ICCAT WORKING DOCUMENT
SCRS/98/58

A Flexible Forward Age-Structured Assessment Program

Christopher M. Legault[1] and Victor R. Restrepo[2]

[1] U.S. Department of Commerce
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
75 Virginia Beach Dr., Miami, Florida, 33149, USA

[2] University of Miami
Rosenstiel School of Marine and Atmospheric Science
Cooperative Unit for Fisheries Education and Research
4600 Rickenbacker Causeway, Miami, Florida 33149, USA

September 1998

Sustainable Fisheries Division Contribution SFD-98/99-16

Summary

This paper documents an age-structured assessment program (ASAP) which incorporates various modeling features that have been discussed by the SCRS in recent years, particularly during meetings of the bluefin tuna species group. The software was developed using the commercial package AD Model Builder, an efficient tool for optimization that uses an automatic differentiation algorithm in order to find a solution quickly using derivatives calculated to within machine precision, even when the number of parameters being estimated is rather large. The model is based on forward computations assuming separability of fishing mortality into year and age components. This assumption is relaxed by allowing for fleet-specific computations and by allowing the selectivity at age to change smoothly over time. The software can also allow the catchability associated with each abundance index to vary smoothly with time. The problem's dimensions (number of ages, years, fleets and abundance indices) are defined at input and limited by hardware only. We illustrate an application of ASAP using data for western Atlantic bluefin tuna.

Introduction

Stock assessment algorithms explain observed data through a statistical estimation procedure based on a number of assumptions. The number and severity of these assumptions are determined by the algorithm and reflect not only the user's paradigms but also the amount and quality of the available data. We present an age-structured assessment program (ASAP) which allows easy comparison of results when certain assumptions are made or relaxed. Specifically, ASAP is a flexible forward program that allows the assumption of separability of gear specific fishing mortality into year and age components to be relaxed and change over time. The assumption of constant catchability coefficients for scaling observed indices of abundance can also be relaxed to change over time. The advantage of this flexibility is an increased ability to fit models and less reliance on assumptions that are thought to be too strict. The disadvantage of such an approach is exactly this ability to explain the data in more (and possibly contradictory) ways through different choices in the amount of variability in the changing parameters. Explicit choices for relative weightings amongst the different parts of the objective function must be made. Slight changes in these parameter weightings in a complex model can produce vastly different results, while a simpler model will be more consistent (not necessarily more accurate) relative to changes in the parameter weightings.

Allowing flexibility in selectivity and catchability greatly increases the number of parameters to be estimated. We use the commercial software package AD Model Builder to estimate the relatively large number of parameters. The software package is based on a C++ library of automatic differentiation code (see Greiwank and Corliss 1991) which allows relatively fast convergence by calculating derivatives to machine precision accuracy. These derivatives are used in a quasi-Newton search routine to minimize the objective function. The array sizes for parameters are defined on input and limited only by hardware. Currently, ASAP is compiled to estimate a maximum of 5,000 parameters, but this can be increased by changing one line of code.

The AD Model Builder software package allows many matrix operations to be programmed easily in its template language and allows for the estimation of parameters to occur in phases. The phases work by estimating only some parameters initially and adding more parameters in a stepwise fashion until all parameters are estimated. When new parameters are added by incrementing the phase, the previously estimated parameters are still estimated, not fixed at the previous values. These phases also allow easy switching between simple and complex models by simply turning on or off phases through the input file. For example, index specific catchability coefficients can be allowed to change or have a constant value over time. An additional feature of the AD Model Builder software is easy likelihood profiling of specified variables, although this can be time consuming for models with large numbers of parameters. We first describe ASAP with all the features and then compare two analyses for bluefin tuna using different levels of complexity in the program.

The Model

Population dynamics

The model's population dynamics follow a standard form common to forward-projection methods such as those of Fournier and Archibald (1982), Deriso et al. (1985), Methot (1998), Ianelli and Fornier (1998), and Porch and Turner (In Press). Catches and fishing mortalities can be modeled as being fleet-specific.

