Evaluation of modeling fishery selectivity with cubic splines, with application towards BSAI Pacific ocean perch

Paul Spencer

Alaska Fisheries Science Center

Introduction

In recent years, the Science and Statistical Committee (SSC) has recommended evaluation of selectivity and for BSAI age-structured rockfish models, which include Pacific ocean perch, northern rockfish, and the blackspotted and rougheye rockfish complex. These recommendations are consistent with those of the 2013 CIE rockfish review panel, which specifically recommended evaluation of cubic splines for modeling fishery selectivity.

Changes in how selectivity are modeled could have important implications on model performance. Additionally, cubic splines are a methodology that, to my knowledge, has not been used to model fishery selectivity for Alaska groundfish. For these reasons, it is useful to introduce the general methodology for review during the September, 2014, Plan Team meeting prior to any potential implementation into the final models developed for the November, 2014, Plan Team meeting. Assuming the general methodology is accepted, the expectation is that it would be introduced in the November Plan Team documents as an alternative to the current methodology.

This document is organized as follows. First, catch and survey data are examined to gain insights on the nature of fishery selectivity (particularly with regard to whether fishery selectivity is time-varying and/or dome-shaped). Second, a description of the methodology for modeling fishery selectivity with cubic splines is presented. Third, the methodology is introduced into BSAI POP stock assessment model, and model performance is evaluated. Finally, a small set models to be evaluated in the final November SAFE document are identified. The process identified above consists of model-independent evaluation of data, and evaluation of stock assessment performance of alternative selectivity models. Assuming this general methodology is accepted, this approach is expected to be applied to BSAI northern rockfish and BSAI blackspotted/rougheye rockfish assessments for the November, 2014, meeting.

Examination of catch and survey data

Information pertaining to temporal variability in fishing selectivity

Prior to 2000, BSAI POP assessments modeled the foreign and domestic fisheries in separate time blocks. From 2001 – 2005, BSAI POP assessments used time-invariant fishery selectivity due to general similarity in depths and areas fished between the foreign and domestic time blocks. In 2006, Observer catch information was presented that indicated depths and areas fished varied not necessarily between the foreign and domestic time blocks, but rather interannually within each of these blocks. This information is updated in Figure 1. For example, during the mid-1980s when the abundance is estimated to be small, a large portion of the POP catch occurred at depths greater than 500 m, whereas in the late 1970s and since 1990 POP were captured primarily between the depths of 200 m and 300 m. Additionally, from 1999 through the early 2000s the proportion caught between 100 m and 200 m increased from ~ 20% in the early to mid 1990s to ~ 30%.

The area of capture has changed as well. POP were predominately captured in the western Aleutians during the late 1970s. From the early 1980s to the mid-1990s POP were captured predominately in the eastern Aleutian Islands, and this high concentration of catch led to area-specific TACs in the mid-1990s that redistributed the POP catch such that about 50% of the current catch is now taken in the western Aleutians (Table 11.5). Note that the extent to which the patterns of observed catch can be used as a proxy for patterns in total catch is dependent upon the degree to which the observer sampling represents the true fishery. In particular, the proportions of total POP caught that were actually sampled by observers were very low in the foreign fishery, due to low sampling ratio prior to 1984 (Megrey and Wespestad 1990).

It is theoretically possible that some of the interannual variability discussed above (i.e., depth of capture) above tracks changes in the stock distribution, which would suggest that availability to the stock is more constant than implied by examining only the Observer catch data. This issue was examined by comparing the catch-weighted mean depth in the fishery (from 1991 – 2013) to that in tows from the Aleutian Islands trawl survey from 1991 to 2010. The survey tows show a relatively constant mean depth of capture across years, whereas the fishery depth of capture show higher of interannual variation (Figure 2), with the ratio of the standard deviation of the point estimates of 2.26, 1.18, and 1.13 for the western AI, central AI, and eastern AI, respectively.

The general pattern of interannual variation in depths and areas fished suggest the potential for variation between the overlap of the population and fishing effort, and thus temporal variability in fishery selectivity. These observations led to modeling fishery selectivity with a time-varying logistic model; currently, parameters are allowed to vary between time blocks of 4 years.

Information pertaining to dome-shaped selectivity

Dome-shaped fishery selectivity indicates a decrease in the proportion of the population captured by the fishery for older-aged fish. Assuming that old fish in the survey are fully selected, a comparison of fishery and survey age compositions can reveal the potential presence of dome-shaped selectivity. The plus group for the POP assessment model is 40 years, and of interest is the relative age composition of the old fish within the plus group. Fishery and survey data were binned across years in each of five periods from 1990 to 2011, the age composition of ages 40 to 70+ are shown in Figure 3. Overall, survey age composition generally exceeds the fishery age composition, and this pattern seems to be more pronounced in the earlier time period. For example, in the 1990-1991 period, the survey age composition exceeded the fishery age composition for each age from 41 – 52 (except age 53); a similar pattern is seen in the 1997-1998 period for ages greater than 58. The pattern can be seen more clearly in the histogram of differences between survey and fishery age proportions (Figure 4); positive differences indicate that the survey proportion exceeded the fishery proportion for a given age. Of the ages with a non-zero difference, the proportion of ages with a positive difference ranged between 0.78 and 0.82 for the four earliest periods, and decreased to 0.6 in the 2009-2011 period. Overall, these data suggest that the some dome-shaped fishery selectivity has likely occurred since 1991, and that it may be diminishing in recent years.

Survey data do not exist to perform a similar analysis in the 1960s and 1970s, but the presence of dome-shaped selectivity can be inferred from the presence of old fish beginning in the AI trawl surveys and fisheries data beginning in the early 1980s. High rates of fishing mortality are estimated in order to account for very high catches in the 1960s and early 1970s. If the fishery selectivity curve during these years was asymptotic and not dome-shaped, then several ages of the 40+ age group would have subjected to high rates of fishing mortality estimated during the late 1960s and 1970s, which is generally not consistent with relatively high abundances of the 40+ group.

Theoretically, the relatively high abundance of older fish in the early 1980s could occur despite the high catches in the 1960 and 1970s if those catches represented very small fishing mortality rates, which could be achieved in the model by scaling the population to be very large such that the survey catchability coefficient was very low. In practice, estimated values for the survey catchability coefficient (q) have ranged between 1 and 2, and further adjusting the catchability coefficient to emphasize the fitting the plus group in the age composition would likely degrade the fits to the trends in population abundance. Nonetheless, this illustrates how age composition data (and modeling of selectivity) can affect estimation of the scale of population abundance.

For these reasons, the estimated abundance of the plus group in the 2012 model with logistic selectivity is near zero in both the fishery and survey data (Figures 5 and 6). This mismatch was noted in the 2012 assessment when age range for the plus group was increased from 25+ to 40+, and further noted by the 2013 CIE review panel, who recommended cubic splines as a parsimonious method for modeling selectivity over both time and age.

Methodology for modeling fishery selectivity with cubic splines

A mathematical definition of a spline is a smooth function that is used for either interpolating between fixed points (referred to as “knots” or “nodes”) or smoothing a dataset. Splines are of interest when the underlying process for which the spline represents is a smooth, nonlinear function. Splines take their name from physical splines used in shipbuilding or mechanical drafting (before the age of computers), where the interest was bending a piece of material (such as strip of wood, or an elastic ruler) between fixed points such as to produce a surface that varies smoothly with a minimal degree of bending. Similarly, mathematical splines are constructed from separate piecewise functions that are joined at the knots, and smoothness is ensured by requiring that at each knot, the two functions joined have equal function values, first derivatives, and second derivatives. These conditions can only be met by using polynomial splines of order 3 or higher, and cubic splines are often used because they limit unnecessary bending between the knots. Splines are implemented in non-parametric modeling such as generalized additive models, and been examined in ecological modeling as an approach for modeling time-varying parameters (Thorson et al. 2013). In stock assessment modeling, non-parameteric selectivity curves (a category that includes splines) performed well in an evaluation of various approaches for modeling fishery selectivity (Thorson and Taylor 2013).

