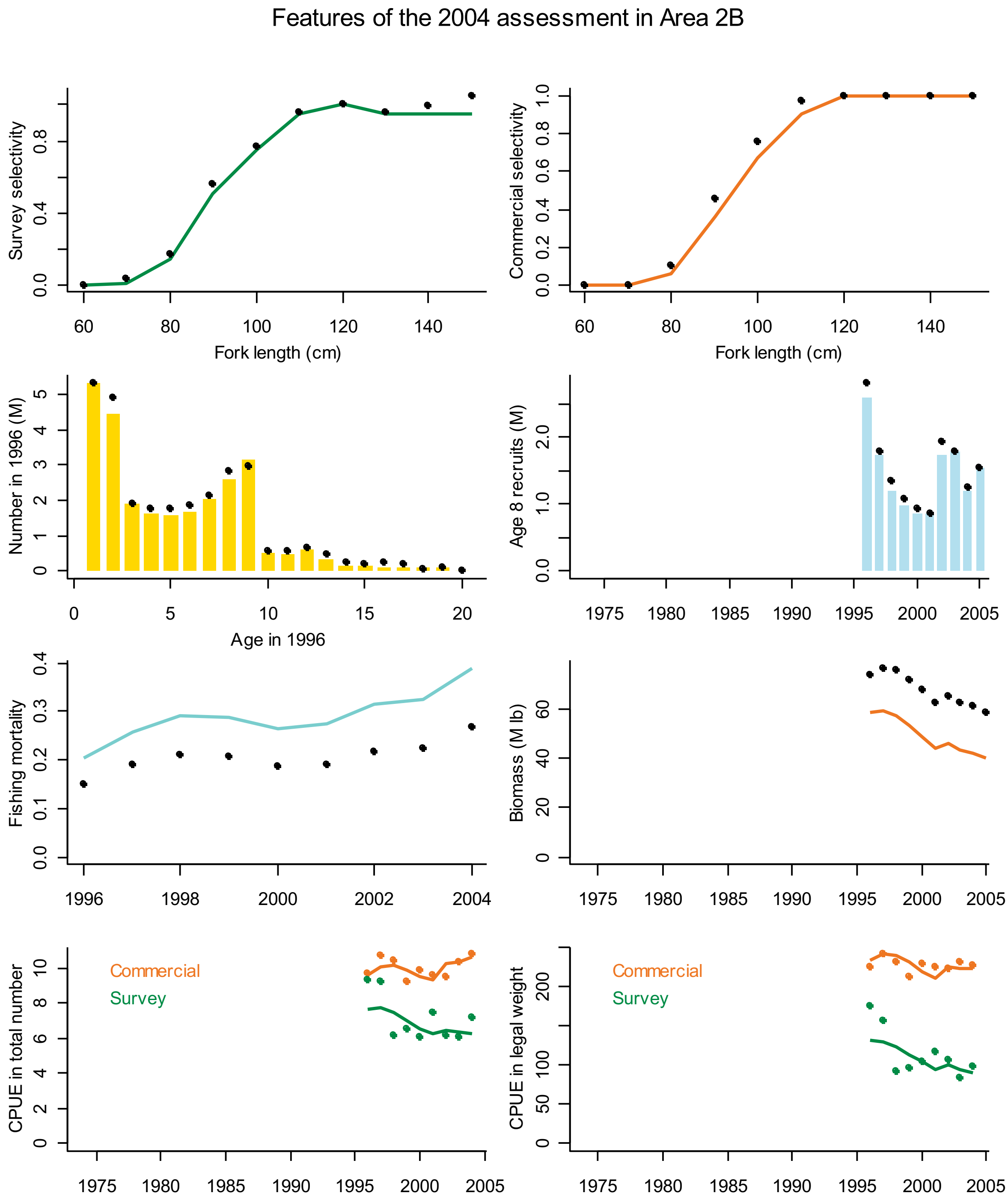


Figure 3c. Estimates of the stock in Area 2B in 2005 from a 1996-2004 fit (lines in upper six panels) and this year's 1974-2004 fit (points in upper six panels). In the bottom two panels the points are observed values and the lines are predictions from the 1996-2004 fit.

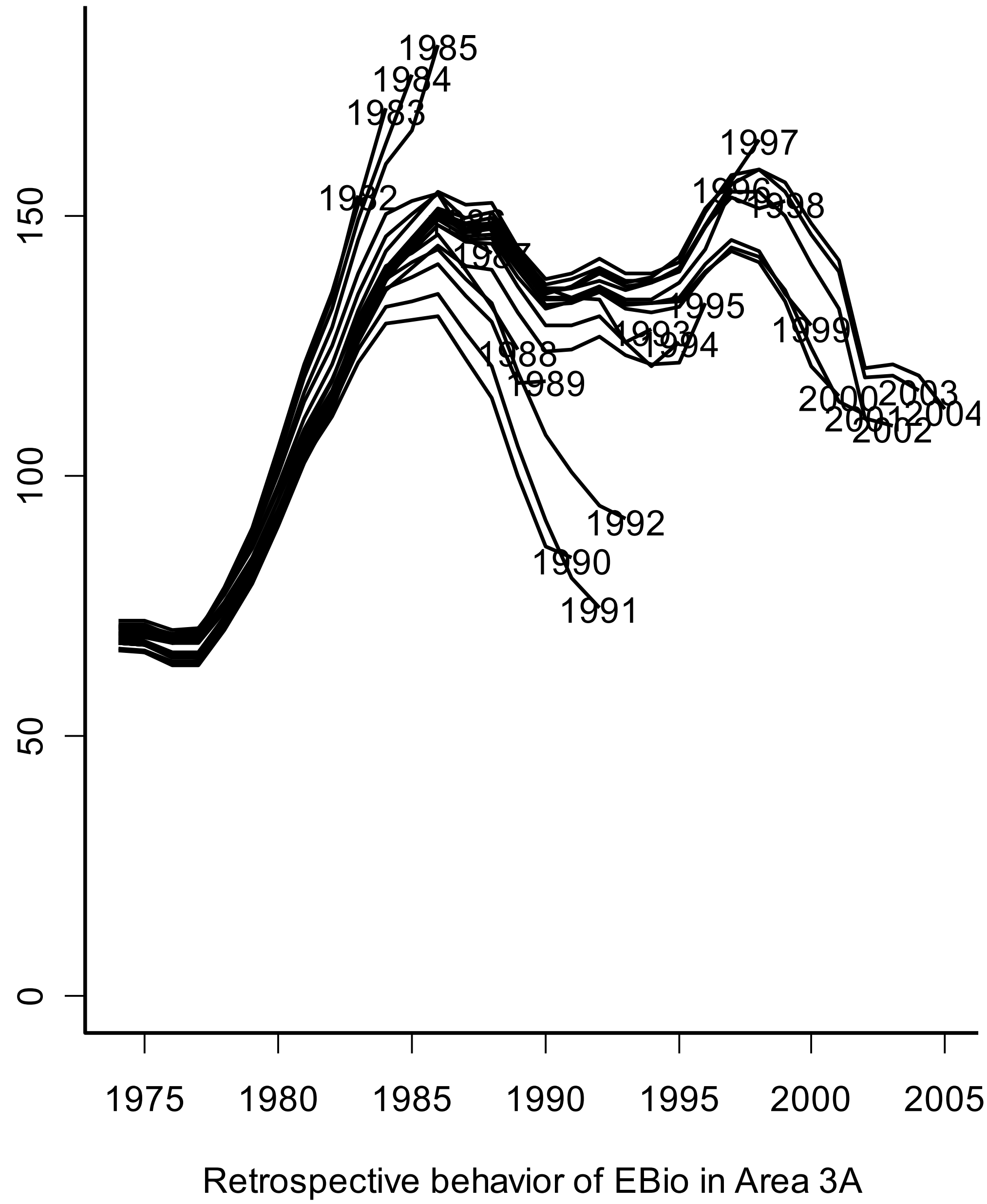


Figure 4. Retrospective behavior of estimates of exploitable biomass by the 2004 assessment model.

