Population Assessment

Assessment of the Pacific halibut stock at the end of 2007

William G. Clark and Steven R. Hare

Abstract

As in 2006, the stock assessment was done by fitting the assessment model to a coastwide dataset to estimate total biomass, and then apportioning the total among regulatory areas in accordance with survey estimates of relative abundance. Coastwide exploitable biomass in 2008 is estimated to be 361 million pounds, down from the 414 million estimated last year. About half of the decrease is due to a change in the parameterization of survey catchability in the model, and the other half to lower commercial and survey catch rates in 2007. Total CEY is 69 million pounds.

Introduction

Each year the International Pacific Halibut Commission (IPHC) staff assesses the abundance and potential yield of Pacific halibut using all available data from the commercial fishery and scientific surveys (Appendix A). A biological target level for total removals from each regulatory area is calculated by applying a fixed harvest rate to the estimate of exploitable biomass in that area. 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.

For many years the staff assessed the stock in each regulatory area by fitting a model to the data from that area (Appendix B). This procedure relied on the assumption that the stock of fish of catchable size in each area was closed, meaning that net migration was negligible. A growing body of evidence from both the assessments (Clark and Hare 2007a) and the ongoing mark-recapture experiment (Webster and Clark 2007) shows that there is probably a continuing eastward net migration of catchable fish from the western Gulf of Alaska (Areas 3B and 4) to the eastern side (Area 2). The effect of this migration on the closed-area stock assessments was to produce underestimates of abundance in the western areas and overestimates in the eastern areas. To some extent this has almost certainly been the case for some time, meaning that exploitation rates were well above the target level in Area 2 and a disproportionate share of the catches have been taken from there.

In order to obtain an unbiased estimate of the coastwide stock in the 2006 assessment, the staff built a coastwide data set and fitted the model to it. Exploitable biomass in each regulatory area was estimated by apportioning the total in proportion to an estimate of stock distribution derived from the setline survey catch rates (CPUE). Specifically, an index of abundance in each area was calculated by multiplying survey CPUE (running 3-year average) by total bottom area between 0 and 300 fm. The logic of this index is that survey CPUE can be regarded as an index of density, so multiplying it by bottom area gives a quantity proportional to total abundance. The estimated proportion in each area is then the index value for that area divided by the sum of the index values. This year's assessment uses the same procedure.

Description of the assessment model

The IPHC assessment model is age- and sex-structured. Commercial and survey selectivity are both estimated as piecewise linear functions of observed mean length at age/sex in survey catches. (There is a 32" minimum size limit in the commercial fishery.) Commercial catchability is normally allowed to vary from year to year with a penalty of 0.03 on log differences. Survey catchability is normally held constant, although some variation was allowed in both this year's and last year's production fits. The model is fitted to commercial and survey catch at age and CPUE. Clark and Hare (2006) provide a full account of model structure and fitting procedures.

The closed-area and coastwide model fits differ in parameterization and likelihood. Some of the closed-area data sets are quite noisy, so the closed-area version is more parsimonious and it is weighted. Specifically, the catchability, selectivity and natural mortality parameters are all unisex; the estimated selectivity schedules are strongly smoothed; the model is fitted only to total CPUE (rather than CPUE at age/sex); and a heavy weight is placed on the CPUE data series to assure satisfactory agreement. The coastwide data are not noisy, so the coastwide version of the model can have sex-specific parameters, weaker selectivity smoothing, and neutral data weighting. It is fitted to CPUE at age/sex as well as total CPUE.

Alternative model fits

In the 2006 coastwide assessment (Clark and Hare 2007b), estimated survey catchability was allowed to vary somewhat because it was found that actual survey catchability had varied substantially. This was shown by model fits in which present abundance was fixed at a range of levels by fixing the terminal fishing mortality rate as in a virtual population analysis (VPA) and then estimating survey catchability as a free parameter in each year (Fig. 1). These fits showed that survey catchability happened to be high in the first year of the data (1997) and low in the last year (2006), resulting in a spurious appearance of a decline in abundance. To neutralize that feature, survey catchability was estimated independently for the first and last years, which effectively meant disregarding those data points and estimating a constant survey catchability from the remaining data (1998-2005).

In this year's assessment some other ways of dealing with variable survey catchability were considered. The candidate models were:

(i) Vanilla: the conventional model, with constant survey catchability in all years.

(ii) HiLoSQ: last year's production model, with three values of survey catchability estimated (1997, 1998-2005, 2006-2007).

(iii) WobbleSQ: survey catchability estimated for each year, but with a penalty of 0.05 on log differences. This is similar to the treatment of commercial catchability.

(iv) TrendlessSQ: same as WobbleSQ, but with the additional requirement that a regression of estimated survey catchability on year have zero slope. This means that survey catchability was allowed to vary but not to show any trend over time.

Table 1 shows features of the candidate model fits and some others. WobbleSQ has the lowest AIC score, but TrendlessSQ is nearly as good, and we think it is appropriate to disallow trends in survey catchability over time, so that is our chosen production model.

The last two fits in Table 1 show the effect of commercial CPUE on the biomass estimate. “No commercial CPUE” is a fit in which commercial CPUE is disregarded, and “CAGEAN” is a fit in which commercial catchability is held constant, so that commercial and survey CPUE are given equal influence. Evidently commercial CPUE tends to increase the biomass estimate, but not greatly.

Effect of the 2007 data on abundance estimates

Coastwide commercial and survey CPUE both declined by 5-10% from 2006 to 2007 (Fig. 2; Appendix A tables A2 and A3). As a result the 2007 coastwide and closed-area model fits mostly revise downward the estimates of abundance at the beginning of 2007 made in the 2006 assessment (Table 2). At the same time the 2007 fits show an increase in abundance between the beginning of 2007 and the beginning of 2008, so last year’s estimates of 2007 biomass and this year’s estimates of 2008 biomass are not very different in most cases. Exceptions are Areas 2C and 4A where the closed-area estimates decrease significantly.

The coastwide estimate of exploitable biomass in 2008 is 361 M lb compared with 414 last year. About half of this difference is due to the change from the HiLoSQ to the TrendlessSQ model fit. The HiLoSQ biomass estimate in 2008 is 386 M lb.

Area-specific biomass and CEY estimates

Area-specific estimates of biomass are calculated by survey apportionment as they were last year, with the difference that this year a depth-stratified mean survey CPUE has been used, which results in about a 40% increase in the Area 2A apportionment, about a 5% decrease in the Area 3A apportionment, and very small increases in most other apportionments. The area-specific estimates from last year’s and this year’s coastwide and closed-area assessments are shown in Tables 3 and 4.

The staff believes that survey apportionment is the most objective and consistent method of estimating the biomass distribution among areas and therefore the best distribution of total CEY, if the aim is proportional harvest. A disproportionate share of the harvest has been taken from Area 2 for decades, so some level of disproportionality was clearly sustainable by the stock with the exploitation pattern that prevailed during that period. Increasing catches from the western portion of the stock in the last decade have altered the exploitation pattern, so the historical high levels of removals from Area 2 may no longer be sustainable. Alternative CEY apportionments under a variety of rules are shown for information in Table 6. The staff does not advocate any of them and would in fact oppose some, such as apportionment on the basis of bottom area alone or an index incorporating commercial CPUE.

Evaluation of the assessment

Quality of fits

The assessment model fits the coastwide data very well. (That is not true of some of the closed-area data sets.) The series of total survey and commercial CPUE are predicted closely (Fig.

3, bottom panels), and so are the commercial catch and survey CPUE at age/sex (Figs. 4a and 4b).

Retrospective performance

Each year's model fit estimates the abundance and other parameters for all years in the data series. One hopes that the present assessment will closely match the biomass trajectory estimated by the previous year's assessment. To the extent that it does not, the assessment is said to have poor retrospective performance.

Our assessment has not tracked very well for the last few years. Each year the assessment has revised downward the previous year's biomass estimates (Fig. 5), meaning that biomass was overestimated then and may be overestimated now if the cause of the retrospective problem lies somewhere within the model. There is some precedent for that; the assessment models in use in the mid 1990s and the early 2000s showed strong retrospective patterns that turned out to be the result of misspecified selectivity (age- rather than length-based). There is also the possibility that the retrospective pattern is caused in some way by the external estimation of the sex composition of the commercial catch, or by the internal prediction of surface age compositions prior to 2002 through the application of an age misclassification matrix (Clark and Hare 2006).

Problems of this sort with the assessment machinery would manifest themselves as systematic revisions of the estimated relative strength of the year-classes present in the stock. That was true of the retrospective patterns caused by the misspecification of selectivity in the past: incoming year-classes would at first be estimated as weak because catch rates were low, but the real reason was low selectivity rather than low abundance. When they were later caught in large numbers, the estimates of relative year-class strength increased.

We can check for patterns of this sort by doing a blind projection of the assessment from, say, 2004 to 2007. This means using the estimates of year-class strength and other parameters from the 2004 assessment and projecting forward to 2007 without benefit of the 2005-2007 data (except for the total catch in number in each year, which determines the annual fishing mortality rate). If there were some problem with the model, the projected age compositions of the survey and commercial catches would differ systematically from the predictions of the 2007 assessment incorporating the 2005-2007 data. But they do not; the two sets of predicted age compositions are nearly the same (Fig. 6a). This is not surprising, given the simplicity of the model and the very good fits to the data.