The splines used in the paper were computed with the “bicubic_spline” function included in recent versions of AD Modelbuilder. This function was contributed by Dr. Steve Martell, and originally developed for use in iSCAM (Integrated Statistical Catch Age Model) from code provided in Press et al. (1992). Briefly, the method allows the user to specify a number of age and year nodes that form a grid in the year-age matrix of time-varying selectivity (with equal grid spacing), and values at these nodes are the log-scale fishery selectivity and estimated as parameters. Fishery selectivity at ages and years between the nodes are interpolated with a bicubic spline. The smoothness of the surface is controlled by the number of nodes, and also by a series of penalties estimated within the model. The original iSCAM model included penalties for: 1) smoothness across the ages (modeled with the second difference); 2) the slope of the rate of decline when selectivity decreases with age (modeled with the first difference); and 3) the smoothness across years (modeled with the second difference). In addition to these penalties, I added a penalty on the interannual variability across years (modeled with the first difference) to address situations in which the selectivity across years was relatively smooth but also non-constant (as would occur with a trend).

Application of the BSAI POP model

The following exploratory models were run, and compared to the base case of the final 2012 assessment model (referred to as Model 0).

Model 1: Same as the 2012 model, except for replacing the time-variant logistic fishery selectivity with the bicubic spline fishery selectivity. Four year nodes and five age nodes were used, for a total of 20 nominal parameters.

Model 2: Same as Model 1, but with the survey biomass estimates and age compositions for the 1980, 1983, and 1986 AI cooperative surveys removed. The removal of these surveys has been suggested by the SSC because of different sampling protocols than those applied in the AFSC trawl survey from 1991 forward.

Model 3: Same as Model 2, but the age and length composition sample size within each data category (i.e., survey age composition, fishery age composition, etc.) are rescaled to a maximum of 200 (for the AI survey length composition, the sample size was set to the average of survey age composition sample sizes). The current model uses the square root of read otoliths as the weights for the fishery and age compositions, and the square root of fish lengthed for the fishery and survey length compositions. Because many fish are lengthed in the fishery and survey, this procedure puts stronger weight on the length composition data. Rescaling within each catch category would give the data categories more equal weights but still allow interannual variation in input sample sizes, which might improve the fits to the age composition data.

The models are compared with respect to how well they fit the age and length composition data. The standard deviation of normalized residuals (SDNR) measures the magnitude of the residuals for the age and length composition fit, relative to the standard deviation assumed for the multinomial distribution (a function of the sample size). Note that this could result in situations in which a particular model resulted in reduced raw residuals an increased SDNR, which would occur if the standard deviation assumed for the multinomial distribution decreased relatively more than the residual. Thus, it is also of interest to examine the standard deviation of the raw residuals (SDRR), as this would provide a simple measure of close the model predictions are to the data points apart from any assumptions on multinomial standard deviation. The estimates of survey catchability and estimated 2013 total biomass are also examined to see how modeling fishery selectivity affects the estimated abundance of the stock.

The time-varying parameters for the ascending portion of the logistic curve in model 0 (Figure 7) allows good fits to the younger and middle-aged fish, resulting in a better overall fit than would be expected given the input sample sizes and likely indicating some overfitting of the fishery data (Table 1). The relatively strong presence of fish between the ages of 33 and 39 in the 1980 survey and 1981 fishery is generally not observed in later years, resulting in a poor fit to these year classes.

The fishery selectivity curve fit with the bicubic spline for Model 1 is shown in Figure 6, and results in improved fits to the plus group for the fishery age composition (Figure 5). The fishery data are still overfit (based on a comparison of effective sample sizes to input samples sizes), but to a lesser extent than in the 2012 final model, and Model 0 and 1 show both show similar (and low) values SDNR values. The fit to the plus group in the survey age composition data is improved as well (Figure 6), and the use of the bicubic spline reduced the SDNR for the survey age compositon data from 1.28 to 0.83. The fits to the fishery length composition data are similar, with the SDNR being reduced slightly from 1.0 in Model 0 to 0.94 in Model 1 (Figure 8).

The removal of the 1980s cooperative Aleutian Islands survey data further improves the fit of the AI survey age composition data (Figure 9), as the SDNR for this data component is reduced to 0.63 in Model 2. Both the SDNR and the SDRR are increased slightly in the fishery age composition data, but the SDNR for this category is the lowest among all the age and length composition data groups, and the high ratio of effective sample size to input sample size indicates that some overfitting may still be occurring. The SDNR and SDRR for the fishery length compositions are nearly equal between Model 2 and Model 1 (Table 1).

The overfitting of the fishery age composition is alleviated by recscaling the sample sizes within each and length composition category (Model 3), which substantially increases the input sample sizes for the age composition data and decreases the input sample size for the length composition data (Table 1). The SDNR of the survey and fishery age composition data is increased from ~0.6 to ~ 0.9 for each data type, which reflects the change in the input sample size. The SDRR are reduced for these data types, indicating suitable fits to the data (Figures 9 and 10). Because the SDNR incorporates the assumed variance, it should be expected that these values should be close to 1 for each data type, which is best achieved in Model 3 (Table 1, Figure 11). Model 3 also achieves the lowest values of SDRR for the survey and fishery age composition.

The estimated survey catchability is increased for each of the models in which the bicubic selectivity curve is used, which lowers the estimates of total biomass. As mentioned above, allowing survey catchability to be estimated in the model allows the overall scale of the population abundance to be adjusted to best fit the trade-off between fitting the age composition data and the trends in population abundance, and also explains why estimates of survey catchability may change when new age and length composition data are introduced to an existing model. The variability in the estimates of q between model formulations, or within an existing model formulation but with updated data, calls into question the utility of attempting to estimate this parameter within the model when the input data may contain inconsistencies between age composition data and population trend data. Temporal stability in the estimates of population would be improved by stronger prior information on survey catchability, and developing this prior information is a current research priority.

Conclusions

The bicubic spline provides an improved method for modeling fishery selectivity than the current method of a logistic curve that is varies between time blocks, and when combined with rescaling the input samples sizes, produces the lowest SDRR and values of SDNR that are relatively constant between the age and length composition data categories. Additionally, the fit to the plus group in the survey and fishery age composition data is improved. The estimated strong dome-shaped selectivity in the early years of the fishery, and less strongly dome-shaped selectivity in recent years, is consistent with examination of the catch and survey data.

For the final models for the November SAFE, the existing assessment model will be compared to that with a bicubic selectivity curve using model selection criteria. In the runs here, the improved performance with the bicubic spline was achieved with 20 fishery selectivity parameters as opposed to 30 for the existing time-varying logistic curve. However, these parameters are not truly independent, as both approaches use penalties to smooth the selectivity curves. Model selection criteria such as deviance information criteria can be used to account for the "effective" number of parameters in these situations. However, model selection criteria may not necessarily reveal a satisfactory residual pattern, which will also be examined.