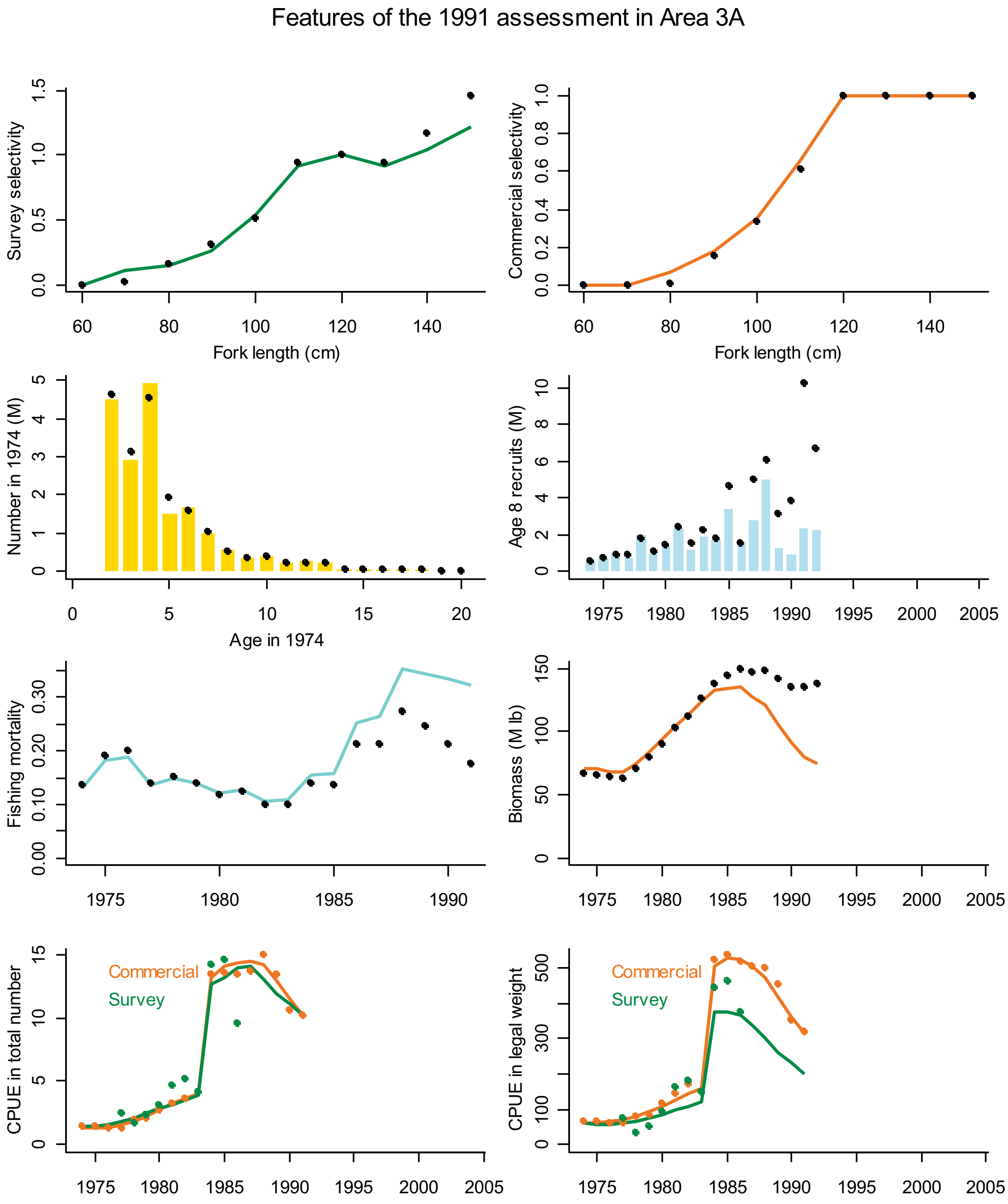


Figure 5. Estimates obtained by fitting the assessment model to data from 1974-1991 (lines in the upper sex panels) and to the full 1974-2004 data set (points in the upper six panels). In the bottom two panels the points are observed CPUE values and the lines are predictions from the 1974-1991 model fit.

Appendix A. Selected fishery and survey data summaries.

Table A1. Commercial catch (million pounds, net weight). Figures include IPHC research catches. Sport catch in Areas 2A and 2B is *not* included in this table.

	2A	2B	2C	3A	3B	4	4A	4B	4C	4D	4E	Total
1974	0.52	4.62	5.60	8.19	1.67	0.71	---	---	---	---	---	21.31
1975	0.46	7.13	6.24	10.60	2.56	0.63	---	---	---	---	---	27.62
1976	0.24	7.28	5.53	11.04	2.73	0.72	---	---	---	---	---	27.54
1977	0.21	5.43	3.19	8.64	3.19	1.22	---	---	---	---	---	21.88
1978	0.10	4.61	4.32	10.30	1.32	1.35	---	---	---	---	---	22.00
1979	0.05	4.86	4.53	11.34	0.39	1.37	---	---	---	---	---	22.54
1980	0.02	5.65	3.24	11.97	0.28	0.71	---	---	---	---	---	21.87
1981	0.20	5.66	4.01	14.23	0.45	---	0.49	0.39	0.30	0.01	0.00	25.74
1982	0.21	5.54	3.50	13.52	4.80	---	1.17	0.01	0.24	0.00	0.01	29.01
1983	0.26	5.44	6.38	14.14	7.75	---	2.50	1.34	0.42	0.15	0.01	38.39
1984	0.43	9.05	5.87	19.77	6.69	---	1.05	1.10	0.58	0.39	0.04	44.97
1985	0.49	10.39	9.21	20.84	10.89	---	1.72	1.24	0.62	0.67	0.04	56.10
1986	0.58	11.22	10.61	32.80	8.82	---	3.38	0.26	0.69	1.22	0.04	69.63
1987	0.59	12.25	10.68	31.31	7.76	---	3.69	1.50	0.88	0.70	0.11	69.47
1988	0.49	12.86	11.36	37.86	7.08	---	1.93	1.59	0.71	0.45	0.01	74.34
1989	0.47	10.43	9.53	33.74	7.84	---	1.02	2.65	0.57	0.67	0.01	66.95
1990	0.32	8.57	9.73	28.85	8.69	---	2.50	1.33	0.53	1.00	0.06	61.60
1991	0.36	7.19	8.69	22.93	11.93	---	2.26	1.51	0.68	1.44	0.10	57.08
1992	0.44	7.63	9.82	26.78	8.62	---	2.70	2.32	0.79	0.73	0.07	59.89
1993	0.50	10.63	11.29	22.74	7.86	---	2.56	1.96	0.83	0.84	0.06	59.27
1994	0.37	9.91	10.38	24.84	3.86	---	1.80	2.02	0.72	0.71	0.12	54.73
1995	0.30	9.62	7.77	18.34	3.12	---	1.62	1.68	0.67	0.64	0.13	43.88
1996	0.30	9.54	8.87	19.69	3.66	---	1.70	2.07	0.68	0.71	0.12	47.34
1997	0.41	12.42	9.92	24.63	9.07	---	2.91	3.32	1.12	1.15	0.25	65.20
1998	0.46	13.17	10.20	25.70	11.16	---	3.42	2.90	1.26	1.31	0.19	69.76
1999	0.45	12.70	10.14	25.32	13.84	---	4.37	3.57	1.76	1.89	0.26	74.31
2000	0.48	10.81	8.44	19.27	15.41	---	5.16	4.69	1.74	1.93	0.35	68.29
2001	0.68	10.29	8.40	21.54	16.34	---	5.01	4.47	1.65	1.84	0.48	70.70
2002	0.85	12.07	8.60	23.13	17.31	---	5.09	4.08	1.21	1.75	0.56	74.66
2003	0.82	11.79	8.41	22.75	17.23	---	5.02	3.86	0.89	1.96	0.42	73.19
2004	0.89	12.16	10.30	25.05	15.61	---	3.48	2.71	0.96	1.67	0.31	73.13

Table A2. Bycatch mortality of legal-sized halibut (80+ cm; in million pounds net weight).

	2A	2B	2C	3A	3B	4	4A	4B	4CDE	Total
1974	0.252	0.900	0.371	4.477	2.816	1.892	---	---	---	10.708
1975	0.252	0.902	0.451	2.610	1.661	1.097	---	---	---	6.973
1976	0.252	0.941	0.503	2.741	1.944	1.181	---	---	---	7.562
1977	0.254	0.725	0.407	3.366	1.544	1.976	---	---	---	8.272
1978	0.253	0.551	0.213	2.441	1.308	3.400	---	---	---	8.166
1979	0.253	0.694	0.638	4.488	0.688	3.446	---	---	---	10.207
1980	0.253	0.514	0.418	4.927	0.870	5.713	---	---	---	12.695
1981	0.252	0.533	0.403	3.989	1.096	4.369	---	---	---	10.642
1982	0.252	0.299	0.199	3.197	1.683	2.944	---	---	---	8.574
1983	0.253	0.291	0.200	2.083	1.218	2.472	---	---	---	6.517
1984	0.252	0.516	0.211	1.508	0.919	2.291	---	---	---	5.697
1985	0.252	0.548	0.201	0.797	0.341	2.246	---	---	---	4.385
1986	0.253	0.558	0.202	0.674	0.197	2.617	---	---	---	4.501
1987	0.253	0.793	0.202	1.588	0.396	2.674	---	---	---	5.906
1988	0.253	0.773	0.202	2.126	0.042	3.273	---	---	---	6.669
1989	0.253	0.720	0.202	1.805	0.437	1.944	---	---	---	5.361
1990	0.253	1.029	0.674	2.633	1.215	---	0.625	0.335	2.385	9.149
1991	0.253	1.221	0.546	3.126	1.035	---	0.731	0.236	2.237	9.385
1992	0.276	1.017	0.574	2.644	1.116	---	0.724	0.655	1.937	8.943
1993	0.276	0.651	0.333	1.919	0.466	---	0.140	0.479	1.407	5.671
1994	0.276	0.571	0.396	2.352	0.848	---	1.197	0.536	1.820	7.996
1995	0.381	0.705	0.219	1.460	0.825	---	1.087	0.149	2.116	6.942
1996	0.473	0.166	0.233	1.403	0.960	---	0.594	0.459	2.991	7.279
1997	0.473	0.109	0.240	1.549	0.729	---	0.844	0.198	2.964	7.106
1998	0.834	0.117	0.238	1.471	0.731	---	1.193	0.327	2.725	7.636
1999	0.761	0.107	0.230	1.283	0.743	---	0.909	0.336	2.642	7.011
2000	0.634	0.128	0.254	1.286	0.646	---	0.808	0.580	2.279	6.615
2001	0.645	0.149	0.184	1.617	0.632	---	0.574	0.387	2.900	7.088
2002	0.286	0.152	0.166	1.073	0.719	---	0.534	0.196	2.735	5.861
2003	0.355	0.133	0.144	1.177	0.500	---	0.515	0.219	2.105	5.148
2004	0.367	0.140	0.149	1.520	0.393	---	0.516	0.294	1.915	5.294

Table A3. Commercial CPUE (net pounds per skate).

Values before 1984 are raw J-hook catch rates, with no hook correction. 1983 is excluded because it consists of a mixture of J- and C-hook data. No value is shown for area/years after 1980 with fewer than 500 skates of reported catch/effort data.