Let a = age, 1…A,
y = year, 1…Y
g = fleet 1….G
u = abundance index series, 1….U

Selectivity (*S*) at age within a year by a fleet can be limited to a range of ages and averages one, as opposed to having a maximum of one,

$$\frac{\sum_{a(g_{start})}^{a(g_{end})} S_{a,y,g}}{a(g_{end}) - a(g_{start}) + 1} = 1.0 \tag{1}$$

where $a(g_{start})$ and $a(g_{end})$ denote the starting and ending ages for the gear's selectivity. The output of the program makes the simple conversion from averaging one to having a maximum of one in order to simplify comparisons with other models.

Fishing mortality is modeled as the product of the selectivity at age within a year by a fleet and a year and fleet specific fishing mortality multiplier ($Fmult_{y,g}$)

$$F_{a,y,g} = S_{a,y,g} Fmult_{y,g} \ . \tag{2}$$

Total fishing mortality at age and year is the sum of the fleet specific fishing mortality rates

$$Ftot_{a,y} = \sum_g F_{a,y,g} \tag{3}$$

and adding the natural mortality rate (*M*) produces the total mortality rate

$$Z_{a,y} = Ftot_{a,y} + M_{a,y} \ . \tag{4}$$

The catch by age, year and fleet is

$$C_{a,y,g} = \frac{N_{a,y} F_{a,y,g} (1 - e^{-Z_{a,y}})}{Z_{a,y}} \tag{5}$$

where *N* denotes population abundance at the start of the year.

The yield by age, year and fleet is

$$Y_{a,y,g} = C_{a,y,g} W_{a,y} \tag{6}$$

where $W_{a,y}$ denotes weight of an individual fish of age *a* in year *y*.

The proportion of catch at age within a year for a fleet is

$$P_{a,y,g} = \frac{C_{a,y,g}}{\sum_a C_{a,y,g}}. \tag{7}$$

The forward projections begin by computing recruitment as deviations from an average value

$$N_{1,y} = \overline{N}_1 e^{u_y} \tag{8}$$

where $?_y$~N(0,s_{Ny}^2) and the other numbers at age in the first year as deviations from equilibrium

$$\begin{aligned} N_{a,1} &= N_{1,1} e^{-\sum_{i=1}^{a-1} Z_{i,1}} e^{y_a} \qquad for\, a < A \\ N_{a,1} &= \frac{N_{1,1} e^{-\sum_{i=1}^{a-1} Z_{i,1}}}{1 - e^{-Z_{A,1}}} e^{y_a} \qquad for\, a = A \end{aligned} \tag{9}$$

where $?_a$~N(0,s_{Na}^2). The remaining population abundance at age and year is then computed

$$\begin{aligned} N_{a,y} &= N_{a-1,y-1} e^{-Z_{a-1,y-1}} \qquad\qquad for\ a < A \\ N_{a,y} &= N_{a-1,y-1} e^{-Z_{a-1,y-1}} + N_{a,y-1} e^{-Z_{a,y-1}} \qquad for\ a = A. \end{aligned} \tag{10}$$

Predicted indices of abundance ($\hat{I}$) are a measure of the population scaled by catchability coefficients (q) and selectivity at age (S)

$$\hat{I}_{u,y} = q_{u,y} \sum_{a(u_{start})}^{a(u_{end})} S_{u,a,y} N^{*}_{a,y} \tag{11}$$

where $a(u_{start})$ and $a(u_{end})$ are the index specific starting and ending ages, respectively, and N^* corresponds to the population abundance in either numbers or weight at a specific time during the year. The abundance index selectivity at age can either be input or linked to a specific fleet. If the latter is chosen, the age range can be smaller than that of the fleet and the annual selectivity patterns are rescaled to equal 1.0 for a specified age (a_{ref})such that the catchability coefficient is linked to this age

$$S_{u,a,y} = \frac{S_{a,y,g}}{S_{a_{ref},y,g}}. \tag{12}$$

Time-varying parameters

Fleet specific selectivity and catchability patterns are allowed to vary over time in the model. Changes in selectivity occur each t_g years through a random walk for every age in a given fleet