What the projection from 2004 fails to predict is the commercial and survey CPUE in 2005-2007 (Fig. 6b). Given the estimates of year-class strength and catchability in 2004, the blind projection shows CPUE bottoming out in 2005 and increasing thereafter. In actuality both declined in 2006 and again in 2007, with the result that the present abundances of all of the year-classes in the stock were revised downward proportionally in the subsequent assessments. So this is a retrospective pattern caused by the data, not by the model.

To some extent the pattern results from the decline in survey catchability mentioned above. VPA-like fits in 2007 show that survey catchability declined every year from 2005 through 2007, by some 20% in total. This is by no means unprecedented, but the run of three declines in a row inevitably affects the biomass estimates. This year's production model ("Trendless") is less affected than a conventional model ("Vanilla") because it allows survey catchability to vary from year to year, but it is affected.

Estimates of uncertainty

There are a number of ways of estimating the uncertainty associated with a given model fit and biomass estimate. They are all unsatisfactory in that they are conditioned on the correctness of the model, and in fact it is the choice of one model rather than another that is the major source of uncertainty in assessments. This is well illustrated by the difference in area-specific biomass estimates between the coastwide and closed-area fits of the IPHC model.

Figure 7 shows probability distributions of the 2008 exploitable biomass obtained in various ways. The Hessian-based estimate of standard deviation is about 20 M lb, and a normal distribution with this amount of dispersion closely approximates a calculated likelihood profile. A straightforward measure of uncertainty is the spread of biomass estimates among plausible models. All of the fits in Table 1 are at least plausible, and they range from 320 to 400 M lb, similar to the Hessian-based normal approximation.

Treatment of process error

The likelihood used in fitting the model is the MULTIFAN scheme developed by Fournier et al. (1990). All errors are treated as being normally distributed, so the likelihood is a sum of squared deviations, each weighted by the inverse of a scaled variance. The variances are the external estimates of sampling variance of each observation, and the scalers are just the root mean squared errors associated with each data type in unscaled fits. This amounts to a one-step reweighting of the data. It succeeds in producing distributions of residuals that are very close to standard normals. The scalers are mostly in the range 4-9, meaning that sampling variance accounts for only a small fraction of the total error variance. The remainder is process error, the result of model misspecification or parameter variation

While the MULTIFAN procedure is clearly effective in standardizing the variances in the halibut assessment, it is somewhat puzzling that process error can be successfully treated as a multiple of sampling error. They arise from different sources and there is really no reason to expect them to be related. One suggestion made during an external review in 2007 was that we consider an additive rather than a multiplicative model of process error. The multiplicative model is $\sigma_p^2 = \left(\tau^2 - 1\right) \cdot \sigma_s^2$, where σ_p^2 is process variance, σ_s^2 is sampling variance, and τ^2 is a scaler. Total variance σ_t^2 is then given by $\sigma_t^2 = \tau^2 \cdot \sigma_s^2$. The additive model is $\sigma_t^2 = \sigma_s^2 + \sigma_p^2$ where σ_p^2 is process error. The suggestion was to estimate a process coefficient of variation (CV) for each data type, so $\sigma_p^2 = \delta^2 \cdot y^2$ where y is the observed value and δ is the CV.

The amount of process error associated with each data point can be estimated as the squared deviation (in an unscaled fit) minus the estimated sampling variance. If the multiplicative model is appropriate, process error should increase with sampling variance, and it does (Fig. 8a). If the additive model is appropriate, process error should increase with the square of the observed value, and it does (Fig. 8b). The reason that both models are appropriate is that most of the observations (commercial and survey catch and CPUE at age/sex) have multinomial sampling variances, so the sampling variances are proportional to the expected values. So while equally appropriate, the additive model would not improve on the multiplicative model.

Use of PIT tag estimates of commercial selectivity in the assessment

Estimates of fishing mortality from the ongoing PIT tag experiment (Webster 2008) are so different from the stock assessment as to be simply incredible, but that is not true of the selectivity estimates. Even when mark-recapture data are not usable for estimating fishing mortality or abundance or migration rates, they can still provide useful estimates of selectivity (Myers and Hoenig 1997, Clark and Kaimmer 2006).

In the stock assessment, commercial selectivity is required to reach 100% at a length of 120 cm and remain there (i.e., commercial selectivity is asymptotic). In model fits, commercial selectivity increases gradually between 80 and 120 cm. At 100 cm it is estimated to be 0.56. The PIT tag data show full commercial selection occurring at a smaller size than the assessment. When a coastwide commercial selectivity is estimated freely from the PIT tag data, it reaches 100% at 100 cm and stays close to that level thereafter (Ray Webster, IPHC, pers. comm.).

The assessment can be made to conform to the PIT tag results by requiring full commercial selection at 100 cm. When that is done, the fit is much worse (AIC = 850 vs 790 for the production model). The exploitable biomass estimate is nearly the same (373 M lb vs 361).

References

Clark, W. G., and Hare, S. R. 2006. Assessment and management of Pacific halibut: data, methods, and policy. Int. Pac. Halibut Comm. Sci. Rep. 83.

Clark, W. G., and Hare, S. R. 2007a. Motivation and plan for a coastwide stock assessment. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006 83-96.

Clark, W. G., and Hare, S. R. 2007b. Assessment of the Pacific halibut stock at the end of 2006. Int. Pac. Halibut Comm. Report of Research and Assessment Activities 2006:97-128.

Clark, W. G., and Kaimmer, S. M. 2006. Estimates of commercial longline selectivity for Pacific halibut (*Hippoglossus stenolepis*) from multiple marking experiments. Fish. Bull. 104:465-467.

Fournier, D. A., Sibert, J. R., Majkowski, J., and Hampton, J. 1990. MULTIFAN, a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data from southern bluefin tuna (*Thunnus maccoyi*). Can. J. Fish. Aquat. Sci. 47:301-317.

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

Webster, R. 2008. Analysis of PIT tag recoveries through 2007. Int. Pac. Halibut Comm. Report of Assessmnent and Research Activities 2007 259-274.

Webster, R., and Clark, W. G. 2007. Analysis of PIT tag reoveries through 2006. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006:129-138.

Table 1. Alternative coastwide model fits. The first two are coastwide fits that have the same parameterization as the closed-area fits.

Model	Number of parameters	Deviance	AIC	Exploitable biomass
Closed-area parameters Closed-area likelihood	121	NA	NA	321
Closed-area parameters Coastwide likelihood	121	716	958	341
Vanilla	136	524	796	337
WobbleSQ	155	479	789	338
HiLoSQ	138	520	796	386
TrendlessSQ	155	480	790	361
No commercial CPUE	145	504	794	344
CAGEAN	134	553	821	400

Table 2. Effect of the 2007 data on closed-area and coastwide abundance estimates.

Area	2007 ebio 2006 assessment Data as of 11/06	2007 ebio 2006 assessment Data as of 11/07	2007 ebio 2007 assessment Data as of 11/07	2008 ebio 2007 assessment Data as of 11/07
Closed-area assessments:				
2A	4.9	5.1	4.0	4.6
2B	39	41	33	37
2C	57	55	45	49
3A	174 [1]	170	169	169
3B	52	53	47	54
4A	17	14	11	11
4B	10	12	15	14
2A-4B sum	354	350	324	339
4CDE	58	52	52	52
Total	412	402	376	391
Coastwide assessment:				
2A-4B sum (90% of total)	339	333	297	325
4CDE	38	37	33	36
Total	377	370	330	361

Notes:

[1] Recalculated to be consistent with present treatment of Area 3A survey CPUE (full-area CPUE = 81% of partial-area CPUE rather than 75%). Value reported last year was 186.

Table 3. Estimates of 2007 exploitable biomass and CEY from the 2006 assessment (2006 RARA, p. 107).

	Area 2A	Area 2B	Area 2C	Area 3A	Area 3B	Area 4A	Area 4B	Area 4CDE	Total
Coastwide assessment [1]									
2007 exploitable biomass	3.7	27	33	176	86	29	19	41	414
Proportion of total	0.009	0.065	0.080	0.423	0.208	0.069	0.045	0.101	1.000
Target harvest rate	0.25	0.25	0.25	0.20	0.20	0.20	0.15	0.15	~0.20
Total CEY	0.93	6.75	8.25	35.20	17.20	5.80	2.85	6.15	83.13
Other removals [2]	0.27	0.53	3.79	7.89	0.43	0.57	0.29	2.30	16.07
2007 fishery CEY [2]	0.66	6.22	4.46	27.31	16.77	5.23	2.56	3.85	67.06
Area assessments [1]									
2007 exploitable biomass	4.9	39	57	186	52	17	10	50	416
Proportion of total	0.012	0.094	0.137	0.447	0.125	0.041	0.024	0.120	1.000
Target harvest rate	0.20	0.20	0.20	0.20	0.20	0.20	0.15	0.15	~0.20
Total CEY	1.00	7.80	11.40	37.20	10.40	3.40	1.50	7.50	80.20
Other removals [2]	0.27	0.53	3.79	7.89	0.43	0.57	0.29	2.30	16.07
2007 fishery CEY [2]	0.73	7.27	7.61	29.31	9.97	2.83	1.21	5.20	64.13
2007 catch limit [3]	1.34	11.47	8.51	26.20	9.22	2.89	1.44	4.10	65.17

Notes:

[1] "Coastwide assessment" refers to the coastwide model fit with survey apportionment of the total biomass estimate among regulatory areas. "Area assessments" are the closed-area model fits.