	2A	2B	2C	3A	3B	4A	4B	4C	4D	4E
J-hook CPUE:										
1974	59	64	57	65	57	---	---	---	---	---
1975	59	68	53	66	68	---	---	---	---	---
1976	33	53	42	60	65	---	---	---	---	---
1977	83	61	45	61	73	---	---	---	---	---
1978	39	63	56	78	53	---	---	---	---	---
1979	50	48	80	86	37	---	---	---	---	---
1980	37	65	79	118	113	---	---	---	---	---
1981	33	67	145	142	160	158	99	110	---	---
1982	22	68	167	170	217	103	---	91	---	---
1983	---	---	---	---	---	---	---	---	---	---
C-hook CPUE:										
1984	63	148	314	524	475	366	161	---	197	---
1985	62	147	370	537	602	333	234	---	330	---
1986	60	120	302	522	515	265	---	427	239	---
1987	57	131	260	504	476	341	220	384	---	---
1988	134	137	281	503	655	453	224	---	201	---
1989	124	134	258	455	590	409	268	331	384	---
1990	168	175	269	353	484	434	209	288	381	---
1991	158	148	233	319	466	471	329	223	398	---
1992	115	171	230	397	440	372	278	249	412	---
1993	147	208	256	393	514	463	218	257	851	---
1994	93	215	207	353	377	463	198	167	480	---
1995	116	219	234	416	476	349	189	---	475	---
1996	159	226	238	473	556	515	269	---	---	---
1997	226	241	246	458	562	483	275	335	671	---
1998	194	232	236	451	611	525	287	287	627	---
1999	---	213	199	437	538	500	310	270	535	---
2000	263	229	186	443	577	547	318	223	556	---
2001	169	226	196	469	431	474	270	203	511	---
2002	181	222	244	507	399	402	245	148	503	---
2003	184	231	233	487	364	355	196	105	389	---
2004	142	227	239	479	326	321	205	124	456	---

Table A4. IPHC setline survey CPUE of legal sized fish in weight (net pounds per skate).

Figures for Area 2B refer to the Charlotte region only. Figures for all other areas refer to all stations fished. The eastward expansion of the 3A survey in 1996 lowered average CPUE by around 25%; the raw values in the table should not be taken at face value. Similarly the 4A value for 1999 is elevated because the Bering Sea edge in 4A was not fished that year. *No corrections* are applied; J-hook values are raw J-hook catch rates.

	2A	2B	2C	3A	3B	4A	4B	4C	4D	4E
J-hook surveys:										
1974	---	---	---	---	---	---	---	---	---	---
1975	---	---	---	---	---	---	---	---	---	---
1976	---	---	---	---	---	---	---	---	---	---
1977	---	15	---	73	---	---	---	---	---	---
1978	---	21	---	34	---	---	---	---	---	---
1979	---	---	---	51	---	---	---	---	---	---
1980	---	28	---	95	---	---	---	---	---	---
1981	---	18	---	162	---	---	---	---	---	---
1982	---	21	145	180	---	---	---	---	---	---
1983	---	20	142	147	---	---	---	---	---	---
1984	---	28	---	217	---	---	---	---	---	---
C-hook surveys:										
1984	---	64	260	446	---	---	---	---	---	---
1985	---	47	260	466	---	---	---	---	---	---
1986	---	42	283	377	---	---	---	---	---	---
1987	---	---	---	---	---	---	---	---	---	---
1988	---	---	---	---	---	---	---	---	---	---
1989	---	---	---	---	---	---	---	---	---	---
1990	---	---	---	---	---	---	---	---	---	---
1991	---	---	---	---	---	---	---	---	---	---
1992	---	---	---	---	---	---	---	---	---	---
1993	---	105	---	323	---	---	---	---	---	---
1994	---	---	---	313	---	---	---	---	---	---
1995	29	166	---	370	---	---	---	---	---	---
1996	---	175	306	317	352	---	---	---	---	---
1997	35	156	411	331	415	237	282	71	111	---
1998	---	92	232	281	435	310	216	---	---	---
1999	37	95	204	241	438	382	203	---	---	---
2000	---	104	233	272	373	286	216	---	213	---
2001	41	117	237	256	357	207	171	---	197	---
2002	33	107	261	299	297	174	119	---	257	---
2003	22	84	223	229	262	159	104	---	195	---
2004	27	99	173	271	236	142	73	---	132	---

Analysis of the constant harvest rate policy for 2005

Steven R. Hare and William G. Clark

Abstract

The constant harvest rate policy is updated and investigation of an appropriate target harvest rate conducted. Spawning biomass limit and threshold reference points are established for IPHC Regulatory Areas 2B, 2C, and 3A. Simulations are conducted using a sex-specific model and selectivities estimated from the 2004 stock assessment. Consideration is given to the impact on yield and spawning biomass. A target harvest rate of 0.225, down slightly from last year's rate of 0.25, is recommended for IPHC Regulatory Areas 2 and 3A. A target harvest rate of 0.225 provides a high fraction of yields at higher harvest rates but greatly decreases the frequency at which the spawning biomass would drop below the threshold.

Introduction

In 2002 and 2003, IPHC staff developed an alternative harvest policy, termed the Conditional Constant Catch (CCC) harvest policy (Hare and Clark 2003, Clark and Hare 2004, Hare and Clark 2005a). Following extended discussion at the 2004 IPHC Annual Meeting, implementation of the CCC policy was delayed until a special workshop could be held and the features of the policy more thoroughly examined. That workshop was held in September 2004 and details of that meeting are summarized in Leaman and Hare (2004, appended). The policy represented a hybrid of a constant catch and a constant harvest rate policy, applying a constant harvest rate at intermediate stock sizes and a constant catch limit, or "cap" at higher stock sizes. In addition, biomass reference points were implemented which scaled down the harvest rate at low spawning biomass levels. The use of harvest caps was deemed undesirable by the industry who felt that the current system is adequate to prevent excessive removals and that the stated motivation of the CCC policy – minimizing variability in annual quotas – was no longer an industry priority. The use of biomass reference points, however, was deemed highly desirable and strong encouragement was given to establishing and using these reference points in all future harvest rate policy evaluations.

The harvest policy now in use at the IPHC for establishing annual quotas is a modified constant harvest rate policy. A range of constant harvest rates are examined to determine the impact on yield, exploitable biomass (EBio) and spawning biomass (SBio). In all analyses, two biomass reference points, termed the "threshold" and the "limit" are employed. Both terms refer to female SBio. The SBio limit is the minimum observed historical SBio in each area: 9 million pounds in Area 2B, 13 million pounds in Area 2C and 42 million pounds in Area 3A. The threshold is 150% of the limit. The full constant (or target) harvest rate applies as long as SBio is above the threshold. The harvest rate is scaled down linearly from the full rate at the threshold to a rate of zero when SBio reaches the limit. In simulation work, this methodology proved completely effective in insuring that SBio never reached the limit and helping return SBio above the threshold relative rapidly (Hare and Clark 2003).

Methods

For the analysis of the harvest rate policy this year, performance over two different timelines is considered: short-term (six years, projecting from current conditions as estimated in the stock assessment) and long term (100 years, projecting from equilibrium conditions). Over the past 20 years, the IPHC has employed a constant harvest rate hanging from 0.20 to 0.38. In this analysis, we consider harvest rates ranging from 0.20 to 0.30.

For the long term simulations, the basic dynamics and modeling remain as described in recent CCC harvest policy analyses (Hare and Clark 2003, Clark and Hare 2004, Hare and Clark 2005a). The simulations are conducted for IPHC Areas 2B, 2C, and 3A, i.e., the areas for which we have long-term reliable estimates of growth and recruitment. A reference set of simulations and results are developed for the "Most Likely" scenario. Under this scenario, recruitment is modeled as alternating between high and low productivity regimes, with a factor of two difference in the respective recruitment levels. The duration of recruitment regimes is randomly varied from 15-30 years and recruitment variability is generated using a variance of 0.16 and lag-1 autocorrelation of 0.1. Growth is modeled as a density dependent process, with annual growth a function of adult halibut (age 10+) numbers in the population. Recruitment to the three areas is assumed to follow the pattern observed over the past 25 years: 20% of total 2B/2C/3A recruits to Area 2B, 20% to Area 2C and 60% to Area 3A. Selectivity is a fixed function of length, linearly increasing from 0 at 80 cm to 1 at 120 cm for Areas 2C and 3A and increasing from 0 at 75 cm to 1 at 110 cm in Area 2B. This different selectivity schedule for Area 2B is inherited from the stock assessment where the higher selectivity (relative to Areas 2C and 3A) has been consistently observed over time.

For each harvest rate and area, simulations run forward for 150 years to establish equilibrium conditions and performance statistics are tabulated for the next 100 years. Two hundred Monte Carlo replicates are run and results are averaged across replicates. Many population and catch indices are tracked in the simulations; for purposes of selecting a harvest rate, four sets of indicators are used: average catch, frequency of SBio reaching the threshold, realized average harvest rate, and long term average SBio relative to unfished level. Other indicators of interest but not reported here include, e.g., female proportion in the catch, numbers of age 20+ fish remaining in the population, average weight of fish in the catch, etc. In addition to reporting results for the "Most Likely" scenario, a second set of results are shown for an alternative scenario – the "Low Growth" scenario. Under this scenario, it is assumed that the current low growth rates – attributed to large numbers of fish in the population - are instead the result of some fundamental ecosystem change. Alternatively, a low growth rate might occur if the halibut population had been "culled" of fish with a genetic disposition towards rapid growth. This alternative scenario is believed to be the most realistic alternative scenario. In previous analysis, other scenarios were examined, including redistributed recruitment among areas and continuous low recruitment levels (Hare and Clark 2003). Consideration of the many alternative scenarios will be revisited in future harvest evaluations.

For the short term simulations, computations were made projecting forward from the numbers and weights at age as estimated in the most recent stock assessment (Clark and Hare 2005). Over this limited time horizon, the values used for incoming recruitment have little impact on estimates as the animals contributing to SBio, EBio, and catches are those aged 8-15 as of the 2004 assessment. Growth rates over the next six years were assumed to be static, reflecting growth rates of the past few years. This assumption also has little effect on the results as rates would not change greatly at the intermediate harvest rates examined here. One hundred simulations were conducted and average catch, SBio and EBio over the next six years were computed for harvest rates from 0.2 to 0.3

Results and Discussion

The results of the simulations for the Most Likely and Low Growth scenarios are summarized in Table 1. The results show the expected pattern of increasing catch and decreasing spawning biomass as harvest rate increases. The question at hand is where to set the harvest rate such that two objectives – obtaining high yield and maintaining a healthy spawning biomass – are met. Before selecting and explaining our choice of a judicious harvest rate, we first discuss in general terms the IPHC's precautionary philosophy.