$$S_{a,y+t,g} = S_{a,y,g} e^{e_{a,y,g}} \tag{13}$$

where $e_{a,y,g}$~N(0,s_{Sg}^2) and are then rescaled to average one following equation (1). If t_g is greater than one, then the selectivity at age for the fleet is the same as previous values until t_g years elapse. The catchability coefficients also follow a random walk

$$q_{u,y+1} = q_{u,y}e^{w_{u,y}} \ , \tag{14}$$

as do the fleet specific fishing mortality rate multipliers

$$Fmult_{y+1,g} = Fmult_{y,g}e^{h_{y,g}} \tag{15}$$

where $?_{u,y}$~N(0,s_{qu}^2) and $?_{y,g}$~N(0,s_{Fg}^2).

<u>Parameter estimation</u>

The number of parameters estimated depends upon the values of t_g and whether or not changes in selectivity or catchability are considered. When time varying selectivity and catchability are not considered the following parameters are estimated: *Y* recruits, *A*-1 population abundance in first year, *YG* fishing mortality rate multipliers, *AG* selectivities (if all ages selected by all gears), *U* catchabilities, and 2 stock recruitment parameters. Inclusion of time varying selectivity and catchability can increase the number of parameters to be estimated by a maximum of (*Y*-1)*AG* + (*Y*-1)*U*. Sensitivity analyses can be conducted to determine the tradeoffs between number of parameters estimated and goodness of fit caused by changes in the t_g values.

The likelihood function to be minimized includes the following components (ignoring constants): total catch in weight by fleet (lognormally distributed)

$$L_1 = l_1[\ln(\sum_a Y_{a,y,g}) - \ln(\sum_a \hat{Y}_{a,y,g})]^2 \ ; \tag{16}$$

catch proportions in numbers of fish by fleet (multinomially distributed)

$$L_2 = -\sum_y \sum_g l_{2,y,g} \sum_a P_{a,y,g} \ln(\hat{P}_{a,y,g}) - P_{a,y,g} \ln(P_{a,y,g}) \ ; \tag{17}$$

and indices of abundance (lognormally distributed)

$$L_3 = \sum_g l_{3,g} \sum_y [\ln(I_{y,g}) - \ln(\hat{I}_{y,g})]^2 / 2s^2_{y,g} + \ln(s_{y,g}) \ , \tag{18}$$

where variables with a hat are estimated by the model and variables without a hat are input as observations. The second term in the catch proportion summation causes the likelihood to equal zero for a perfect fit. The sigmas in equation 18 are input by the user and can optionally be set to all equal 1.0 for equal weighting of all index points. The weights (*?*) assigned to each component of the likelihood function correspond to the inverse of the variance assumed to be associated with that component. Note that the year and fleet subscripts for the catch proportion lambdas allow zero weights to be assigned to specific year and fleet combinations such that only the total catch in weight by that fleet and year would be incorporated in the objective function. Priors for the variances of the time varying parameters are also included in the likelihood by setting *?* equal to the inverse of the assumed variance for each component

$$L_4 = \sum_g l_{4,g} \sum_a \sum_y e^2_{a,y,g} \qquad (selectivity) \tag{19}$$

$$L_5 = \sum_u l_{5,u} \sum_y w^2_{u,y} \qquad (catchability) \tag{20}$$

$$L_6 = \sum_g l_{6,g} \sum_y h^2_{y,g} \qquad (F\ multipliers) \tag{21}$$

$$L_7 = l_7 \sum_y u^2_y \qquad (recruitment) \tag{22}$$

$$L_8 = l_8 \sum_y y^2_y \qquad (N\ year 1)\,. \tag{23}$$

Additionally, there is a prior for fitting a Beverton and Holt type stock-recruitment relationship

$$L_9 = l_9 \sum_y [\ln(N_{1,y}) - \ln\left(\frac{a\ SSB_{y-1}}{b + SSB_{y-1}}\right)]^2 \tag{24}$$