[2] "Other removals" comprise legal-sized wastage, legal-sized bycatch, personal use, and in most areas sport catch. In Areas 2A and 2B sport catch is included in fishery CEY rather than in other removals.

[3] "Catch limit" includes sport as well as commercial catch in Areas 2A and 2B.

Table 4. Estimates of 2008 exploitable biomass and CEY from the 2007 assessment.

	Area 2A	Area 2B	Area 2C	Area 3A	Area 3B	Area 4A	Area 4B	Area 4CDE	Total
Coastwide assessment [1]									
2008 exploitable biomass	4.7	25.6	32.5	144.8	74.0	21.3	20.2	37.9	361
Proportion of total	0.013	0.071	0.090	0.401	0.205	0.059	0.056	0.105	1.000
Target harvest rate	0.20	0.20	0.20	0.20	0.20	0.20	0.15	0.15	<0.20
Total CEY	0.94	5.12	6.50	28.96	14.80	4.26	3.03	5.69	69.30
Other removals [2]	0.29	0.47	2.59[3]	6.71[3]	0.53	0.75	0.33	2.01	13.68
2008 fishery CEY [2]	0.65	4.65	3.92	22.25	14.27	3.51	2.71	3.68	55.62
Area assessments [1]									
2008 exploitable biomass	4.6	37	49	169	54	11	14	52	391
Proportion of total	0.012	0.095	0.125	0.432	0.138	0.028	0.036	0.133	0.999
Target harvest rate	0.20	0.20	0.20	0.20	0.20	0.20	0.15	0.15	<0.20
Total CEY	0.92	7.40	9.80	33.80	10.80	2.20	2.10	7.80	74.82
Other removals [2]	0.29	0.47	2.59[3]	6.71[3]	0.53	0.75	0.33	2.01	13.68
2008 fishery CEY [2]	0.63	6.93	7.21	27.09	10.27	1.45	1.77	5.79	61.14

Notes:

[1] "Coastwide assessment" refers to the coastwide model fit with survey apportionment of the total biomass estimate among regulatory areas. "Area assessments" are the closed-area model fits.

[2] "Other removals" comprise legal-sized wastage, legal-sized bycatch, personal use, and in most areas sport catch. In Areas 2A and 2B sport catch is included in fishery CEY rather than in other removals.

[3] The sport catch component in these figures is the adopted guideline harvest level (GHL), which is lower than the actual 2007 catch in Area 2C and higher in Aea 3A.

Table 5. Other removals in detail. Sport catch figures for Areas 2C and 3A are actual catches not GHL levels as in Table 4.

	Area 2A	Area 2B	Area 2C	Area 3A	Area 3B	Area 4A	Area 4B	Area 4CDE	Total
Sport catch	0.52	1.75	2.55	5.05	0.01	0.05	0.00	0.00	9.93
Legal-sized bycatch	0.25	0.15	0.21	0.99	0.45	0.66	0.32	1.90	4.93
Pesonal use	0.04	0.30	0.58	0.38	0.05	0.03	0.00	0.11	1.49
Legal-sized wastage	0.00	0.02	0.02	0.05	0.02	0.01	0.01	0.00	0.13
Total	0.81	2.22	3.36	6.47	0.53	0.75	0.33	2.01	16.48
Total excl.sport catch in Areas 2A and 2B	0.29	0.47	3.36	6.47	0.53	0.75	0.33	2.01	14.21
Sublegal discard mortality (shown for information; not taken off total CEY)	0.02	0.44	0.27	0.92	0.42	0.13	0.02	0.07	2.29

Table 6. Shares of total CEY by area according to various apportionment rules.

Rule	Area 2A	Area 2B	Area 2C	Area 3A	Area 3B	Area 4A	Area 4B	Area 4CDE	Total
Survey apportionment (CPUE x bottom area)	0.013	0.071	0.090	0.401	0.205	0.059	0.056	0.105	1.000
2008 exploitable biomass from 2007 closed-area assessments	0.012	0.095	0.125	0.432	0.138	0.028	0.036	0.133	0.999
Historical recruitment from 2007 closed-area assessments (1987-1996)	0.02?	0.107	0.098	0.451	0.161	0.046	0.018	0.10?	1.001
Share of total catch (1990-2007)	0.017	0.144	0.140	0.366	0.142	0.065	0.035	0.091	1.000
Share of bottom 0-300 fm (excl. EBS shelf outside 4C)	0.066	0.160	0.082	0.256	0.154	0.094	0.078	0.111	1.001
Commercial apportionment (CPUE x bottom area)	0.035	0.113	0.055	0.401	0.162	0.088	0.066	0.080	1.000

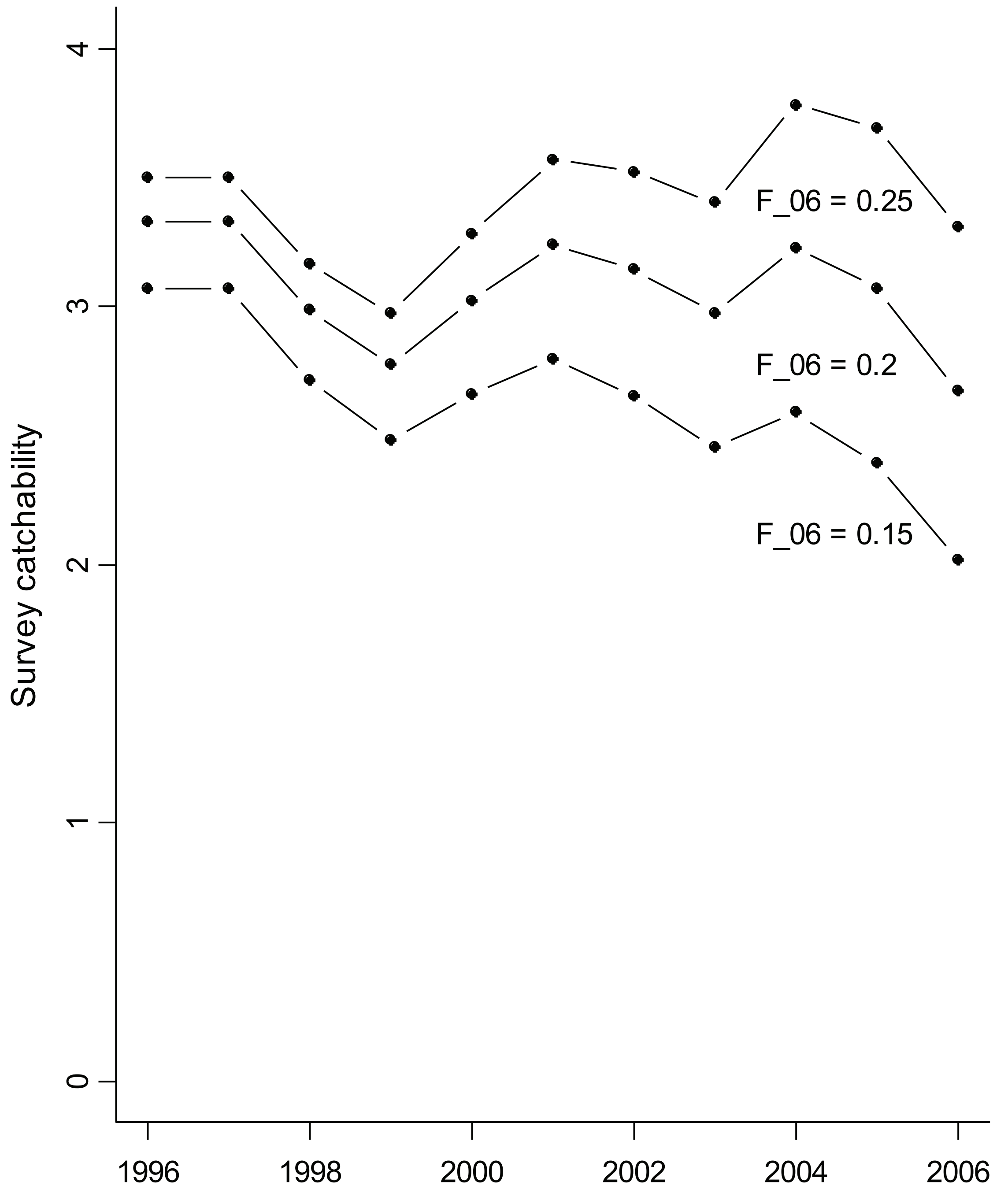

Figure 1. Calculated values of survey catchability in VPA-like fits of the model in the 2006 assessment. The labels refer to the value of the fixed terminal fishing mortality rate; e.g. "F_06 = 0.2" means that the fishing mortality rate in 2006 was set to 0.20.