A time-varying double logistic curve could also be potentially investigated, although from the results seen here it not expected that the performance would be an improvement upon the non-parametric spline. Using a functional form such as the double logistic would also force the user to define the time blocks for the temporal changes, which could result in model misspecification (Martell and Stewart 2014).

Assuming that the general methodology of the bicubic spline is accepted, the general approach illustrated here for BSAI POP is expected to be followed for the BSAI northern and blackspotted/rougheye assessments, as this would address recent SSC requests.

References

Martell, S. and I. Stewart. 2014. Towards defining good practices for modeling time-varying selectivity. Fish. Res. 158:84-95.

Megrey, B.A. and V.G. Wespestad. 1990. Alaskan groundfish resources: 10 years of management under the Magnuson Fishery Conservation and Management Act. North American Journal of Fisheries Management 10:125-143.

Thorson, J.T., S. Zhou, A.E. Punt, and A.D.M. Smith. 2013. A stepwise-selected spline approximation to time-varying parameters, with application to occupancy modeling. Methods in Ecology and Evolution 4:123-132.

Thorson, J.T. and I.G. Taylor. 2014. A comparison of parametric, semi-parametric, and non-parametric approaches to selectivity in age-structured assessment models. Fish. Res. 158:74-83.

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. 1992. Numerical recipies in C: the art of scientific computing, 2nd ed. Cambridge University Press. 994 p.

Table 1. Estimated quantities and performance metrics for the evaluated models.

	Model 0	Model 1	Model 2	Model 3
Nominal fishery selectivity parameters	30	20	20	20
Likelihood				
Component				
Fishery age composition	39.56	41.34	45.94	217.45
Fishery length composition	290.41	299.79	293.09	230.21
AI survey age composition	105.82	71.47	34.45	101.61
AI survey length composition	7.15	8.62	8.05	8.06
Average Effective				
Sample Size				
Fishery age composition	227.95	162.89	160.60	208.23
Fishery length composition	227.90	174.55	180.05	148.13
AI survey age composition	102.76	114.21	121.59	195.30
AI survey length composition	174.74	132.11	141.84	142.42
Average Sample Sizes				
Fishery age composition	23.13	23.13	23.13	159.60
Fishery length composition	147.70	147.70	147.70	94.41
AI survey age composition	32.36	32.36	29.13	166.50
AI survey length composition	177.00	177.00	177.00	166.00
Standard Deviation of Normalized Residuals				
Fishery age composition	0.44	0.45	0.61	0.93
Fishery length composition	1.00	0.94	0.94	0.81
AI survey age composition	1.28	0.83	0.63	0.87
AI survey length composition	0.76	0.82	0.79	0.78
AI trawl survey	1.28	1.10	1.00	1.24
CPUE index	2.63	2.61	2.57	2.56
	1.31	1.30	1.28	1.28
Standard Deviation of Raw Residuals				
Fishery age composition	0.014	0.016	0.017	0.013
Fishery length composition	0.024	0.023	0.023	0.024
AI survey age composition	0.021	0.019	0.015	0.012
AI survey length composition	0.014	0.016	0.016	0.016
Survey catchability	1.10	1.52	1.74	1.63
2013 biomass estimate (t)	661,440	507,087	490,613	458,454

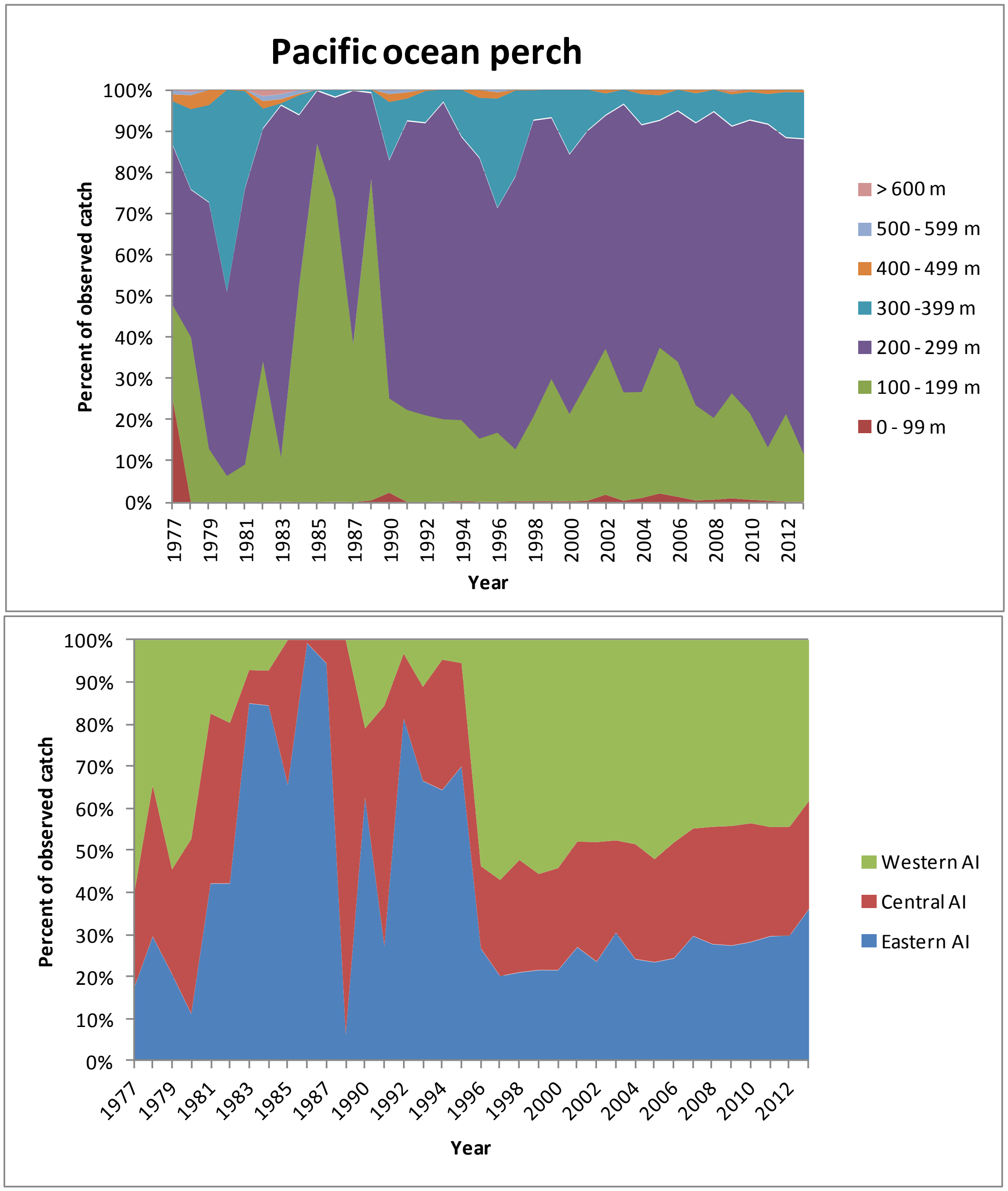

Figure 1. Temporal variability in the depth and area of capture of Pacific ocean perch in the Aleutian Islands, based upon fishery observer data.