The IPHC considers first and foremost the impact of a harvest level on SBio. The approach taken is one of avoidance of dropping below the minimum observed historical level. This is different from the philosophy adopted by NMFS which has a harvest control rule based on a more theoretical construct: spawning biomass per recruit. Within the three areas being analyzed, halibut populations rebounded from the minimum SBios of the early 1970s to the high levels observed for the past 15-20 years. We can have some confidence therefore of stock dynamics at those levels of SBio, but no experience at lower levels. There is no compelling reason to allow SBio to drop below the minimum limit. Recent work, based on our current conception of halibut population dynamics, indicates that high harvest rates too often result in sustained periods of low spawning biomass. In Clark and Hare (2004), a maximum harvest rate of 0.40 was investigated; values higher than this sometimes drove SBio below the limit, but values equal to or less did not. Thus, a harvest rate of 0.40 functions in the same manner as the "maximum fishing mortality threshold" that is defined under National Standard 1 for NMFS' managed groundfish stocks. By that definition, harvest rates above the reference value of 0.40 would constitute "overfishing". By restricting allowable harvest rates to the range of 0.20 to 0.30, allowance is made for observation error in estimates of exploitable biomass. Analysis of retrospective patterns in halibut assessments indicate initial stock biomass errors have a coefficient of variation of 10-15% (Clark and Hare 2005). Thus, even with a persistent overestimate of the true stock biomass, restriction of harvest rates to a maximum of 0.30 would ensure that the maximum rate of 0.40 would not be reached.,

The IPHC adopted a constant harvest rate policy in 1985 and initially adopted a rate of 0.35. On the basis of subsequent analyses, the rate was lowered, first to 0.30 in 1993 and then to 0.20 in 1996. For the 2004 season, a provisional rate of 0.25 was adopted pending the results of further analyses using a sex-specific simulation model.

For 2005, we recommend a return to a slightly more conservative harvest rate of 0.225. Under the Most Likely scenario, harvesting at or below a rate of 0.20 maintains SBio above the threshold almost 100% of the time Above a rate of 0.20, the frequency of dropping below the threshold begins to rise steeply (Figure 1). Depending on the area, at a harvest rate of 0.25, the SBio would drop below the threshold between 21 and 29% of the time. At the slightly lower harvest rate of 0.225, the percentage of time SBio is below the threshold is 50-66% lower than at a rate of 0.25. If one considers the Low Growth scenario, even a rate of 0.225 would result in SBio dropping below the threshold 20-30% of the time and a rate of 0.25 increases the frequency even further. The expected annual yield amounts for the three areas are 4-9% lower at a harvest rate of 0.25 compared to a rate of 0.225. This would appear to be a modest tradeoff of yield for increased assurance of a SBio more often above the threshold. The actual average harvest rate over time is lower than the target harvest rate if the SBio threshold is reached and reduction in harvest rate triggered. Harvest rates above 0.25 result in actual harvest rates considerably lower than the target rate. Target harvest rates of 0.225 and 0.25 lead to actual harvest rates that are 96-99% of the target rate. One other indicator that we considered is average long term SBio as a fraction of unfished SBio. A number of published studies have suggested that for groundfish, average SBio should remain in the range of 20-60% of unfished level (Clark 1991). These simulations show that at harvest rates around 0.25, the long term average SBio drops below 20% in 2B and is close in 2C. It should be noted that even a harvest rate of 0.20 pushes average SBio down to 25-35% of unfished level.

Finally, we also note that a Yield Per Recruit and Spawning Biomass Per Recruit analysis was conducted and on that basis a harvest rate in the range of 0.15-0.20 would be recommended using the harvest control rule adopted by NMFS (Hare and Clark 2005a). As argued in that analysis, however, we feel that a rate based on the dynamic simulations performed here provided a better basis for selecting an appropriate harvest rate.

Results of the short term simulations are given in Figures 2-4. In Area 2B, SBio and EBio are both projected to remain at approximately current levels at harvest rates around 0.225. Expected annual yield is projected to rise slightly from 12.9 to 13.2 million pounds between 2005 and 2010. In Area 2C, SBio, EBio and annual yield are all projected to slowly decline over the next six years for harvest rates of 0.2 and higher. At a harvest rate of 0.225, expected yield declines from 14.8 to 14.4 million pounds over the six-year projection window. A relatively strong decline in SBio, EBio and yield is projected for the next six years in Area 3A. At a harvest rate of 0.225, annual yield declines by more than four million pounds between 2005 and 2010. A higher harvest rate would result in an even steeper decline in yield. For all three areas, SBio is projected to remain well above the threshold that would trigger a reduction in harvest rate.

References

Clark, W.G. 1991. Groundfish exploitation rates based on life history parameters. Can. J. Fish. Aquat. Sci. 48: 734-750.

Clark, W.G. and Hare, S. R. 2004. A conditional constant catch policy for managing the Pacific halibut fishery. N. Am. J. Fish. Mgmt. 24:106-113.

Clark, W.G. and Hare, S.R. 2005. Assessment of the Pacific halibut stock at the end of 2004. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2004: this volume.

Hare, S.R. and Clark, W.G. 2003. Issues and tradeoffs in the implementation of a conditional constant catch harvest policy. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2002: 121-161.

Hare, S.R. and Clark, W.G. 2005a. Yield per recruit analysis for a sex specific halibut assessment model. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2004: this volume.

Hare, S.R. and Clark, W.G. 2005b. Updated simulation analysis of the CCC harvest policy with separate accounting of males and females. Report of Assessment and Research Activities 2004: this volume.

Table 1. Performance statistics for a range of harvest rates under the Most Likely (Density Dependent (DD) Growth) and Low Growth scenarios.

DD Growth

	Average Annual Yield (million pounds)		
HR	2B	2C	3A
0.000	0	0	0
0.200	15.3	16.2	45.0
0.225	16.0	17.1	49.2
0.250	16.6	17.8	52.9
0.275	17.1	18.3	56.1
0.300	17.5	18.8	58.9

	% of time Sbio reaches Threshold		
HR	2B	2C	3A
0.000	0	0	0
0.200	0	2	4
0.225	7	12	15
0.250	21	27	29
0.275	34	37	39
0.300	42	43	46

	Average Actual harvest rate		
HR	2B	2C	3A
0.000	0.000	0.000	0.000
0.200	0.200	0.200	0.199
0.225	0.224	0.222	0.222
0.250	0.244	0.242	0.240
0.275	0.261	0.258	0.256
0.300	0.276	0.273	0.270

	Average Spawning Biomass (fraction of HR=0.00 biomass)		
HR	2B	2C	3A
0.000	105.0	122.6	268.0
0.200	0.24	0.27	0.36
0.225	0.20	0.24	0.32
0.250	0.18	0.22	0.30
0.275	0.16	0.20	0.28
0.300	0.15	0.18	0.26

Low Growth

	Average Annual Yield (million pounds)		
HR	2B	2C	3A
0.000	0	0	0
0.200	11.5	11.5	23.5
0.225	11.9	11.9	24.4
0.250	12.2	12.2	25.2
0.275	12.5	12.4	25.8
0.300	12.7	12.6	26.4

	% of time Sbio reaches Threshold		
HR	2B	2C	3A
0.000	0	0	0
0.200	6	10	22
0.225	19	21	31
0.250	31	32	37
0.275	39	39	41
0.300	45	44	45

	Average Actual harvest rate		
HR	2B	2C	3A
0.000	0.000	0.000	0.000
0.200	0.199	0.198	0.193
0.225	0.220	0.218	0.211
0.250	0.238	0.236	0.228
0.275	0.253	0.252	0.244
0.300	0.268	0.267	0.258

	Average Spawning Biomass (fraction of HR=0.00 biomass)		
HR	2B	2C	3A
0.000	100.4	117.3	250.7
0.200	0.22	0.27	0.36
0.225	0.20	0.24	0.34
0.250	0.18	0.22	0.32
0.275	0.17	0.21	0.30
0.300	0.15	0.19	0.29

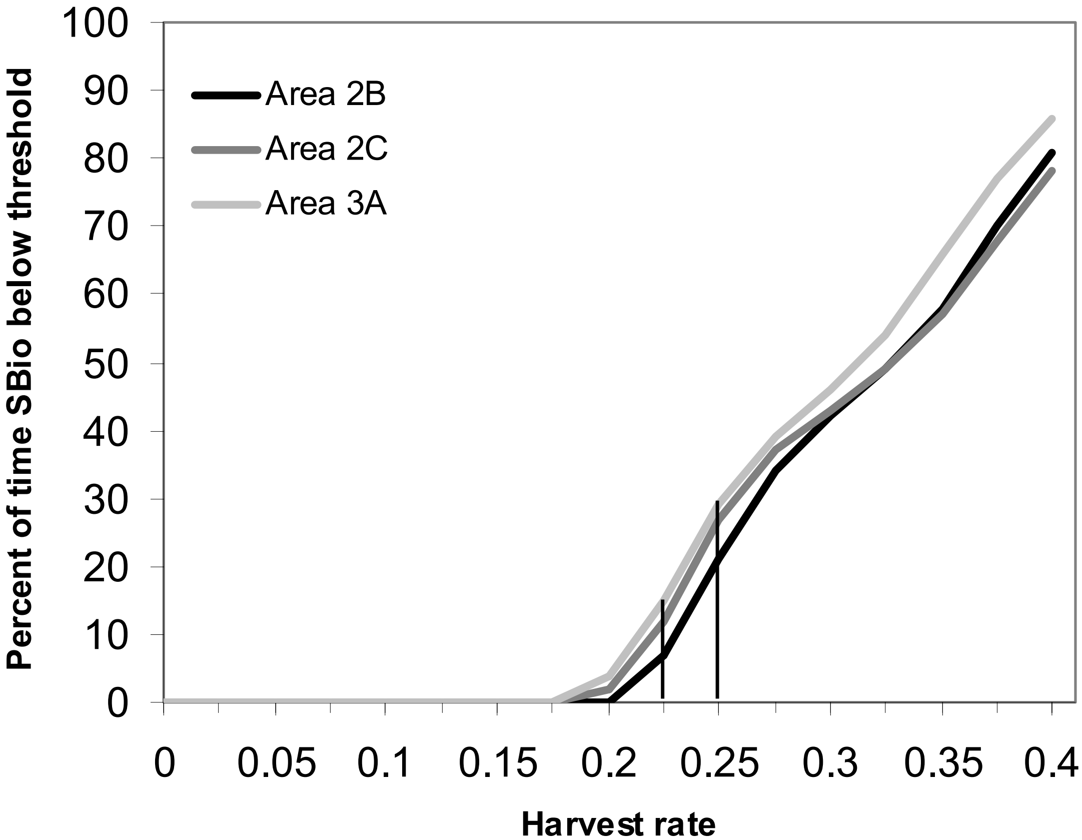


Figure 1. Percent of time that spawning biomass (SBio) drops below the threshold with increasing harvest rate. Vertical lines are drawn at harvest rates of 0.225 and 0.25.