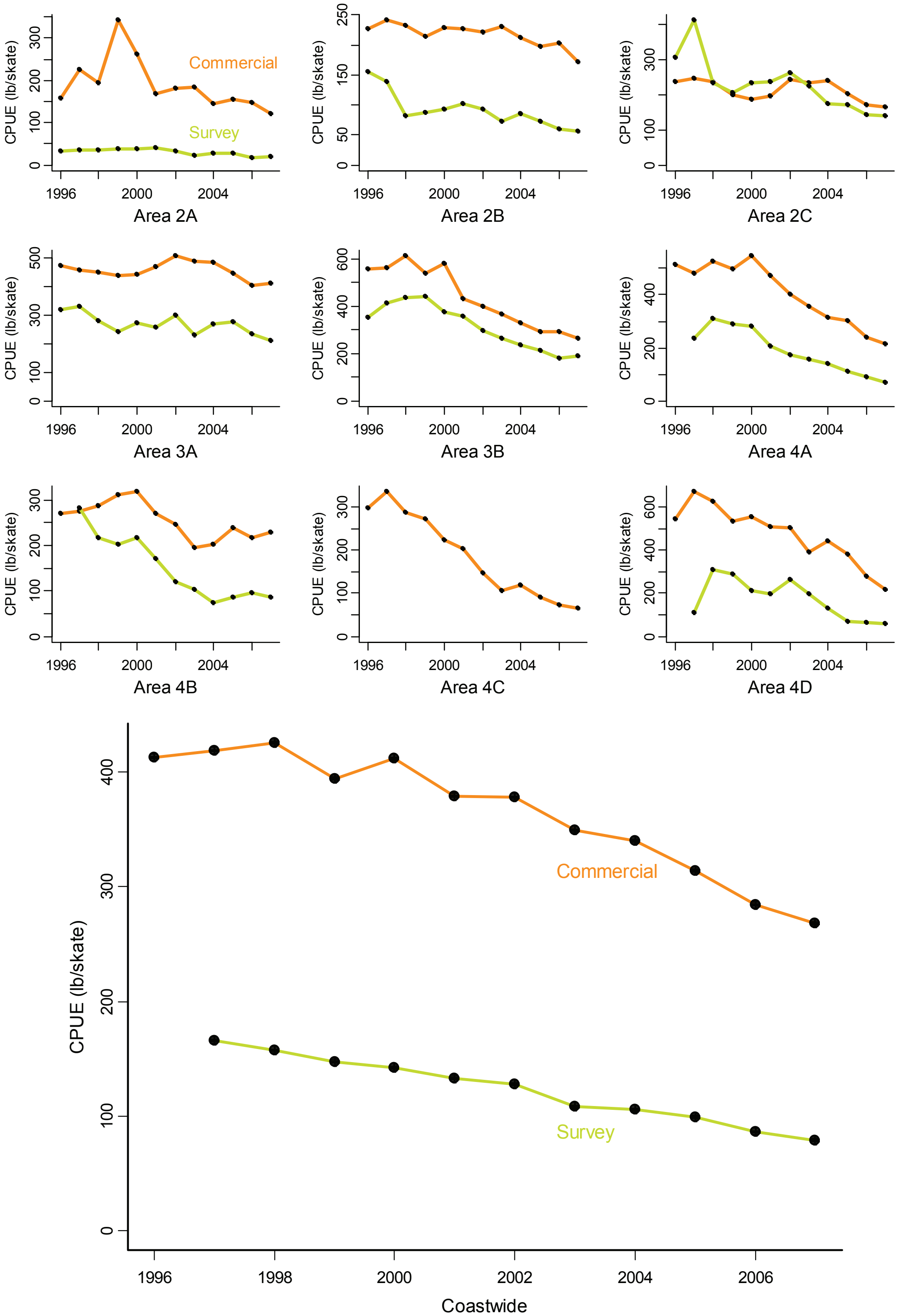

Figure 2. Commercial and survey CPUE by area (above) and coastwide (below).

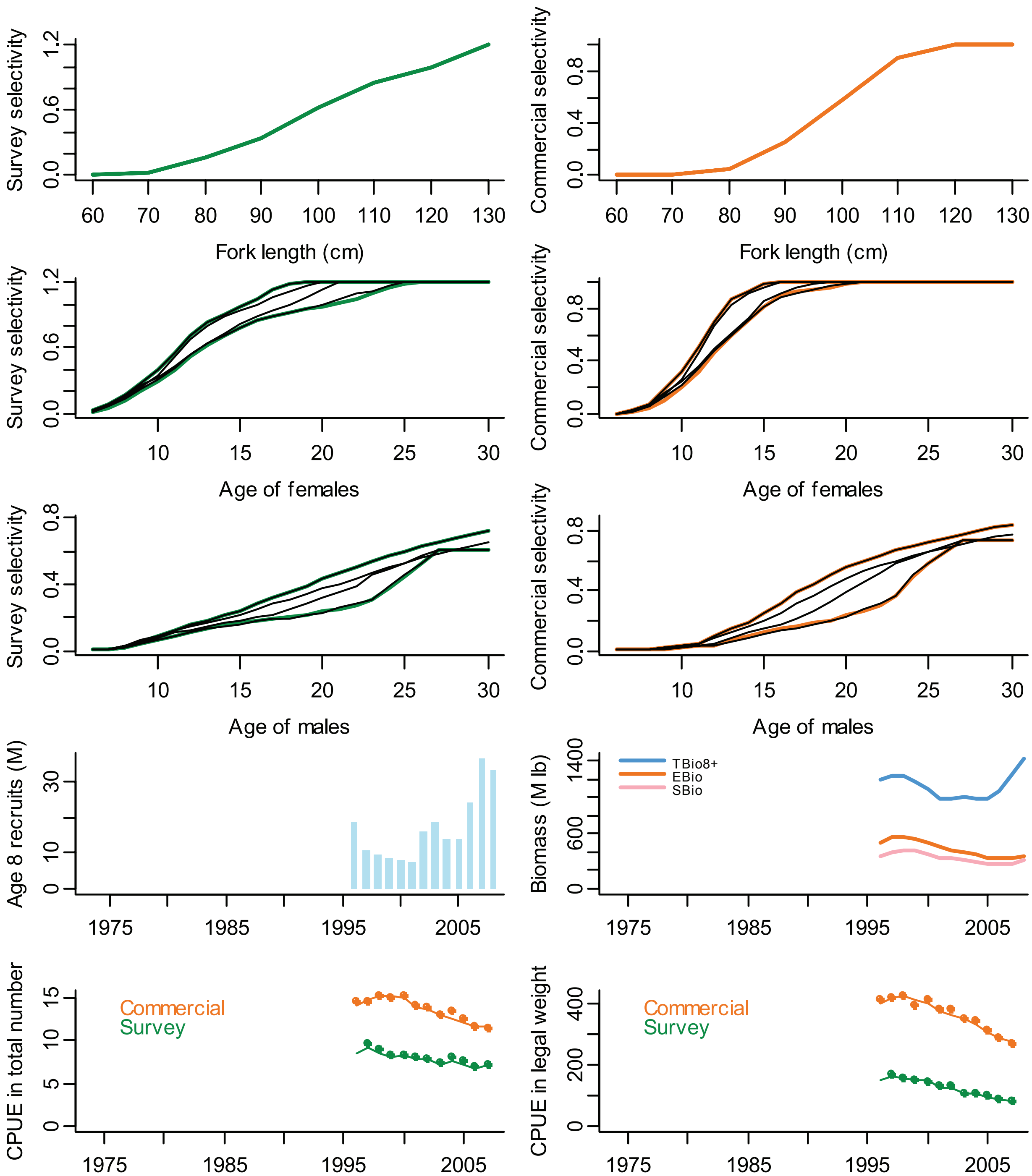

Figure 3. Features of the 2007 coastwide assessment. Age-specific selectivities are plotted for every third year plus the last.