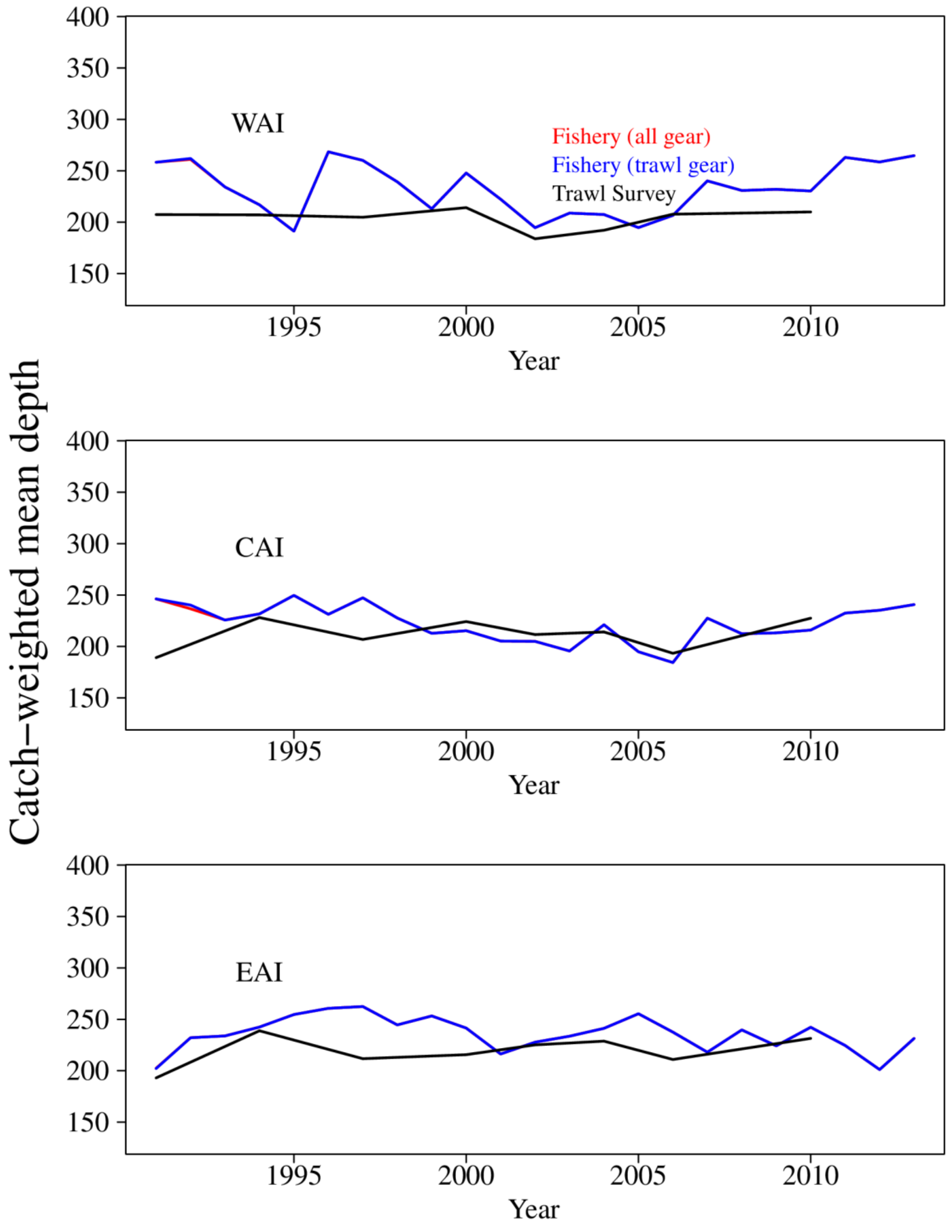

Figure 2. Mean depth of catch in the Aleutian Islands fishery and survey tows, weighted by catch (in numbers).

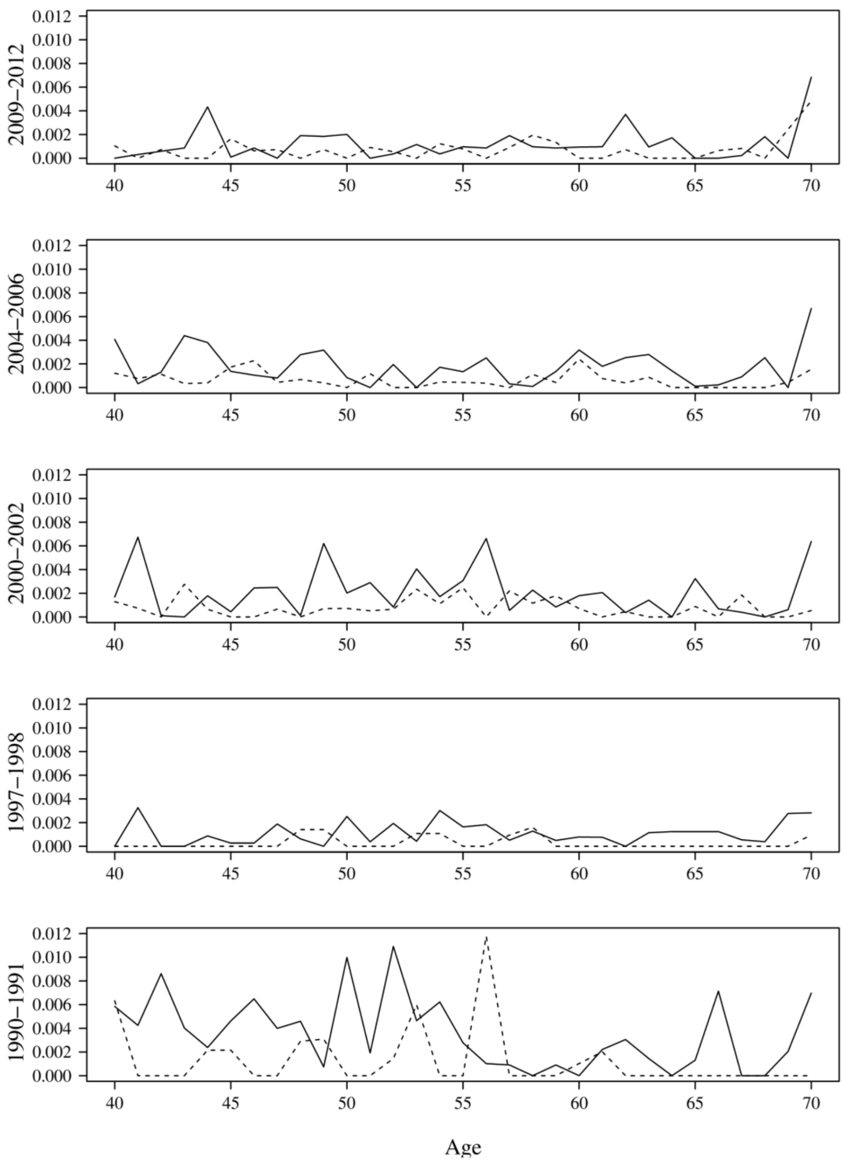

Figure 3. Age compostions in the Aleutian Islands survey (solid line) and fishery (dashed line) for ages 40 to 70+ for five time periods.