where *SSB* denotes the spawning stock biomass and *a* and *ß* are parameters to be estimated. Penalties are used to determine the amount of curvature allowed in the fleet selectivity patterns, both at age

$$r_1 = l_{r1} \sum_y \sum_g \sum_{a(g_{start})}^{a(g_{end})-2} (S_{a,y,g} - 2S_{a+1,y,g} + S_{a+2,y,g})^2 \tag{25}$$

and over time

$$r_2 = l_{r2} \sum_a \sum_g \sum_{y=1}^{Y-2} (S_{a,y,g} - 2S_{a,y+1,g} + S_{a,y+2,g})^2\,. \tag{26}$$

The function to be minimized is then the sum of the likelihoods and penalties

$$L = L_1 + L_2 + L_3 + L_4 + L_5 + L_6 + L_7 + L_8 + L_9 + r_1 + r_2\,. \tag{27}$$

An additional penalty is utilized in early phases of the minimization to keep the average total fishing mortality rate close to the natural morality rate. This penalty ensures the population abundance estimates do not get exceedingly large during early phases of the minimization. The final penalty added to the objective function forces the parameters for fleet selectivities in the first year to average 1.0. This penalty prevents multiple parameter sets from having the same objective function value, which would cause difficulty for the minimization routine. Each component of the objective function is reported in the output file along with the corresponding number of observations, weight assigned to that component, and residual sum of squared deviations (if appropriate).

Additional Features

The model optionally does some additional computations once the likelihood function has been minimized. These "extras" do not impact the solution, they are merely provided for reference. Each fleet can be designated as either directed or nondirected for the projections and F reference point calculations, with the option to modify the nondirected F in the future. The directed fleets are combined to form an overall selectivity pattern that is used to solve for common fishing mortality rate reference points ($F_{0.1}$, F_{max}, $F_{30\%SPR}$, $F_{40\%SPR}$ and F_{msy}) and compared to the terminal year F estimate. The inverse of the SPR for each of these points is also given so replacement lines corresponding to these reference values can be plotted on the spawner-recruit relationship. Projections are computed using either the stock-recruitment relationship or input values to generate future recruitment. The projections for each successive year can be made using either a total catch in weight or the application of a static $F_{X\%SPR}$, where X is input. A reference year is also input that allows comparison of the spawning stock biomass (SSB) in the terminal year and that in the final projection year as SSB_y/SSB_{ref}. Likelihood profiles for these SSB ratios can optionally be generated.

Example: Western Atlantic Bluefin Tuna

Two analyses of western Atlantic bluefin tuna data using ASAP are presented here. The first analysis (simple) did not allow selectivity and catchability to change over time (225 parameters estimated). The second analysis (complex) used the full complexity allowed by the model, with fleet selectivities allowed to change every two years and index catchabilities allowed to change every year (914 parameters estimated). In both analyses the model was structured for years 1970-1995, ages 1-10+, five fleets, and seven tuning indices (each point input with a variance) with all likelihood component weightings equal between the analyses. The natural mortality rate was set at 0.14 for all ages (for data details see Restrepo and Legault In Press). The number of observations associated with, and the weights given to, each part of the likelihood function are shown in Table 1. In this example, the weights assigned to each component were chosen arbitrarily. In an actual assessment, these weights will need to be selected by the assessment working group.