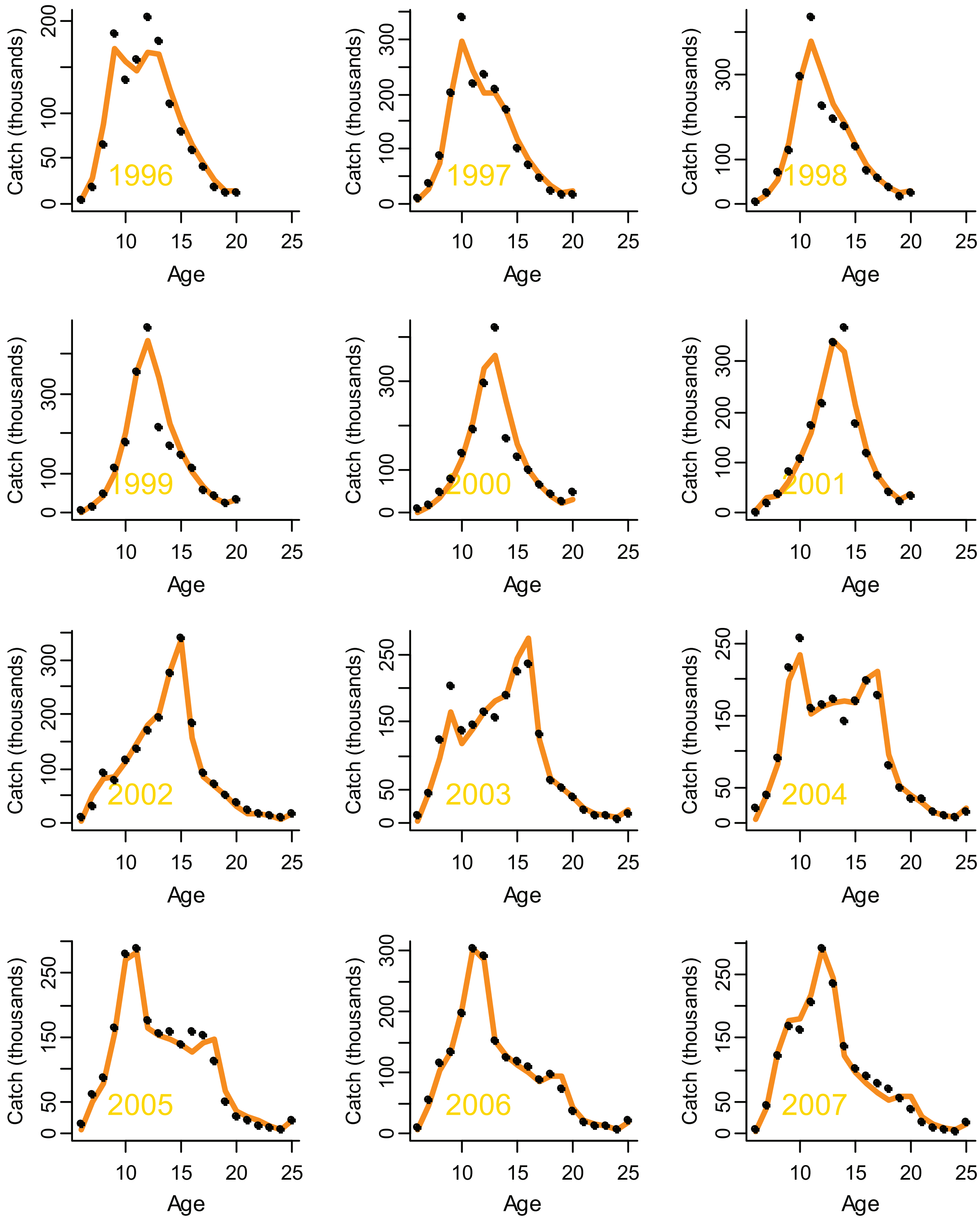

Figure 4a. Observed (points) and predicted (lines) commercial catch at age of females in the 2007 coastwide model fit.

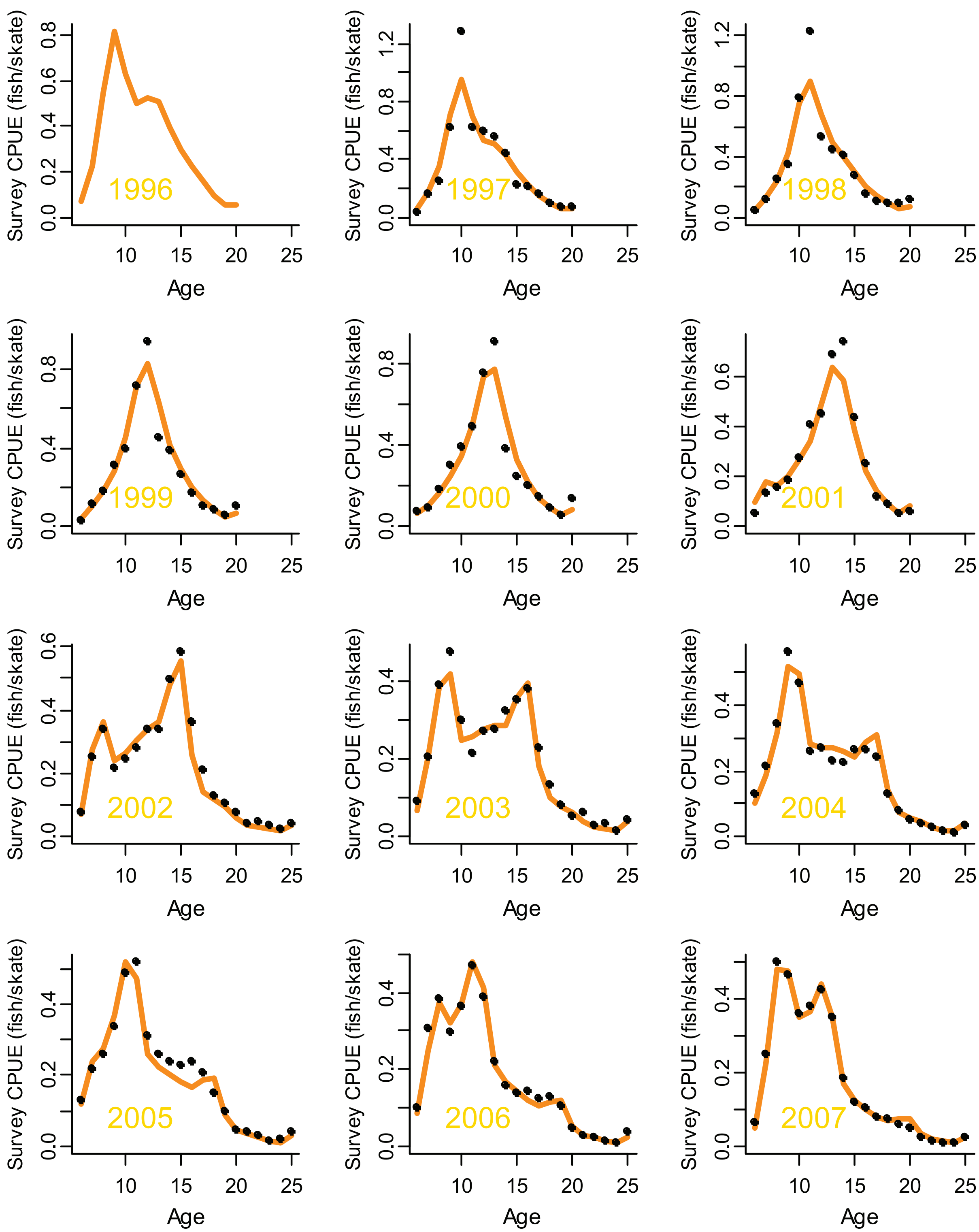

Figure 4b. Observed (points) and predicted (lines) survey CPUE at age of females in the 2007 coastwide model fit.

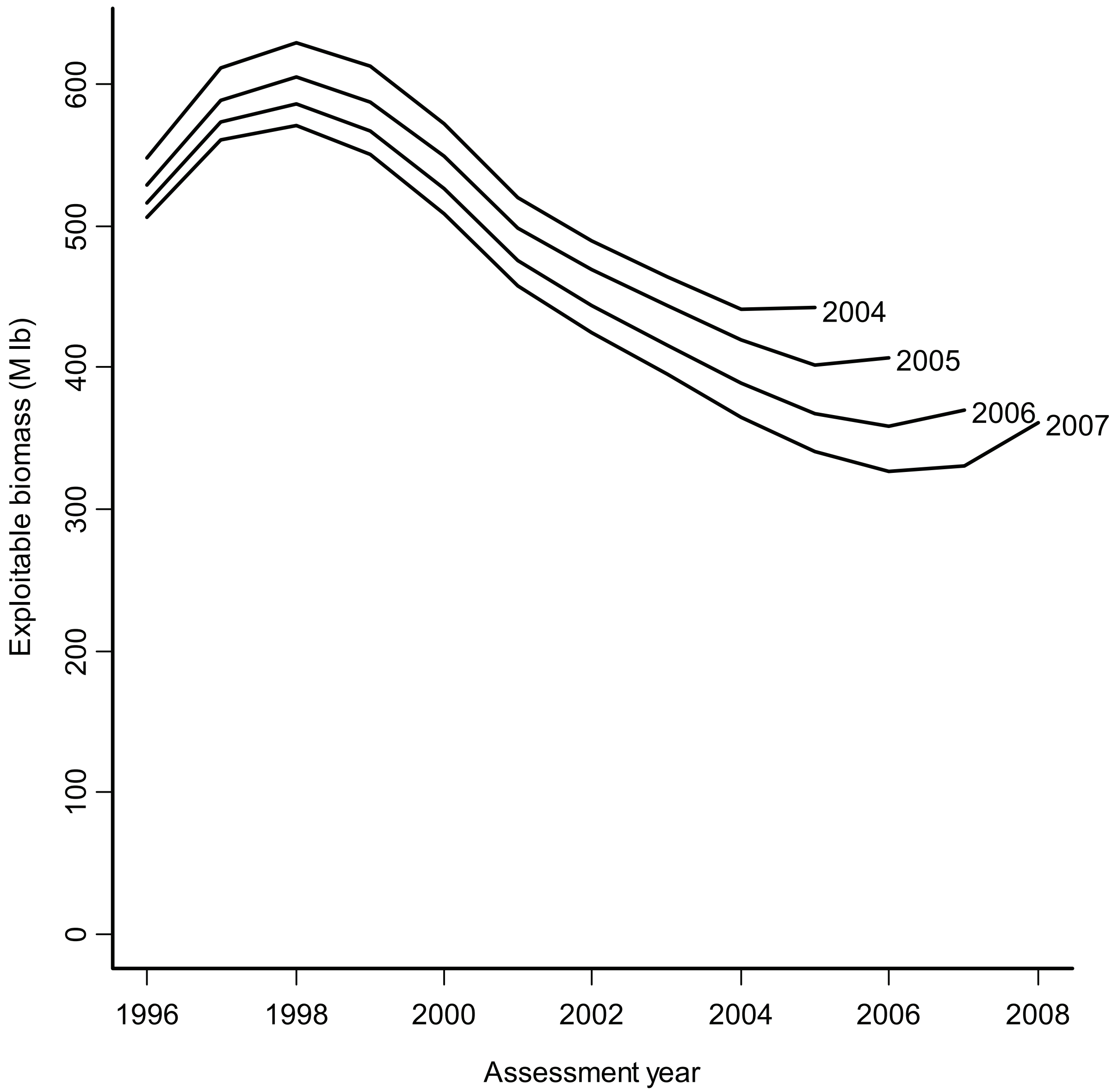

Figure. 5. Retrospective performance of the assessment. Each line is the biomass trajectory estimated by the model fitted to data from 1996 through the labeled last year.