	Average Annual Yield (M net lbs.)				
HR	0.200	0.225	0.250	0.275	0.300
2005	11.5	12.9	14.3	15.8	17.2
2006	11.7	13.0	14.2	15.4	16.5
2007	12.0	13.1	14.2	15.1	16.0
2008	12.2	13.1	14.0	14.8	15.5
2009	12.3	13.1	13.9	14.6	15.1
2010	12.4	13.2	13.9	14.4	14.9
6 yr. avg.	12.0	13.1	14.1	15.0	15.9

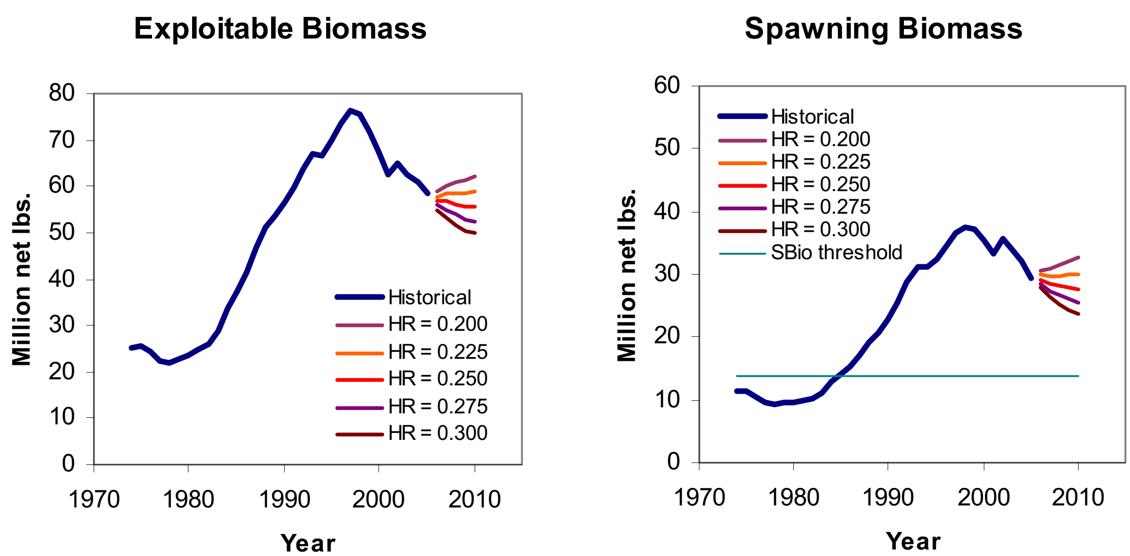
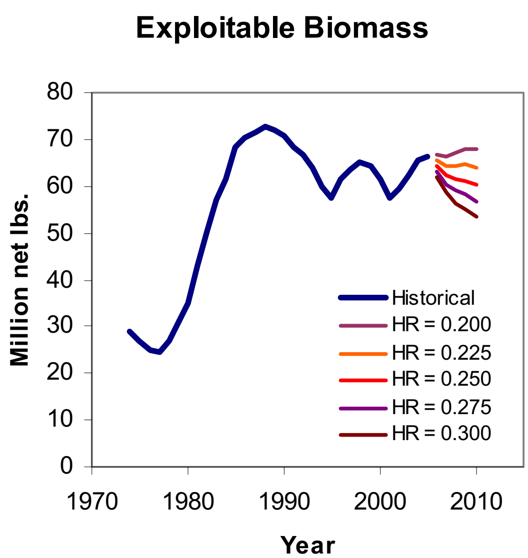


Figure 2. Six year projections of impact on annual yield, exploitable biomass and spawning biomass at a range of harvest rates for Area 2B.

HR	Average Annual Yield (M net lbs.) 0.200	0.225	0.250	0.275	0.300
2005	13.2	14.8	16.5	18.1	19.8
2006	13.3	14.7	16.1	17.4	18.6
2007	13.3	14.5	15.6	16.6	17.6
2008	13.4	14.5	15.4	16.2	16.9
2009	13.6	14.5	15.3	16.0	16.6
2010	13.6	14.4	15.1	15.6	16.1
6 yr. avg.	13.4	14.6	15.7	16.7	17.6



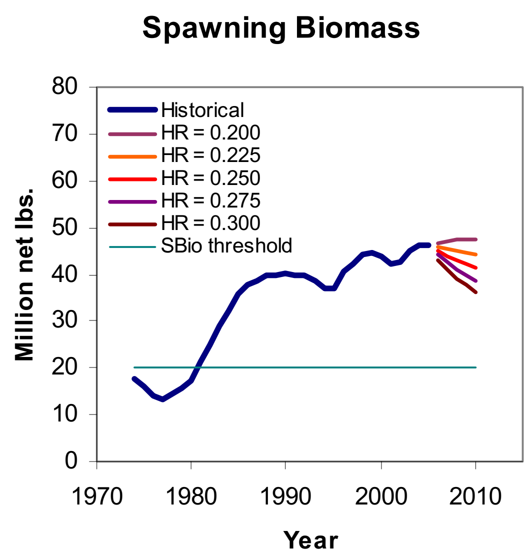
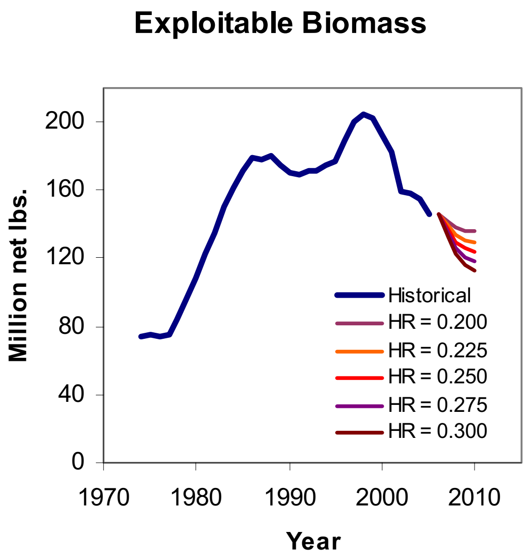


Figure 3. Six year projections of impact on annual yield, exploitable biomass and spawning biomass at a range of harvest rates for Area 2C.

	Average Annual Yield (M net lbs.)				
HR	0.200	0.225	0.250	0.275	0.300
2005	29.2	32.9	36.5	40.2	43.8
2006	28.2	31.2	34.2	37.0	39.7
2007	27.5	30.1	32.5	34.7	36.7
2008	27.2	29.5	31.5	33.3	34.9
2009	27.1	29.1	30.9	32.4	33.8
2010	27.0	28.8	30.4	31.8	33.0
6 yr. avg.	27.7	30.3	32.7	34.9	37.0



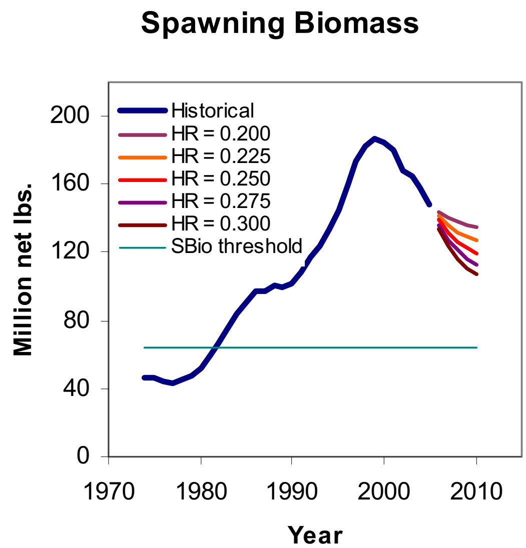


Figure 4. Six year projections of impact on annual yield, exploitable biomass and spawning biomass at a range of harvest rates for Area 3A.

Appendix:
Report of the Conditional Constant Catch harvest policy workshop

Bruce M. Leaman and Steven R. Hare

Introduction

A new harvest policy, called the Conditional Constant Catch (CCC) policy (Clark and Hare, 2004a; Hare and Clark 2003), was developed and introduced by staff at the 2003 International Pacific Halibut Commission (IPHC) Annual Meeting. For the 2004 Annual Meeting, the CCC policy was incorporated into the IPHC staff catch limit recommendations (Leaman and Gilroy 2004, Hare et al. 2004). At that meeting there was considerable uncertainty about the benefits and tradeoffs of the policy, resulting in an industry recommendation to conduct additional evaluation of the policy at a workshop during 2004. Industry and staff agreed that a workshop, isolated from the pressure of deciding on upcoming catch limits, would permit a more thorough examination of the CCC policy attributes. In addition, staff would have the opportunity to continue its development of the policy based on a sex-differentiated population model. The original CCC policy was developed using simulation modeling of a halibut population that was not differentiated by sex. In 2004, a new population assessment model was introduced that provides sex-specific estimates of halibut numbers at age and size at age. Two additional changes were made that strongly affected the population assessment. Selectivity is now modeled as a function of length and all halibut are now aged using the break and burn method. The updated presentation of the CCC policy was completed for the workshop (Hare and Clark 2004).