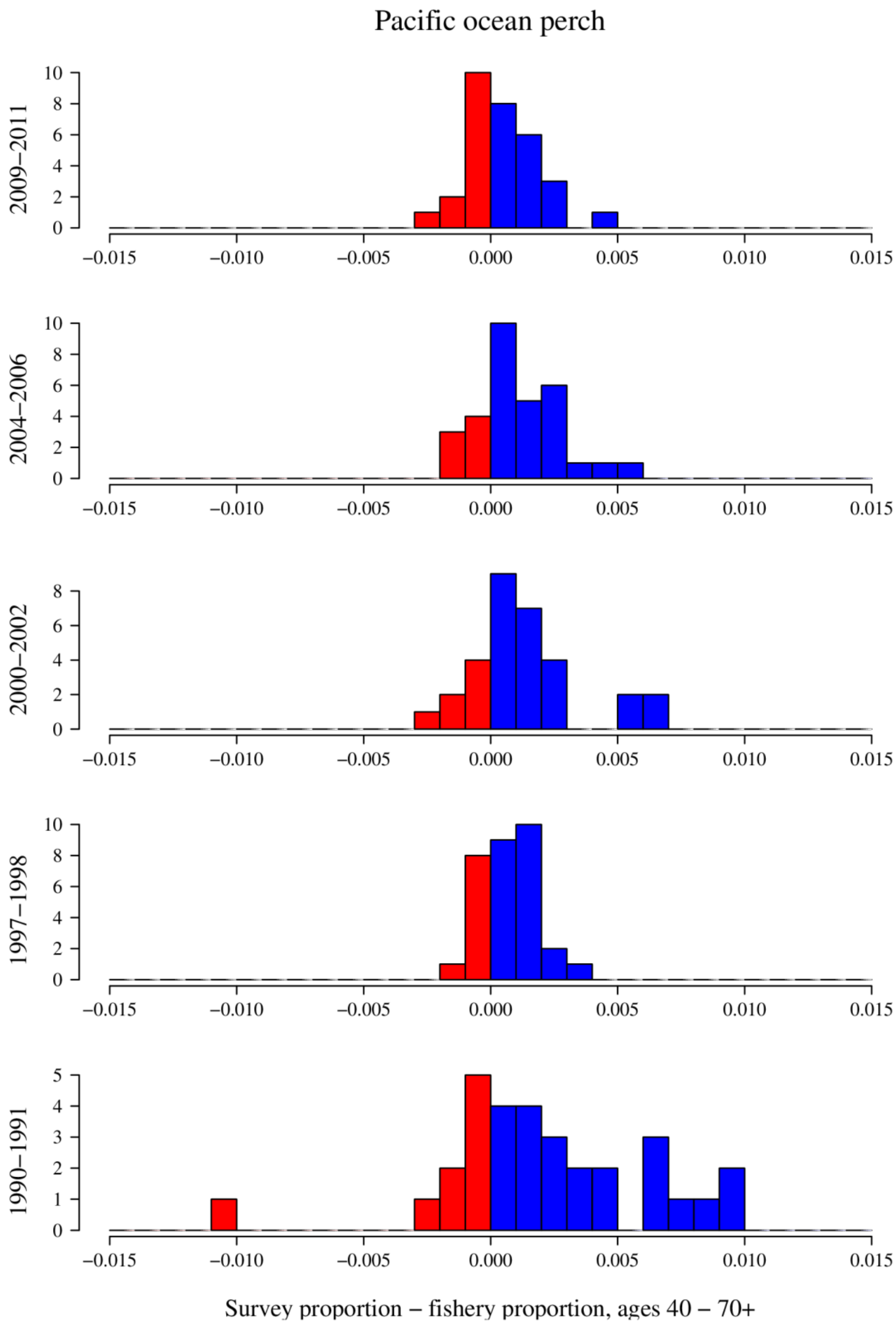

Figure 4. Histograms of the difference (survey proportion – fishery proportion) for ages 40 to 70+ for five time periods.

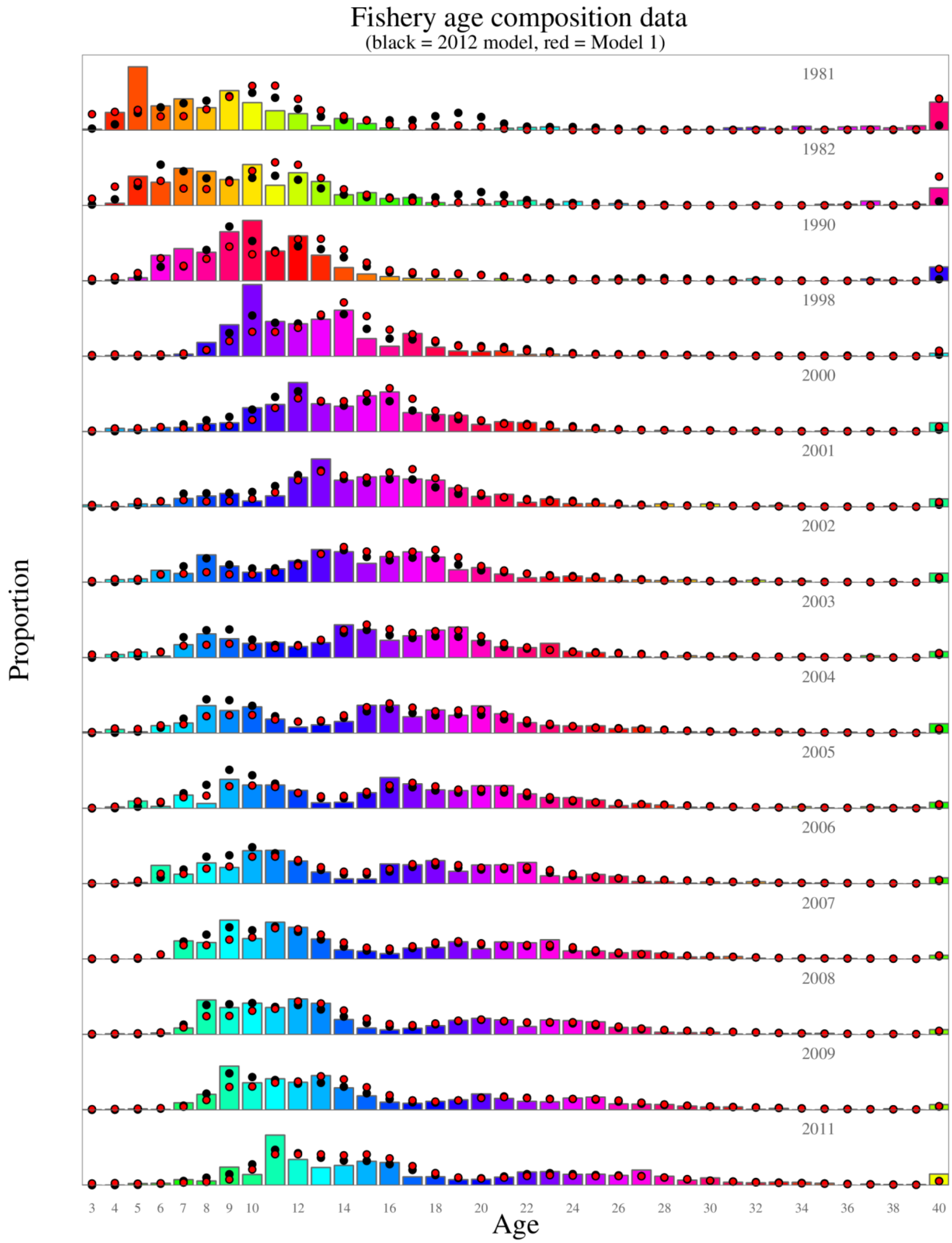

Figure 5. Fits to fishery age composition data from the 2012 assessment model (Model 0), and with the bicubic spline (Model 1).