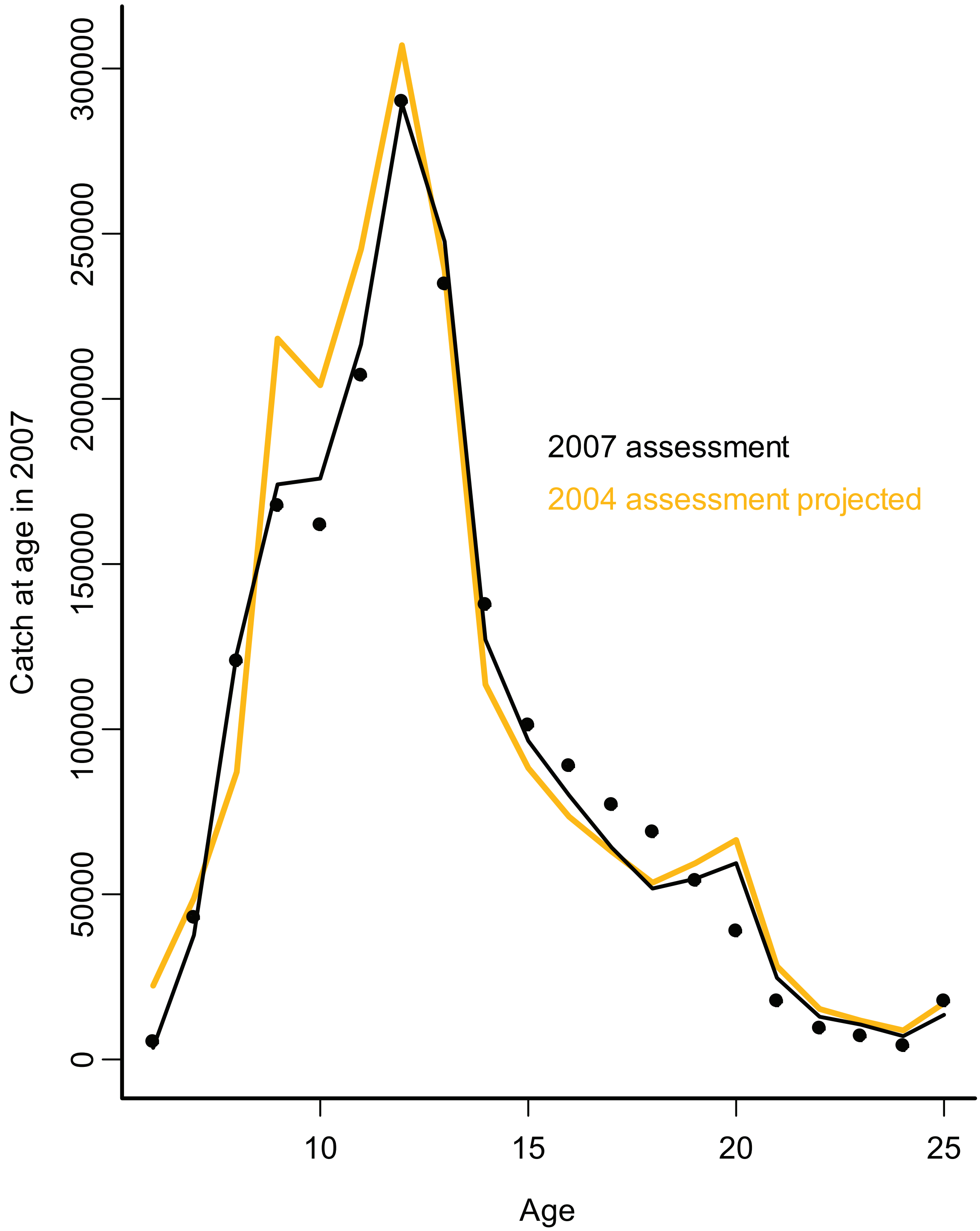

Figure 6a. Observed commercial catch at age of females in 2007 (points) and predicted catch at age from the 2007 assessment and from a blind projection of the 2004 assessment.

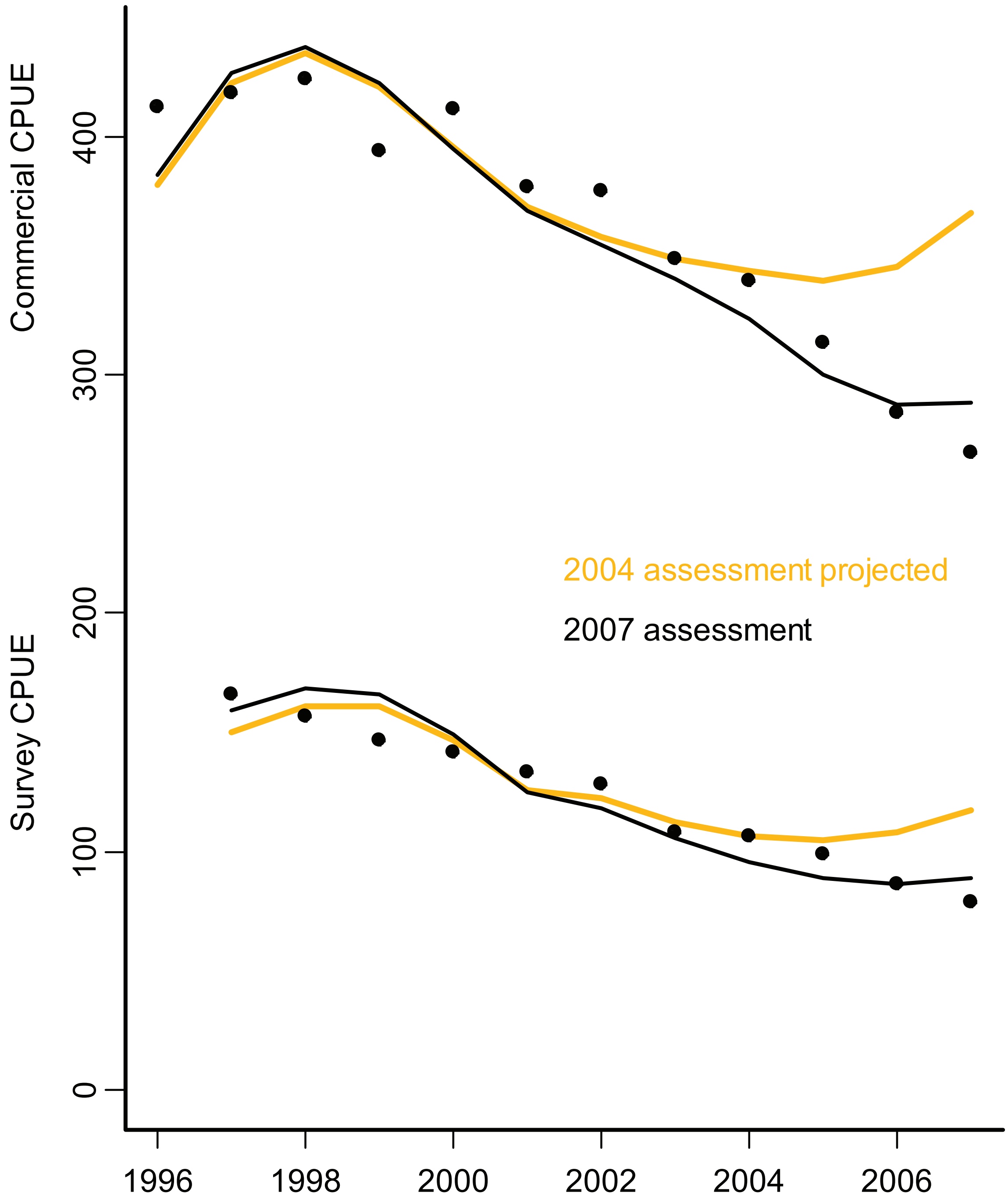

Figure 6b. Points are observed commercial (above) and survey (below) CPUE. Lines are predicted values from the 2007 assessment and a blind projection of the 2004 assessment.