Workshop attendance and format

The IPHC Conference Board (CB) and Processors' Advisory Group (PAG) nominated members to participate in the workshop. In addition, IPHC Commissioners and Science Advisors were invited to attend. The workshop was held in Seattle on September 27th, 2004. Attending were:

Dr. Jim Balsiger (IPHC Commissioner)	Eric Olsen (CB)
Dan Falvey (CB)	Chris Sporer (CB)
Jim Lane (CB)	Scott Stevenson (CB)
Dr. Loh-Lee Low (U.S. Science Advisor)	Rob Wurm (CB)

The Canadian Science Advisor, Dr. Max Stocker, was unable to attend due to weather effects on travel from Vancouver Island on the day of the meeting. No representatives from the PAG attended, however Executive Director Leaman had met separately with the Halibut Association of North America (HANA) on June 3, 2004. Most members of HANA also sit on the IPHC PAG.

The workshop opened with the updated presentation of the staff's analysis of the CCC policy, based on the sex-specific population assessment model (Hare and Clark 2004). Staff reviewed the operation and benefits of this policy. We used the new simulation model to examine the performance of the CCC harvest policy across a range of catch ceiling and ceiling harvest rates. The range of catch ceilings was updated from the original analyses reflecting new recruitment and size at age estimates. Participants then queried staff on aspects of the policy and expressed their views on its merits. The workshop participants were then censused concerning the features of a harvest policy that they believed to be the most important, from both conservation and operational viewpoints. Lastly, the participants examined potential research projects to address questions raised during the workshop.

Evaluation of the CCC policy and determination of desirable harvest policy features

The CCC harvest policy was developed to provide more stable catch quotas than the constant harvest rate policy used by the IPHC for the past 20 years. A major attraction of the policy is that it is based to a large extent on the long-term productivity of the stock and less so on the annual estimate of exploitable biomass. The reanalysis of the CCC policy, which tracks both sexes and incorporates other new features of the revised stock assessment, generates basically the same results as the original analysis. The policy results in a tradeoff between yield and stability but also incorporates features that provide strong protection for the spawning biomass.

The results of the analysis support a higher ceiling harvest rate than is recommended by a yield per recruit (ypr) analysis (Hare and Clark 2005). That analysis supports harvest rates between 0.15 and 0.20 in order to maintain spawning biomass per recruit above a level of 30-40% of that with no fishing. However, there are a number of limitations to a ypr analysis, the most serious of which is that it does not take into account the dynamic nature of the population in response to changing population size and alternating regimes of recruitment. For that reason, staff believes that the dynamic analysis is the most appropriate and realistic for determining the appropriate harvest policy and policy parameter settings.

Workshop participants agreed that the protection for spawning biomass afforded by the CCC policy was highly desirable but did not agree that the tradeoff between yield and stability was favourable. Harvesters expressed the view that, in the absence of clear conservation benefits, the loss in revenue associated with stability in yield was not a highly valued trait for the harvest policy. Concerns about the impacts of market-driven changes in ex-vessel price per pound are not as acute as they were in 1998, when changes in estimates of available yield, in addition to large cold storage holdings, resulted in dramatic changes in ex-vessel price. Since 1998, changes in the marketing capability of the halibut industry has resulted in less sensitivity of ex-vessel price to changes in halibut landings.

Harvesters believed that the threshold and limit reference point features of the policy should be maintained. They also noted that staff recommendations and the present IPHC catch limit setting procedures act to guarantee an adequate degree of precaution in halibut management. Staff recommendations incorporate a ‘slow up – quick down’ policy when faced with changes in estimates of available yield from the stock, as well as the staff’s judgements on the reliability of the assessment results in each area. The existing constant exploitation yield (CEY) policy developed by the staff incorporates the effects of environmental variability on halibut recruitment, and a harvest rate that is robust to model uncertainty has been chosen. In addition to this formal procedure, the Commission’s advisory bodies have traditionally been conservative about yield changes and the Commission itself has also adopted a conservative approach to stock management. Workshop participants therefore believed that the multiple layers of caution in the IPHC procedures would continue to provide the necessary conservation benefits, and that the stability associated with the CCC policy was not required. However, harvesters did note that shorter-term (e.g. 5-year) projections of recruitment and biomass would be beneficial to their evaluations of harvest policy.

In summary, the workshop participants did not endorse application of the maximum yield caps as part of a harvest policy, but did agree with retention of the threshold and limit reference points.

Dr. Leaman reported on his meeting with HANA. He noted a similar conclusion from the processor representatives as was expressed by harvesters at the workshop. That is, in the absence of any improved conservation benefits associated with the stability features of the CCC policy, stability of yield on its own was not highly beneficial to the processors. HANA members also endorsed the generally conservative IPHC decision-making procedures already in place and believed that these procedures provided adequate precaution in management. Processors further expressed the view that a small degree of uncertainty in the expectation of annual yield might even be beneficial, since there was an element of 'newness' for each year's yield that could be incorporated into marketing.

Workshop participants then turned to consideration of desirable features that should be incorporated in any harvest policy/stock assessment. All participants believed that conservation of spawning biomass was the paramount feature required in any harvest policy and that the limit and threshold reference points should be an integral part of IPHC harvest policy. They also believed that incorporating the most recent understanding of stock status in recommended catch limits was important to industry. The use of maximum yield caps within the CCC policy is somewhat insensitive to present stock status. However, participants noted that application of any policy that is sensitive to annual stock status means either that the assessment must be very accurate or that the harvest rate needs to be sufficiently conservative that mistakes in model configuration will not result in yield estimates that will imperil the stock. This calls for a harvest rate that is conservative. In the same vein, it was thought to be desirable that the recommended catch limits not be overly sensitive to technological changes in the assessment. Lastly, harvesters believed that dialogue on the development of policy was extremely important, both to arrive at the best policy and to achieve understanding of it.

On the topic of assessment, participants want incorporation of the best information and the use of a comprehensive analysis of all areas. They were particularly concerned that the western portions of the stock be subject to as thorough an assessment as possible. While a stable assessment model was viewed as very desirable, it was recognized that stock assessment needs to be a dynamic process and that staff must continue to pursue the best understanding of the stock and its dynamics.

Future work

The workshop participants believed that more understanding of short and medium-term stock behaviour would be valuable and that 3-5 yr projections will help the industry in its evaluation of catch limit recommendations. The optimum harvest rate used in the CEY policy should be re-evaluated with the sex-specific assessment model and the new understanding of selectivity. The impact of harvest policy on the age-specific reproductive characteristics of the stock was discussed by participants and is an element of long-term harvest policy that staff would like to examine. Participants also supported the staff's desire to conduct experiments that evaluate the impact of changes in hook sizes and spacing on commercial data used in the stock assessment. Lastly, the harvesters wished to see estimates of long-term productivity for each area, so that they could evaluate the current status and trajectory of abundance relative to this long-term average.

References

Clark, W.G. and Hare, S.R. 2004a. A conditional constant catch policy for managing the Pacific halibut fishery. N. Am. J. Fish. Mgmt. 24:106-113.

Hare, S.R. and Clark, W.G. 2003. Issues and tradeoffs in the implementation of a conditional constant catch harvest policy. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2002: 121-161.

Hare, S.R. and Clark, W.G. 2004. Updated simulation analysis of the CCC harvest policy with separate accounting of males and females. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2004: this volume.

Hare, S.R. and Clark, W.G. 2005. Yield per recruit analysis for a sex specific halibut assessment model. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2004: this volume.

Hare, S.R., Clark, W.G, and Leaman, B.M. 2004. The conditional constant catch (CCC) harvest policy: summary and estimated CCC yield for 2004. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2003: 201-212

Leaman, B.M. and Gilroy, H.L. 2004. Staff regulatory proposals: 2004. Int. Pac. Halibut Comm. Annual Meeting Bluebook 2004: 89-94.

Analysis of the 1985 hook spacing experiment

Din Chen

Abstract

This paper analyzes the 1985 experiment on the effects of C-hook spacing of 13, 21 and 26 feet on CPUE (lb/100 hooks). We found that the number of fish caught decreased with increased spacing, based on the fixed-length experimental design, whereas the CPUE increases with increasing hook spacing. Statistically there was no significant difference in CPUE among the three spacings in Area 2B. However there was a significant difference in CPUE among the three spacings in Area 3A. We caution that the results are based on a small sample size and the statistical power of the test is low.

Introduction

In 1931, the International Pacific Halibut Commission (IPHC) defined a unit of fishing effort as 1,800 feet of longline gear with the assumption that catch was proportional to length of groundline regardless of the number of hooks. By 1940, the catch was assumed proportional to the number of hooks regardless of the length of groundline, and the unit of effort was redefined as a 6-line skate with 120 hooks. This proportionality had also been accepted in other longline fisheries and in theoretical studies.

From 1970 to 1973, fourteen experiments to test the standard units of effort used in the halibut fishery. Longlines with different J-hook spacings were fished at the same time on the same grounds. The results showed that catch was dependent on the hook spacing and that effort was not proportional to the number of hooks. A new unit of effort, 100 hooks of 18-foot gear, was defined and effort by other longline gear was adjusted to this standard by an empirically determined curvilinear relation between catch per hook and hook-spacing. The results were documented in Hamley and Skud (1978).

Until 1983, the halibut fishery used J-shaped hooks. In that year, a new C-shaped hook which proved to be much more effective in hooking halibut was introduced. Scientific experiments conducted in Alaska and British Columbia in 1984 proved that the C-hook caught 2.2 times as much legal-sized halibut in weight and 2.4 times as much as in number as the J-hook. The stock assessment used a factor 0.45 (1/2.2) to convert the J-hook effort to the present C-hook standard.