The overall fit of the complex analysis was better than the simple analysis (lower objective function value) as expected due to the greater number of parameters (Table 1). The complex analysis fits the indices better than the simple analysis, especially the US Rod and Reel Large, US Longline Gulf of Mexico, and the Japan Longline Gulf of Mexico indices. (Figure 1). Recruitment estimates from the two analyses are similar to the estimates from the 1996 SCRS assessment, which used virtual population analysis (VPA) with the main differences occurring in the early years of the time series (Figure 2). The estimates of spawning stock biomass (SSB) differ between the analyses, the complex one is similar in magnitude to the SCRS96 results, while the simple analysis estimates larger values (Figure 3). However, standardizing the SSB trends (dividing by the SSB in 1975) produces similar trends for all three analyses (Figure 3). The resulting stock-recruitment relationship is shown in figure 4. The total fishing mortality rates by year and age differ in both magnitude and pattern, with the complex analysis more closely matching the 1996 SCRS assessment (Figure 5). These differences in F are due to the assumptions about selectivity, fixed for the simple analysis and allowed to vary for the complex one (Figure 6). Note in particular the large change in selectivity of the purse seine fleet, mainly young fish in the early years and old fish in recent years. The catchability values also reflect the difference in assumptions, constant for the simple analysis and allowed to vary in the complex analysis (Figure 7). Note the large lambda given to the larval index causes the catchability coefficients to vary only slightly in the complex analysis. The catch at age proportions are fit relatively well in both analyses, the input and effective sample sizes are similar, even though this is the largest part of the total likelihood. The estimated effective sample size can be computed as

$$Effective\,N_g = \frac{\sum_a \sum_y \hat{p}_{a,y,g}(1-\hat{p}_{a,y,g})}{\sum_a \sum_y (p_{a,y,g}-\hat{p}_{a,y,g})^2} \qquad (28)$$

(for details see McAllister and Ianelli, 1997 Appendix 2).

Discussion

The flexibility afforded by ASAP is a continuation of the trend in stock assessment programs from the relatively simple structure of Fournier and Archibald (1982) to the more flexible structure found in Methot (1998), Ianelli and Fournier (1998), and Porch and Turner (In Press). In fact, ASAP is based on the same logic as these more flexible programs, but combines the advantages of the AD Model Builder software with the more general input flexibility of stock synthesis and CATCHEM. J. Ianelli (NMFS, Seattle, pers. comm.) also provided guidance in the formulation of certain model components, specifically the logic of linking fleet specific indices with a specific age in the tuning process (see equation 12). The distinguishing feature between this approach and that found in virtual population analysis (VPA) (Gavaris 1988, Powers and Restrepo 1992) is that VPA assumes the catch at age is measured without error, while ASAP assumes the observed catch at age varies about its true value.

The flexibility of ASAP can also cause problems however. Slight changes in the weights assigned to each likelihood component can produce different results, both in magnitude and trend. The large number of parameters, in the complex model especially, required the solutions in each phase to progress towards a satisfactory region in the solution space. If any phase led the solution away from this region, the final result will not be believable (e.g. total F<1e-5). This problem was not found in multiple tests using simulated data that did not contain errors or only small observation errors. Thus, the ability to fit highly complex models depends upon the quality of the data available, especially the consistency between the catch at age and the tuning indices. Nevertheless, the flexible nature of ASAP allows for easy exploration of the data to determine what level of complexity can appropriately be modeled.

Acknowledgments

The conclusions presented here are entirely the authors' and are not necessarily endorsed by NMFS or UM. We are grateful to Jim Ianelli, Clay Porch, Joe Powers, Gerry Scott and Steve Turner for helpful discussions.

References

Deriso, R.B., T.J. Quinn II and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Fournier, D. and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Greiwank, A. and G.F. Corliss (eds). 1991. Automatic differentiation of algorithms: theory, implementation and application. Proceedings of the SIAM Workshop on the Automatic Differentiation of Algorithms, held Jan. 6-8, Breckenridge, CO. Soc. Indust. and Applied Mathematics, Philadelphia.

ICCAT. 1997. Report for biennial period 1996-1997. Part I (1996), Vol. 2. Int. Comm. Cons. Atl. Tunas. 204 pp.

Ianelli, J.N. and D.A. Fournier. 1998. Alternative age-structured analyses of the NRC simulated stock assessment data. NOAA Tech. Memo. NMFS-F/SPO-30. pp. 81-96.

McAllister, M.K. and J.N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54:284-300.

Methot, R. 1998. Application of stock synthesis to NRC test data sets. NOAA Tech. Memo. NMFS-F/SPO-30. pp. 59-80.