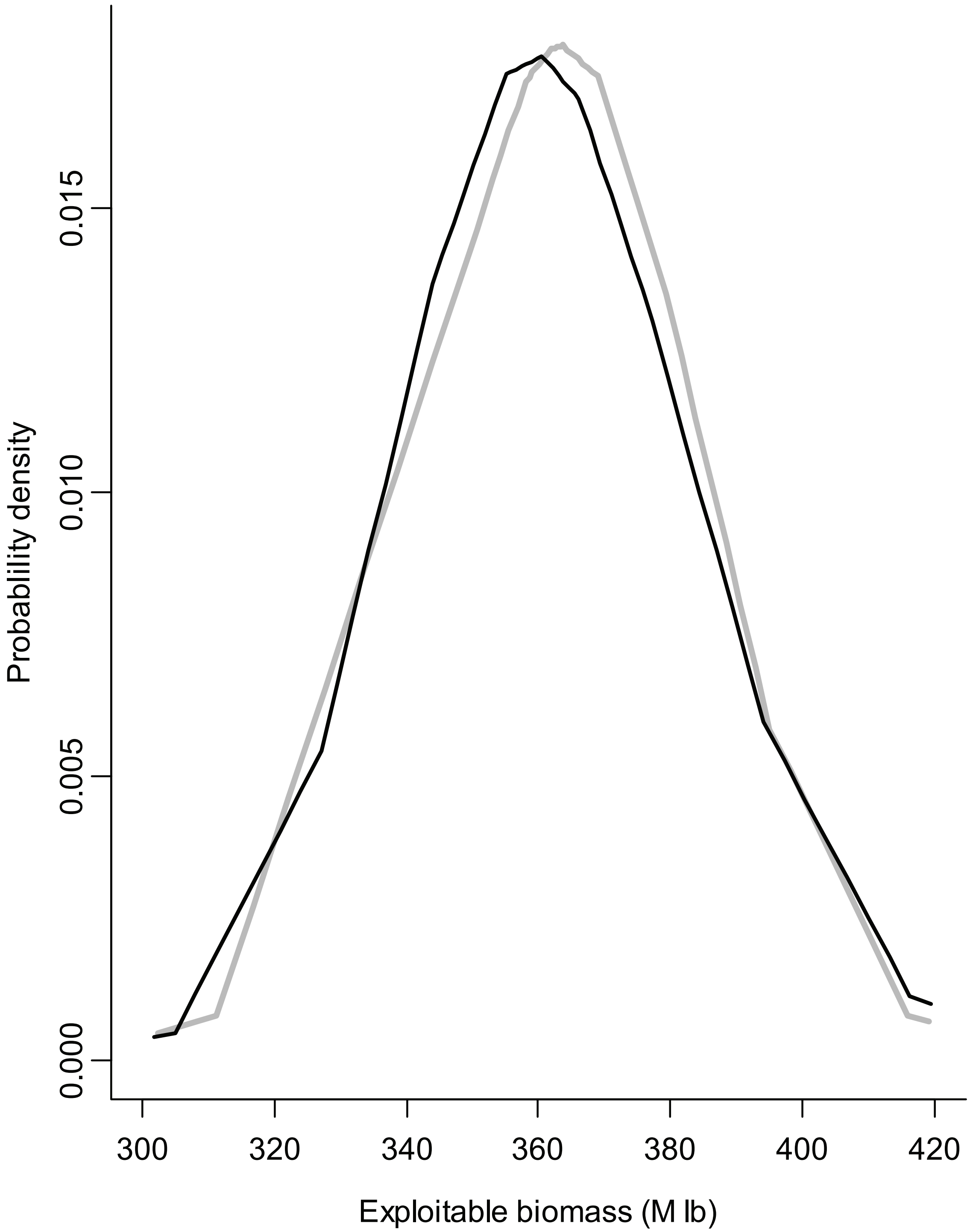

Figure 7. Estimates of uncertainty in the estimate of 2008 exploitable biomass: normal approximation based on the Hessian (gray line) and calculated likelihood profile (black line).

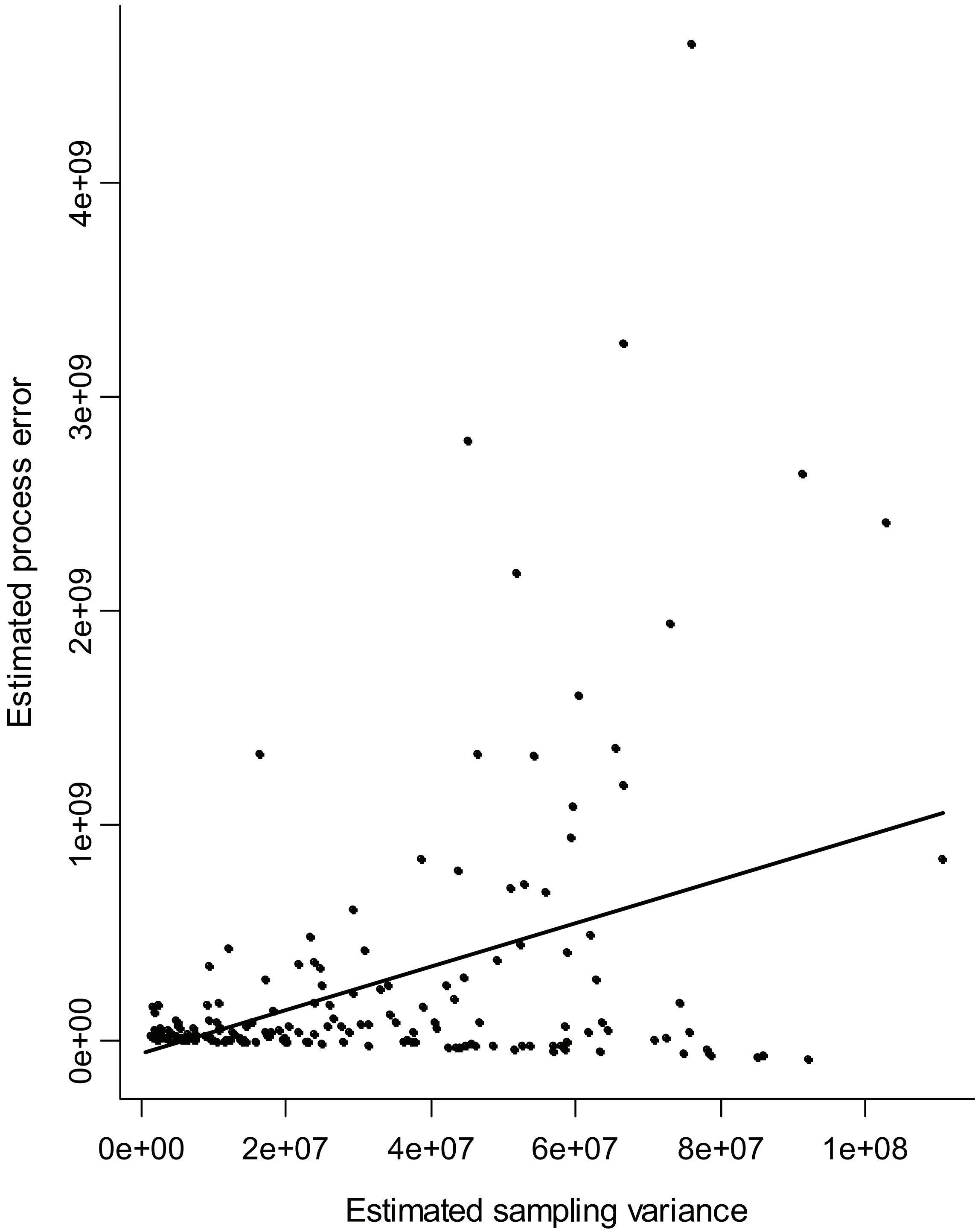

Figure 8a. Estimated process error (squared deviation minus estimated sampling variance) plotted against estimated sampling variance of female catch at age.