In 1985, the IPHC conducted a circle-hook spacing experiment to determine standardization factors for C-hook setline gear of different hook spacing. Two vessels were contracted for this experiment. The *F/V Chelsea* was used for the experiment at the east of Kodiak Island between Albatross and Portlock Banks (Area 3A), from August 23 to August 31. The *F/V Star Wars II* was used in Carpenter Bay and north of Cape Scott (Area 2B), from June 26 to July 4. Three spacings of 13, 21 and 26 feet were fished side-by-side at each of three fishing locations in the two regions.

The purpose of this experiment was to obtain data to compare the CPUE of halibut gear using circle hooks spaced at the three different intervals. The secondary goal of the experiment was to study local depletion of halibut by repeatedly fishing the same location.

Data and methods

For the first goal, the experiment was designed as a Latin-Square experiment (Montgomery 1991) and for the second goal the experiment was repeated on three consecutive days at each location. The data were extracted from IPHC database. Data from *F/V Chelsea* in Area 3A are uniquely identified by set numbers 98-151 and the data from *F/V Star Wars II* in Area 2B are from set numbers 81-134.

Three fishing locations were chosen. At each fishing location, a total six fixed-length skates were fished with two skates for each spacing. Each fishing location was comprised of two of the following hook spacings: 13, 21 or 26 feet. The 'hkspc' column identified daily treatments, with two sets daily for each hook spacing. Stations were designated by a five digit alphanumeric code that spanned both the "stnno" and the "stnpos" fields. The first three digits were numeric comprised the station number: the first was always "9" identifying the experiment as hook spacing; the second was from 1-3, indicating the replicate of the three days; and the third was from 1-6 indicating berth position. The fourth and fifth digits were alphabetic for the station position: the fourth was A-E to distinguish sample sets at each fishing location, and the last digit was either H (Hecate) for sets in Area 2B, or K (Kodiak) for Area 3A. For example 925BK meant hook spacing experiment, 2nd repetition of the 5th berth position in sample set B in the Kodiak area.

The experiment was completed in 1985. However an analysis of the experiment was not conducted. As a component of planning for an experiment in 2005, the data from the 1985 experiment was retrieved for a formal analysis.

Table 1 describes the experimental design for both vessels. For each location (A, B, C for *F/V Star Wars* and D, E, F for *F/V Chelsea*), the experimental fishing was completed on each day based on a Latin-Square design (Montgomery 1991). For the second experiment goal (the depletion), the experimental fishing was repeated for three days at the exact same fishing location. However, the treatments for the *F/V Star Wars* were conducted incorrectly for the second and third days and therefore only the data from the first day can be used for this analysis.

The statistical model as a Latin-Square with replicates can then be formulated as:

$$y_{ijkl} = \mu + \alpha_i + \tau_j + \beta_k + \varepsilon_{ijkl}$$

where y_{ijkl} is the log(CPUE) for i^{th} location, j^{th} treatment, k^{th} berth alignment, l^{th} repetition, with random error ε. In Latin Square terminology, location is the row effect (α), treatment is denoted τ, berth alignment is the column effect (β) with this setting repeated twice. The indices *i*, *j*, *k*, and *l* assume values:

i = location: A, B, C for *F/V Star Wars* and D, E, F for *F/V Chelsea*
j = spacing 13, 21, or 26 feet
k = set 1, 2, or 3.
l = replication 1 to 2 for *F/V Star Wars* and 1 to 6 for *F/V Chelsea*.

It should be noted that the CPUE used here is calculated as the total pounds in 100 hooks per skate not adjusted for hook spacing, in order to compare the spacing effect which is defined as:

$$CPUE = \frac{Catch}{\frac{Num\,Skates \times Num\,Hooks}{100}}.$$

Therefore, the CPUE used here is different from the IPHC standard CPUE which is adjusted to 18 feet hook spacing (Sullivan et al. 1999).

Results

The depletion effect can be observed from the decreasing CPUE series over successive days of the experiments (Figs. 1 and 2). The general trend of increasing CPUE with increasing spacing is also seen in Figures 1 and 2. While there is an increasing trend in CPUE with increasing hook spacing, the total fish caught in number decreases (Fig. 3) with decreasing number of hooks (Tables 2 and 3).

To examine statistical significance, an ANOVA was conducted (Section 5.2 in Montgomery 1991) as in Table 4. There was no statistically significantly difference in CPUE (lb/100 hooks) among the three spacings for the *F/V Star Wars* fishing in Area 2B (Fig. 1). However, there was a significant difference in CPUE among the three spacings in Area 3A for the *F/V Chelsea* (Fig. 2).

Discussion

This analysis indicated no statistically significantly difference in CPUE in Area 2B and a statistically significant difference in 3A, among spacings of 13, 21 and 26 feet. However, this experiment involved a very small sample size and the statistical power is relative low. For a more statistically powerful analysis, it is recommended that an experiment with a larger sample size over a broader area should be conducted. Such an experiment is planned for 2005.

References

Hamley. J. M. and Skud, B. E. 1978. Hooking spacing. Part II in Int. Pac. Halibut Comm. Sci. Rep. 64.

Montgomery, D, C. 1991. Design and Analysis of Experiments. John Wiley & Sons, Inc.

Sullivan, P. J., Parma, A. M. and Clark, W. G. 1999. The Pacific halibut stock assessment of 1997. Int. Pac. Halibut Comm. Sci. Rep. 79.

Table 1. Experimental design for both *F/V Star Wars II* and *F/V Chelsea*. "Location" denotes the fishing locations. "Day" denotes the fishing days of 1, 2 and 3. The numbers under the "Treatment" are the spacing of 13, 21 and 26 feet.

F/V	Location	Day	Treatment					
Star War II	A	1	13	26	21	21	13	26
	A	2	13	21	26	26	21	13
	A	3	21	26	13	13	26	21
	B	1	26	21	13	13	26	21
	B	2	13	26	21	21	13	26
	B	3	21	13	26	26	21	13
	C	1	21	13	26	26	21	13
	C	2	26	21	13	13	26	21
	C	3	13	26	21	21	13	26
Chelsea	D	1	26	13	21	21	13	26
	D	2	13	21	26	26	21	13
	D	3	21	26	13	13	26	21
	E	1	13	21	26	26	21	13
	E	2	21	26	13	13	26	21
	E	3	26	13	21	21	13	26
	F	1	21	26	13	13	26	21
	F	2	26	13	21	21	13	26
	F	3	13	21	26	26	21	13

Table 2. Total catch in weight (lb) from the experiment design corresponding to Table 1.

F/V	Location	Day	Treatment					
Star War II	A	1	633.9	155.3	310.8	373.9	155.9	221.6
	A	2	284.7	317.7	459.3	290.1	310.3	690.2
	A	3	213.1	104.2	316.7	559.6	181.5	949.3
	B	1	221.2	380.3	703.1	2210.5	1972.8	1812.2
	B	2	383.2	122.2	311.2	364.0	1601.4	1311.2
	B	3	285.5	135.5	376.0	729.0	756.3	1211.6
	C	1	263.1	555.2	197.7	880.4	688.6	1458.7
	C	2	600.5	260.1	266.1	361.8	247.3	388.4
	C	3	168.5	112.2	276.1	743.9	687.4	506.3
Chelsea	D	1	3532.6	5871.2	4275.7	4310.0	7916.3	5369.1
	D	2	3822.4	1182.0	1454.3	1526.8	2409.7	5318.5
	D	3	1905.6	1321.7	2861.0	2064.0	1803.7	3113.4
	E	1	4398.4	2251.2	3129.2	2525.4	2576.9	4053.0
	E	2	2207.2	1900.3	1300.4	1137.9	1014.3	2094.3
	E	3	770.1	1043.2	1264.2	1112.9	1908.7	1665.3
	F	1	2601.9	1837.9	2082.3	3022.5	4106.9	5252.8
	F	2	1666.7	1578.4	1287.4	1684.6	3833.4	3275.0
	F	3	1704.2	1394.2	1465.0	1812.8	2979.3	2721.0

Table 3. Total catch in number corresponding to the design in Table 1.

F/V	Location	Day	Treatment					
Star War II	A	1	34	10	17	23	10	13
	A	2	17	19	26	17	22	43
	A	3	12	6	15	38	14	47
	B	1	9	9	23	65	61	66
	B	2	13	5	12	15	53	52
	B	3	10	5	14	26	25	44
	C	1	11	20	8	21	23	43
	C	2	11	11	11	12	9	16
	C	3	6	5	12	23	23	22
Chelsea	D	1	70	83	78	86	139	82
	D	2	84	29	29	29	47	103
	D	3	46	30	50	41	34	62
	E	1	88	48	55	48	60	82
	E	2	53	46	31	40	32	53
	E	3	29	38	40	31	55	38
	F	1	42	37	44	71	52	81
	F	2	27	29	28	39	72	52
	F	3	42	35	36	37	67	61

Table 4. The analysis of variance table for CPUE (lb/100 hooks) on the 1985 hook spacing experiment.