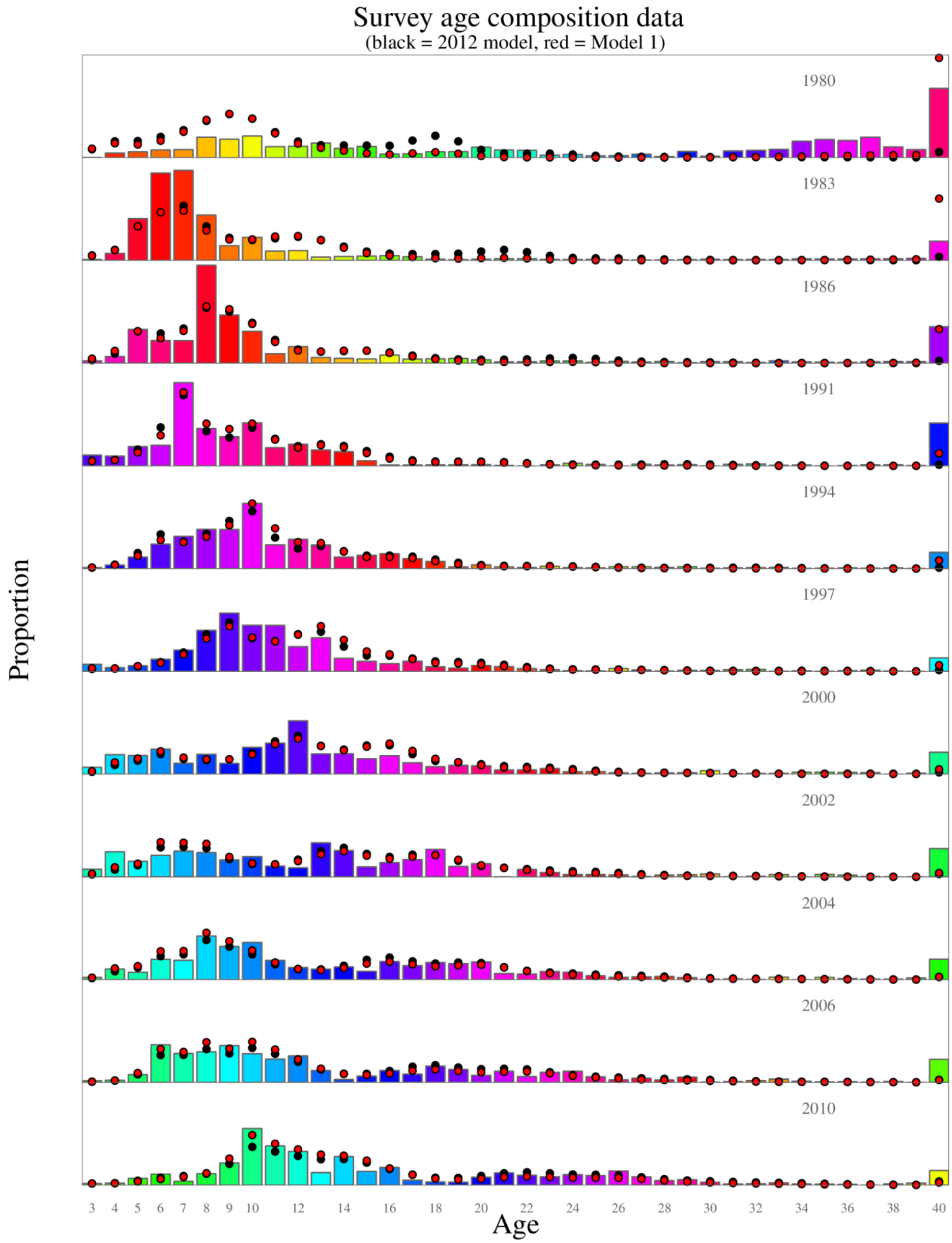

Figure 6. Fits to survey age composition data from the 2012 assessment model (Model 0), and with the bicubic spline (Model 1).

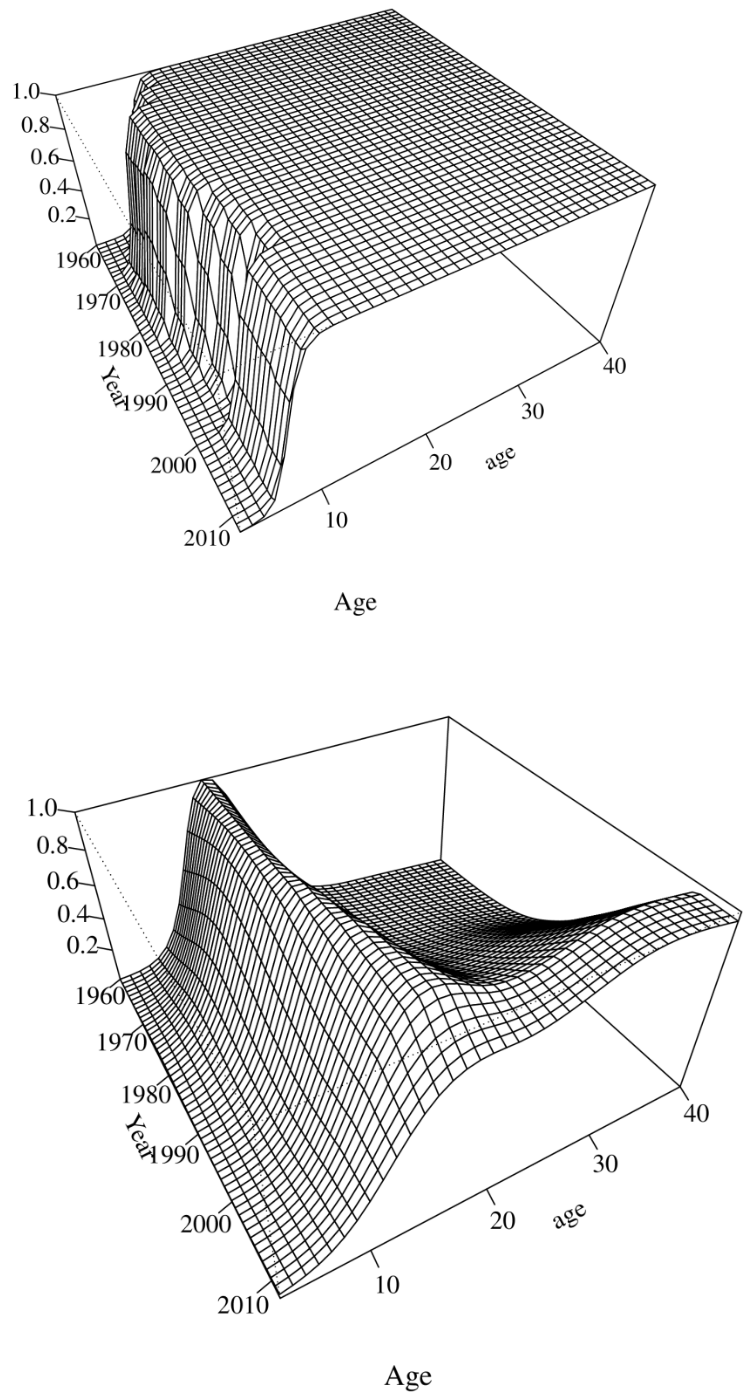

Figure 7. Estimated fishery selective from the 2012 model (model 0; top) and with the bicubic spline (Model 1, bottom).