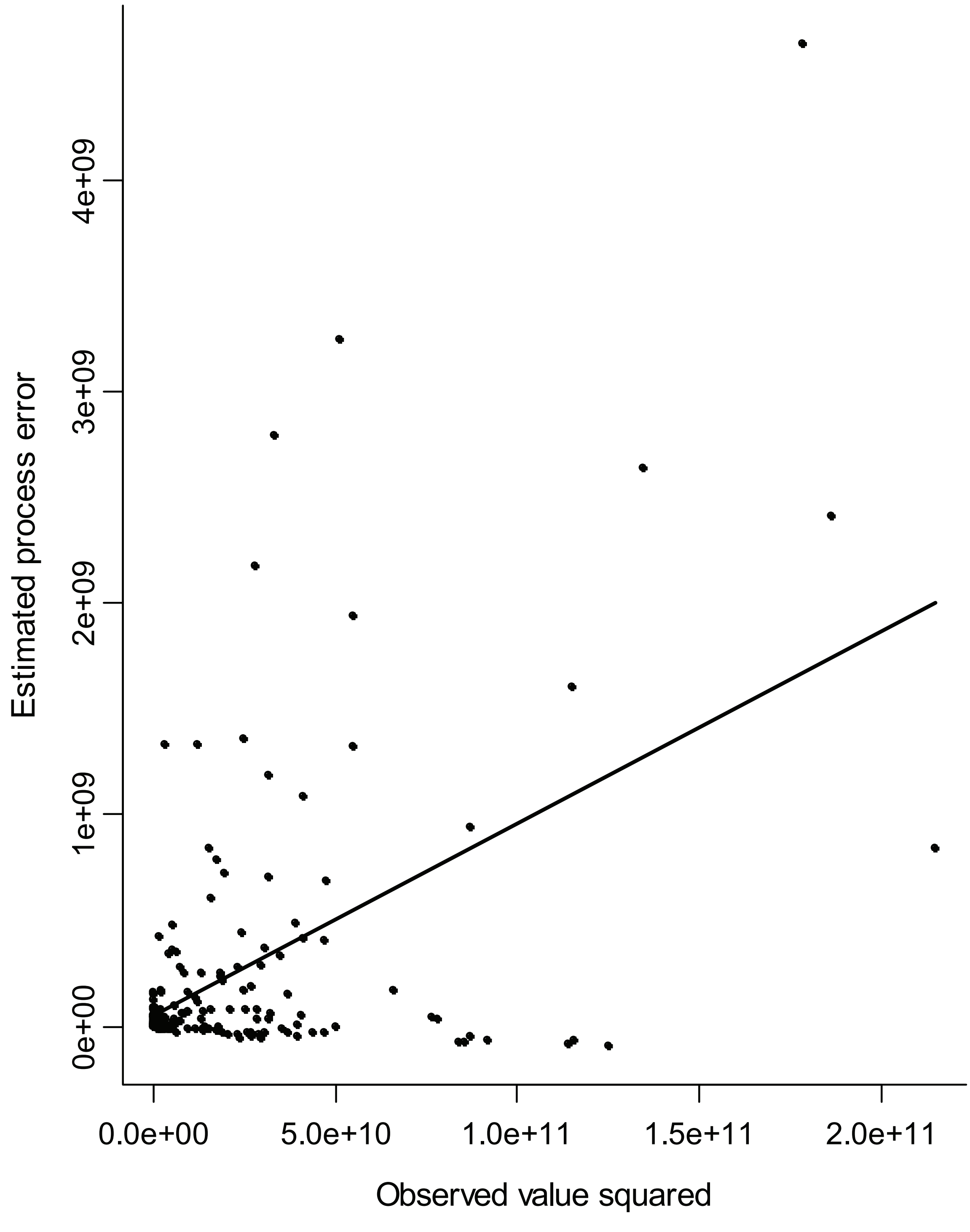

Figure 8b. Estimated process error (squared deviation minus estimated sampling variance) plotted against the square of the observed value of female catch at age.

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	0.88	12.16	10.23	25.17	15.46	---	3.56	2.72	0.95	1.66	0.31	73.11
2005	0.80	12.33	10.63	26.03	13.17		3.40	1.98	0.53	2.58	0.37	71.82
2006	0.83	12.01	10.49	25.71	10.79	---	3.33	1.59	0.49	2.37	0.37	67.98
2007	0.78	9.74	8.49	26.31	9.42	---	2.81	1.41	0.55	2.72	0.58	62.81

Table A2. 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. Total column recomputed in 2007 with new bottom area numbers.

	2A	2B	2C	3A	3B	4A	4B	4C	4D	4E	Total
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	---	357
1985	62	147	370	537	602	333	234	---	330	---	400
1986	60	120	302	522	515	265	---	427	239	---	356
1987	57	131	260	504	476	341	220	384	---	---	349
1988	134	137	281	503	655	453	224	---	201	---	392
1989	124	134	258	455	590	409	268	331	384	---	376
1990	168	175	269	353	484	434	209	288	381	---	334
1991	158	148	233	319	466	471	329	223	398	---	328
1992	115	171	230	397	440	372	278	249	412	---	336
1993	147	208	256	393	514	463	218	257	851	---	392
1994	93	215	207	353	377	463	198	167	480	---	326
1995	116	219	234	416	476	349	189	---	475	---	351
1996	159	226	238	473	556	515	269	---	---	---	413
1997	226	241	246	458	562	483	275	335	671	---	419
1998	194	232	236	451	611	525	287	287	627	---	425
1999	---	213	199	437	538	500	310	270	535	---	394
2000	263	229	186	443	577	547	318	223	556	---	412
2001	169	226	196	469	431	474	270	203	511	---	379
2002	181	222	244	507	399	402	245	148	503	---	378
2003	184	231	233	487	364	355	196	105	389	---	349
2004	145	212	240	485	328	315	202	120	444	---	340
2005	155	197	203	446	293	301	238	91	379	---	314
2006	147	202	170	403	292	241	218	72	280	---	284
2007	121	172	164	410	261	213	230	66	216	---	268

Table A3. IPHC setline survey CPUE of legal sized fish in weight (net pounds per skate).

Figures refer to entire areas. For cases where only part of an area was fished (e.g., northern 2B, western 3A), the CPUE shown is an adjusted value. *No hook corrections* are applied; J-hook values are raw J-hook catch rates. Area 4EBS is the eastern Bering Sea shelf, first surveyed in 2006. For other years, the 4EBS CPUE is a constructed value based on the NMFS trawl survey and the single 2006 setline data point.

	2A	2B	2C	3A	3B	4A	4B	4C	4D	4EBS	Total
J-hook surveys:											
1974	---	---	---	---	---	---	---	---	---	---	---
1975	---	---	---	---	---	---	---	---	---	---	---
1976	---	---	---	---	---	---	---	---	---	---	---
1977	---	13	---	58	---	---	---	---	---	---	---
1978	---	18	---	27	---	---	---	---	---	---	---
1979	---	NA	---	41	---	---	---	---	---	---	---
1980	---	25	---	76	---	---	---	---	---	---	---
1981	---	16	---	131	---	---	---	---	---	---	---
1982	---	21	114	130	---	---	---	---	---	---	---
1983	---	18	142	119	---	---	---	---	---	---	---
1984	---	25	---	176	---	---	---	---	---	---	---
C-hook surveys:											
1984	---	57	260	361	---	---	---	---	---	7	---
1985	---	42	260	378	---	---	---	---	---	8	---
1986	---	38	283	305	---	---	---	---	---	9	---
1987	---	NA	---	---	---	---	---	---	---	10	---
1988	---	NA	---	---	---	---	---	---	---	20	---
1989	---	NA	---	---	---	---	---	---	---	13	---
1990	---	NA	---	---	---	---	---	---	---	14	---
1991	---	NA	---	---	---	---	---	---	---	12	---
1992	---	NA	---	---	---	---	---	---	---	11	---
1993	---	93	---	261	---	---	---	---	---	22	---
1994	---	NA	---	254	---	---	---	---	---	17	---
1995	29	148	---	300	---	---	---	---	---	20	---
1996	---	156	306	317	352	---	---	---	---	25	---
1997	35	139	411	331	414	237	282	71	111	23	166
1998	---	82	232	281	435	310	216	---	---	30	157
1999	37	88	204	241	438	290	203	---	---	27	147
2000	---	93	233	272	373	282	216	---	215	20	142
2001	41	102	237	256	357	205	171	---	197	21	133
2002	33	92	261	299	297	174	119	---	263	13	128
2003	22	73	223	229	262	158	104	---	195	18	108
2004	27	86	173	270	236	142	73	---	132	18	106
2005	28	72	171	276	211	111	86	---	69	17	99
2006	16	59	144	232	181	88	95	---	63	18	86
2007	19	57	140	212	191	69	87	---	57	13	79

Appendix B. Evolution of IPHC assessment methods, 1982-2007

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 (Quinn et al. 1985). 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. 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 2003).

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.

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, 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.

Apart from a few minor and inconsequential corrections and alterations, the 2005 analytical assessment was the same as the 2004 assessment. The only important change in procedure was the use of the NMFS trawl survey to estimate biomass in Area 4CDE where an analytical assessment was not done.

In 2006, growing concerns about migration of legal-sized fish from western to eastern areas led the staff to doubt the validity of the closed-area assessments that had been done for many years (Clark and Hare 2007a). The staff therefore estimated coastwide abundance by fitting the model to a coastwide dataset, and estimated biomass in each area in accordance with survey estimates of relative abundance (Clark and Hare 2007b). The 2007 assessment followed the same procedure. Sublegal discard mortality in the halibut fishery was added to the removals included in the assessment; it had the effect of decreasing the present biomass estimate by less than 1%.

References

Clark, W. G. 2005. Estimates of length-specific commercial selectivity from historical marking experiments. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2004:91-102.

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 2004:199-212.

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 Hare, S. R. 2007a. Motivation and plan for a coastwide assessment. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006:83-96.

Clark, W. G., and Hare, S. R. 2007b. Assessment of the Pacific halibut stock at the end of 2006. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006:97-128.

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.

Quinn, T. J. II, Deriso, R. B., and Hoag, S. H. 1985. Methods of assessment of Pacific halibut. Int. Pacific Halibut Comm. Sci. Rep. 72.

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.