F/V	Source	SSQ	DF	MS	F	p(F)
Star Wars II	**Spacing**	0.081	2	0.040	0.091	**0.913**
	Row	3.968	2	1.984	4.502	0.040
	Cols	0.253	2	0.126	0.287	0.757
	Rep	3.257	1	3.257		
	Error	4.407	10	0.441		
	Total	11.966	17			
Chelsea	**Spacing**	1.407	2	0.703	8.173	**0.001**
	Row	1.505	2	0.753	8.747	0.001
	Cols	1.145	2	0.572	6.652	0.003
	Rep	5.712	5	1.142		
	Error	3.614	42	0.086		
	Total	13.383	53			

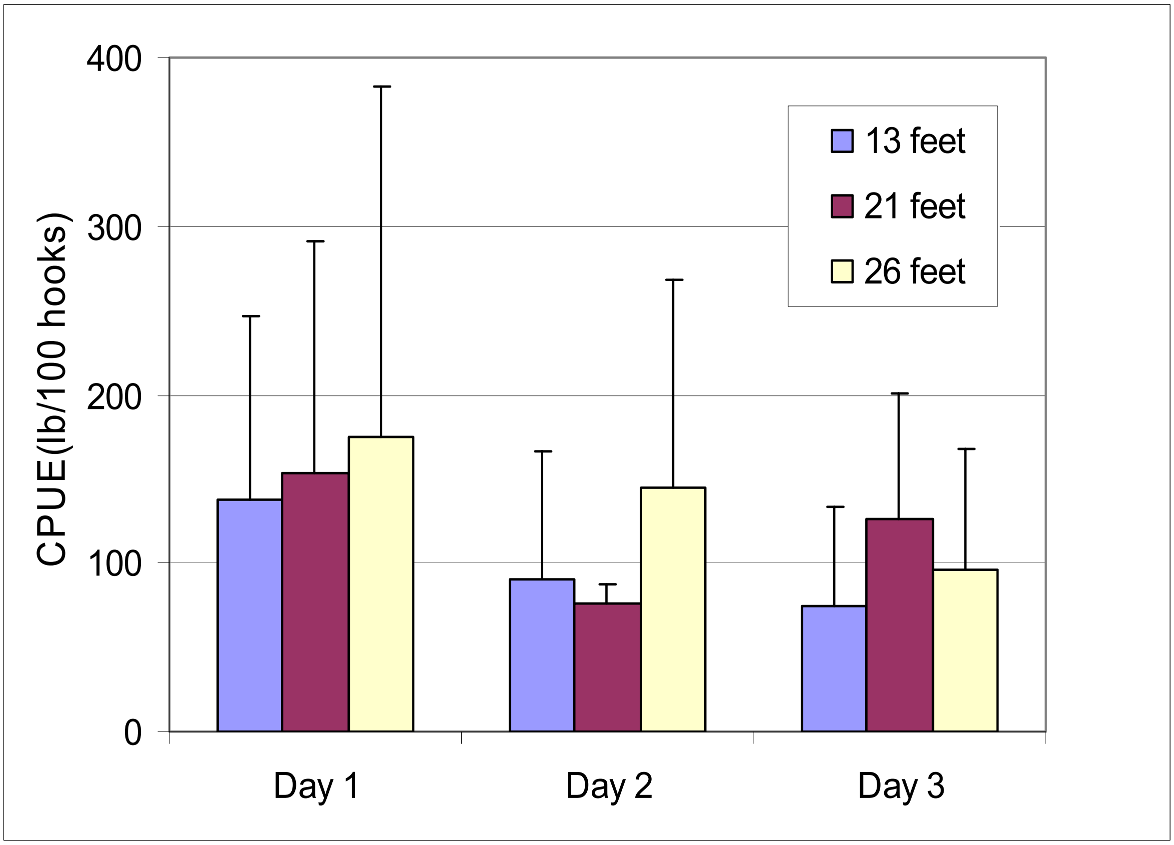


Figure 1. Mean and standard error for CPUE, F/V Star Wars in Area 2B, at hook spacings of 13, 21, and 26 feet.

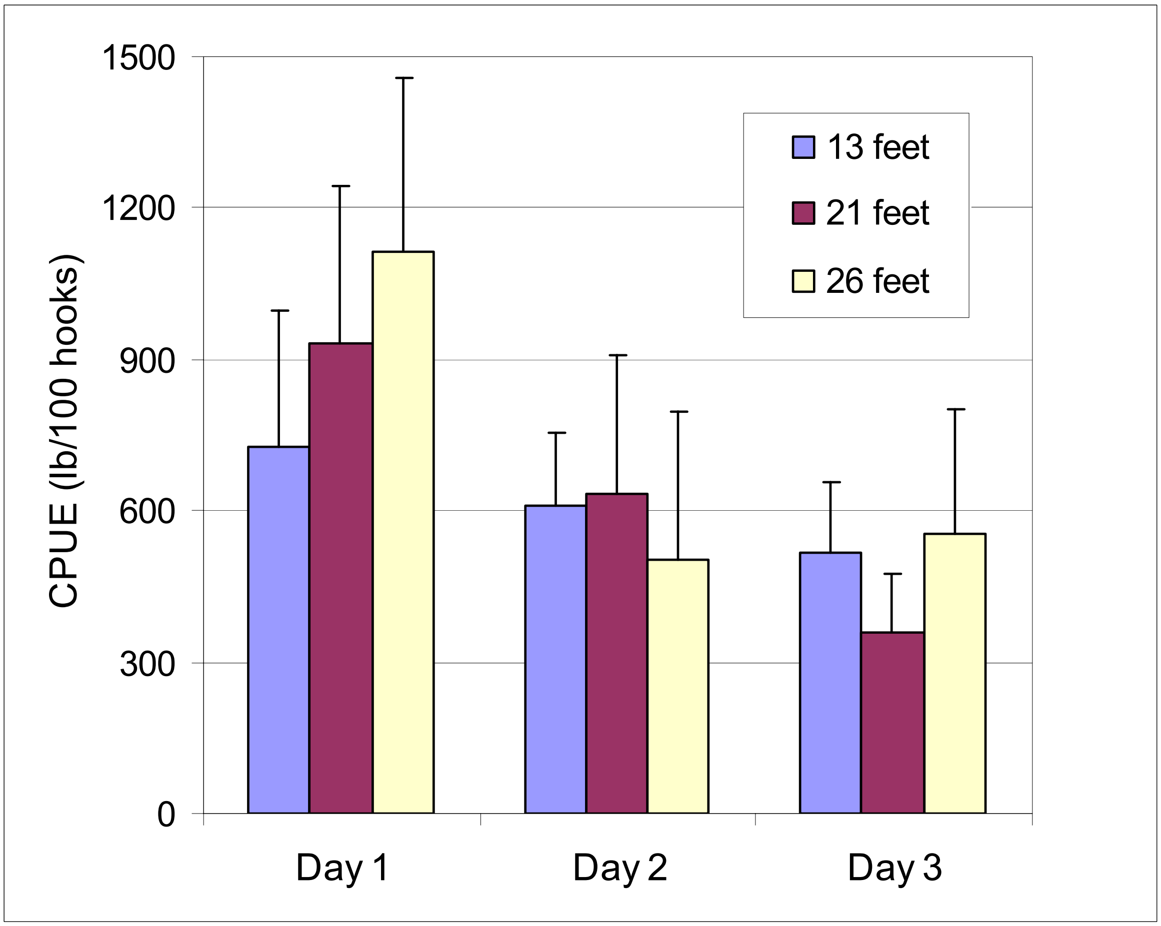


Figure 2. Mean and standard error for CPUE, F/V Chelsea in Area 3A, for hook spacings of 13, 21, and 26 feet.

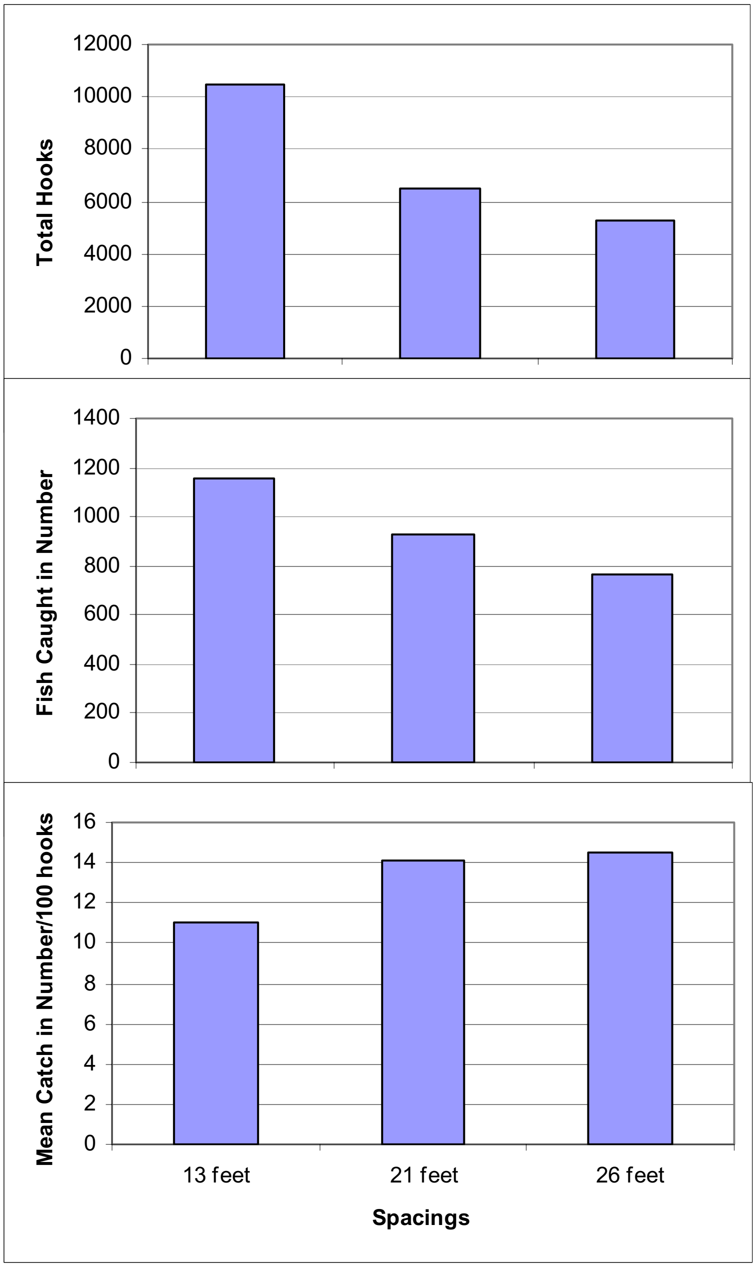


Figure 3. Summary of "Total Hooks", the total "Fish Caught in Number" and the "Mean Catch in Number/100 Hooks" for the three spacings for data from F/V Chelsea.