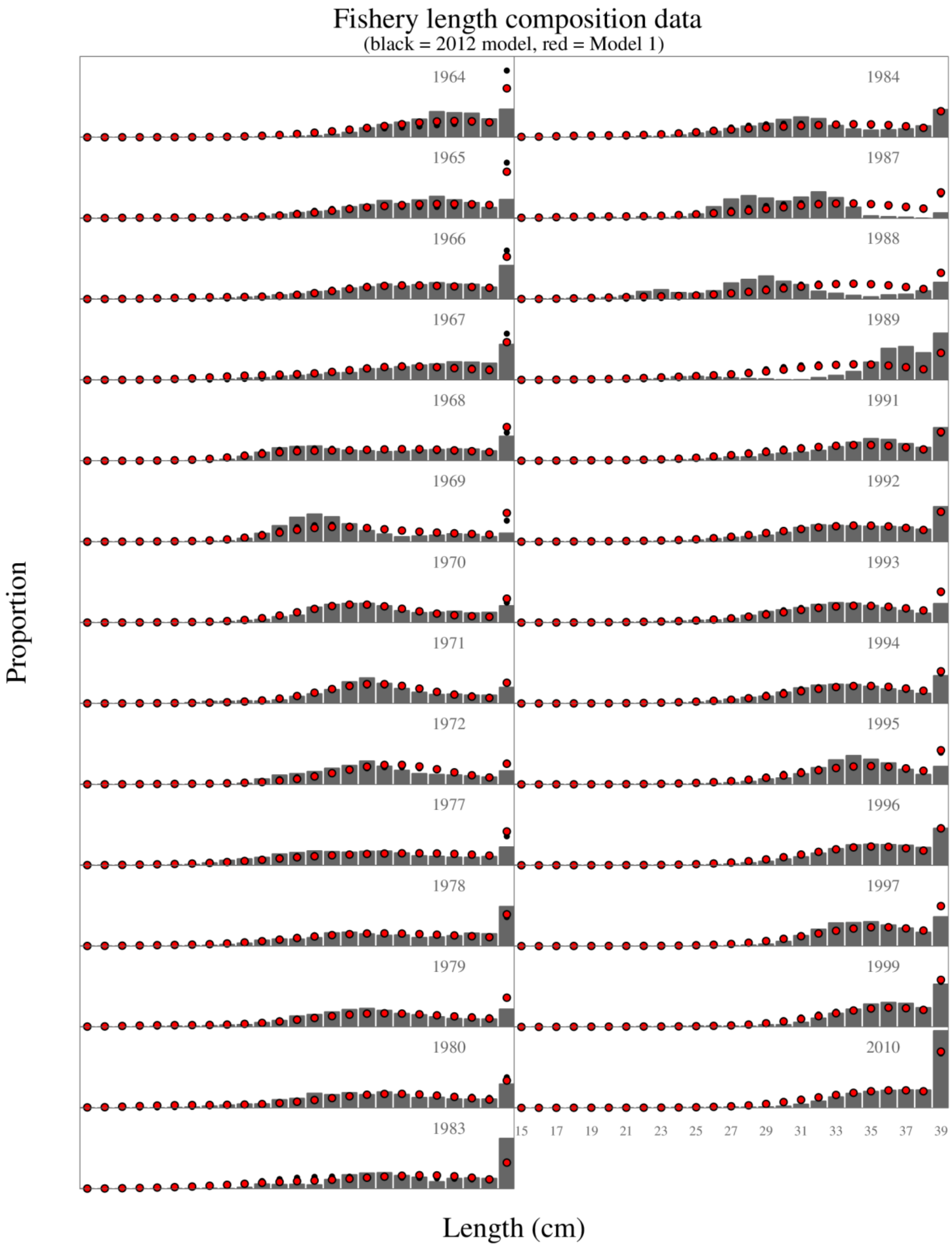

Figure 8. Fits to fishery length composition data from the 2012 assessment model (Model 0), and with the bicubic spline (Model 1).

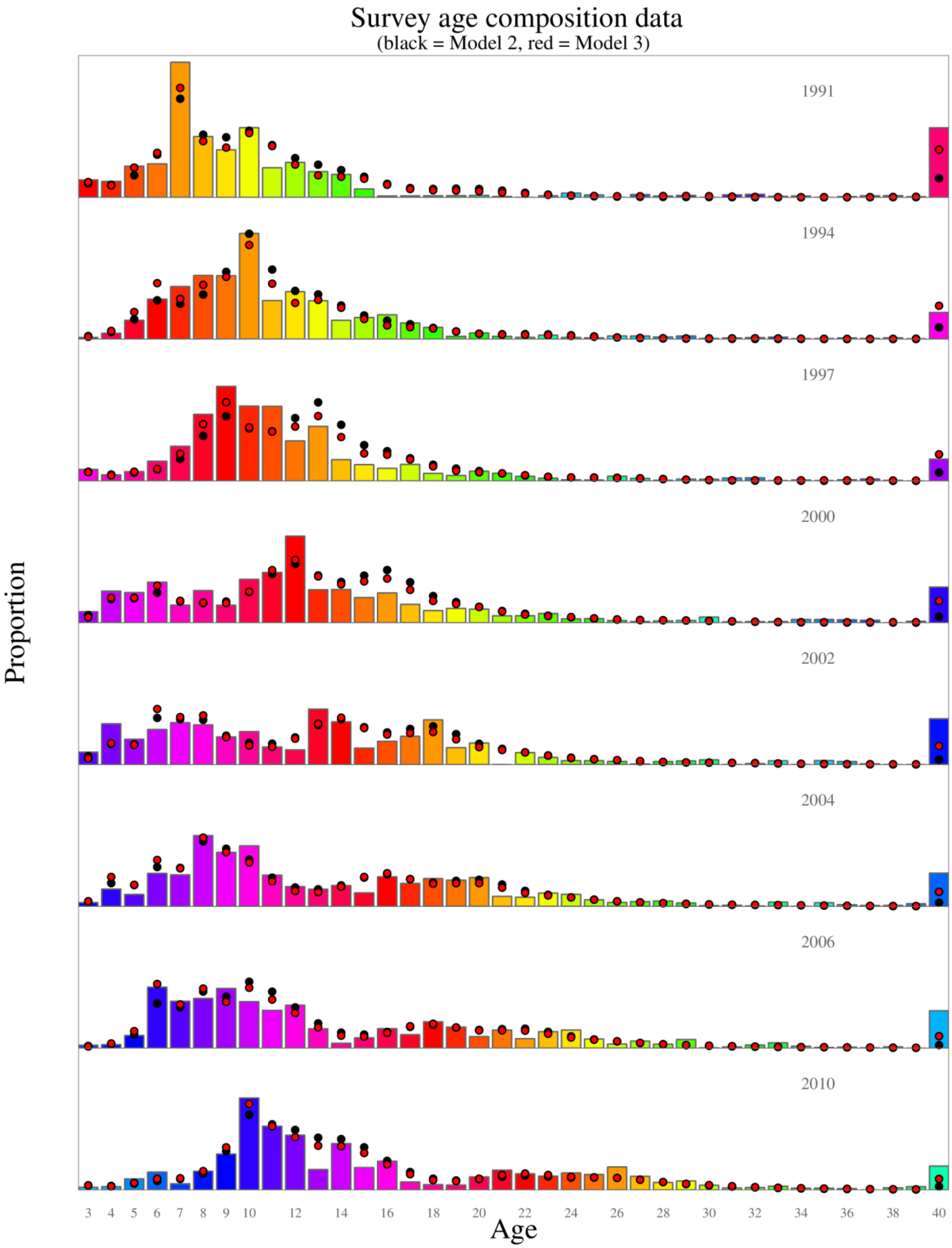

Figure 9. Fits to survey age composition data for Models 2 and 3.

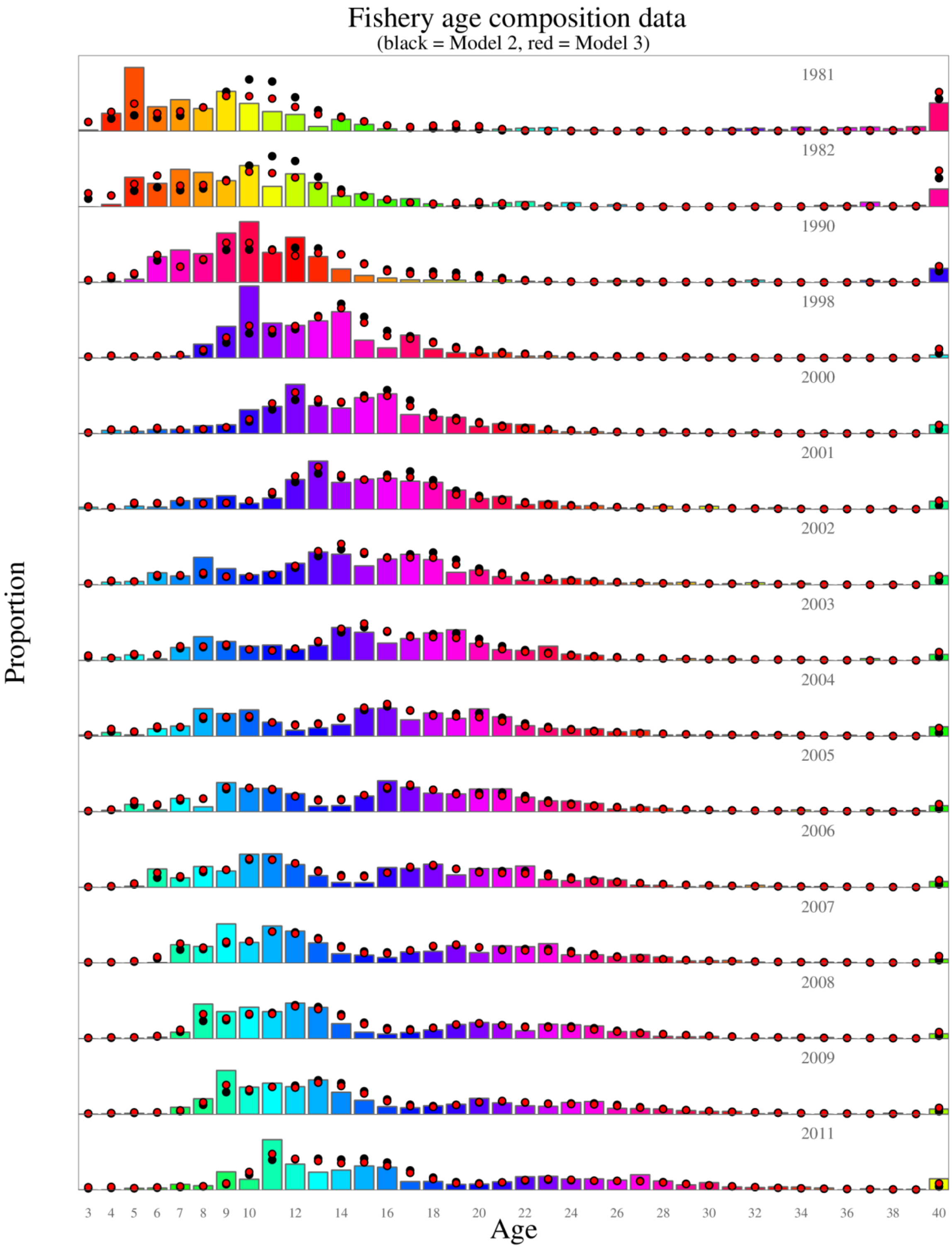

Figure 10. Fits to fishery age composition data for Models 2 and 3.

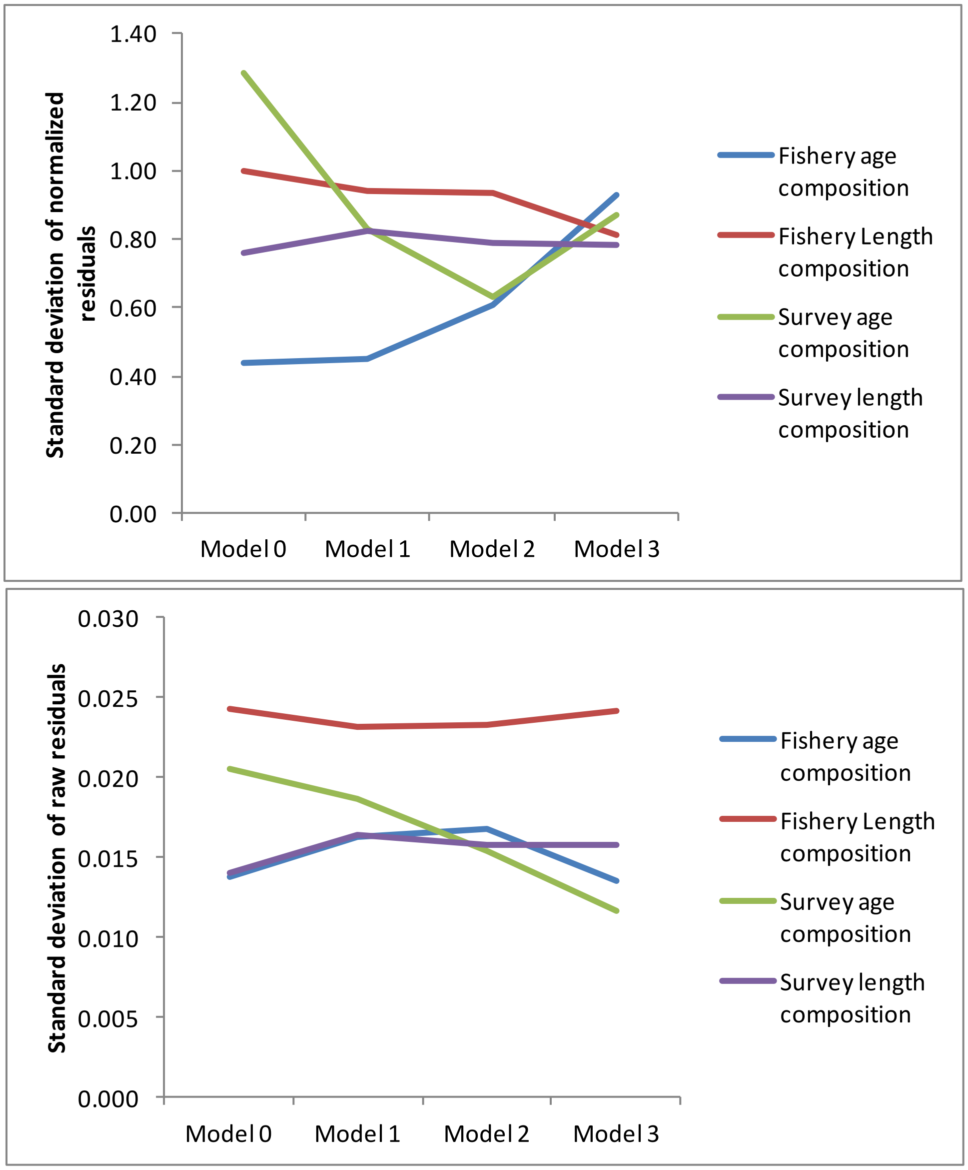

Figure 11. The standard deviations of the normalized and raw residuals of the age and length composition data for the four models evaluated.