Porch, C.E. and S.C. Turner. In Press. Catch-at-age analyses of west Atlantic bluefin tuna incorporating data from 1960 to 1994 (preliminary results). Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. (Working Document SCRS/96/119).

Powers, J.E. and V.R. Restrepo. 1992. Additional options for age-sequenced analysis. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 39:346-354.

Restrepo, V.R. and C.M. Legault. In Press. A stochastic implementation of an age-structured production model. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. (Working Document SCRS/97/59).

Table 1. Likelihood function components for two ASAP analyses. nobs=number of observations in that component, ?=weight given to that component, RSS=residual sum of squared deviations, L=likelihood value

			Simple		**Complex**	
Component	nobs	?	RSS	L	RSS	L
Total Catch in Weight						
Rod and Reel	26	100.5	0.0005	0.0479	0.0001	0.0147
Japan Longline	26	100.5	0.0015	0.1558	0.0003	0.0322
Other Longline	26	100.5	0.0001	0.0069	0.0001	0.0070
Purse Seine	26	100.5	0.0002	0.0183	0.0039	0.3913
Other	26	100.5	0.0001	0.0065	0.0000	0.0026
Total	130	100.5	0.0023	0.2353	0.0045	0.4477
Catch at Age Proportions	1300	N/A	N/A	874.40	N/A	396.47
Index Fits						
Larval Index	16	1	5.26	11.95	5.29	11.61
US Rod and Reel Small	15	1	3.95	9.33	2.02	-1.02
Canadian Tended Line	15	1	2.08	3.05	0.64	-5.95
US Rod and Reel Large	13	1	1.76	1.22	0.39	-5.74
US Longline Gulf of Mexico	9	1	6.13	15.26	0.31	-3.79
Japan Longline Gulf of Mexico	8	1	0.74	1.10	0.58	1.05
Japan Longline NW Atlantic	20	1	3.22	9.51	0.58	-9.19
Total	96	7	23.15	51.43	9.80	-13.02
Selectivity Deviations						
Rod and Reel	12	0.1	0	0	2.52	0.25
Japan Longline	12	0.1	0	0	4.42	0.44
Other Longline	12	0.1	0	0	3.56	0.36
Purse Seine	12	0.1	0	0	8.74	0.87
Other	12	0.1	0	0	3.00	0.30
Total	60	0.5	0	0	22.25	2.22
Catchability Deviations						
Larval Index	16	1000	0	0	0.00	0.29
US Rod and Reel Small	15	6.7	0	0	0.51	3.43
Canadian Tended Line	15	6.7	0	0	0.37	2.45
US Rod and Reel Large	13	6.7	0	0	0.18	1.20
US Longline Gulf of Mexico	9	6.7	0	0	0.21	1.39
Japan Longline Gulf of Mexico	8	6.7	0	0	0.00	0.03
Japan Longline NW Atlantic	20	6.7	0	0	0.35	2.35
Total	96	1040.2	0	0	1.62	11.14
Fmult Deviations						
Rod and Reel	25	0.1	5.26	0.53	5.01	0.50
Japan Longline	25	0.1	21.44	2.14	19.67	1.97
Other Longline	25	0.1	24.30	2.43	23.97	2.40
Purse Seine	25	0.1	5.24	0.52	8.07	0.81
Other	25	0.1	5.60	0.56	6.84	0.68
Total	125	0.1	61.84	6.18	63.56	6.36
Recruitment	26	0.01	10.14	0.10	14.51	0.15
N in Year 1	9	1.44	3.34	4.82	3.08	4.43
Stock-Recruit Fit	25	0.001	9.47	0.01	3.94	0.00
Selectivity Curvature over Age	40	1.44	12.03	17.32	17.19	24.76
Selectivity Curvature over Time	1200	1.44	0	0	52.03	74.92
F penalty	260	0.001	3.0E-01	3.0E-4	2.3E-02	2.3E-02
Mean Sel Year 1 Penalty	50	1	4.5E-12	4.5E-12	4.7E-12	4.7E-12
Objective Function Value				**954.50**		**507.87**

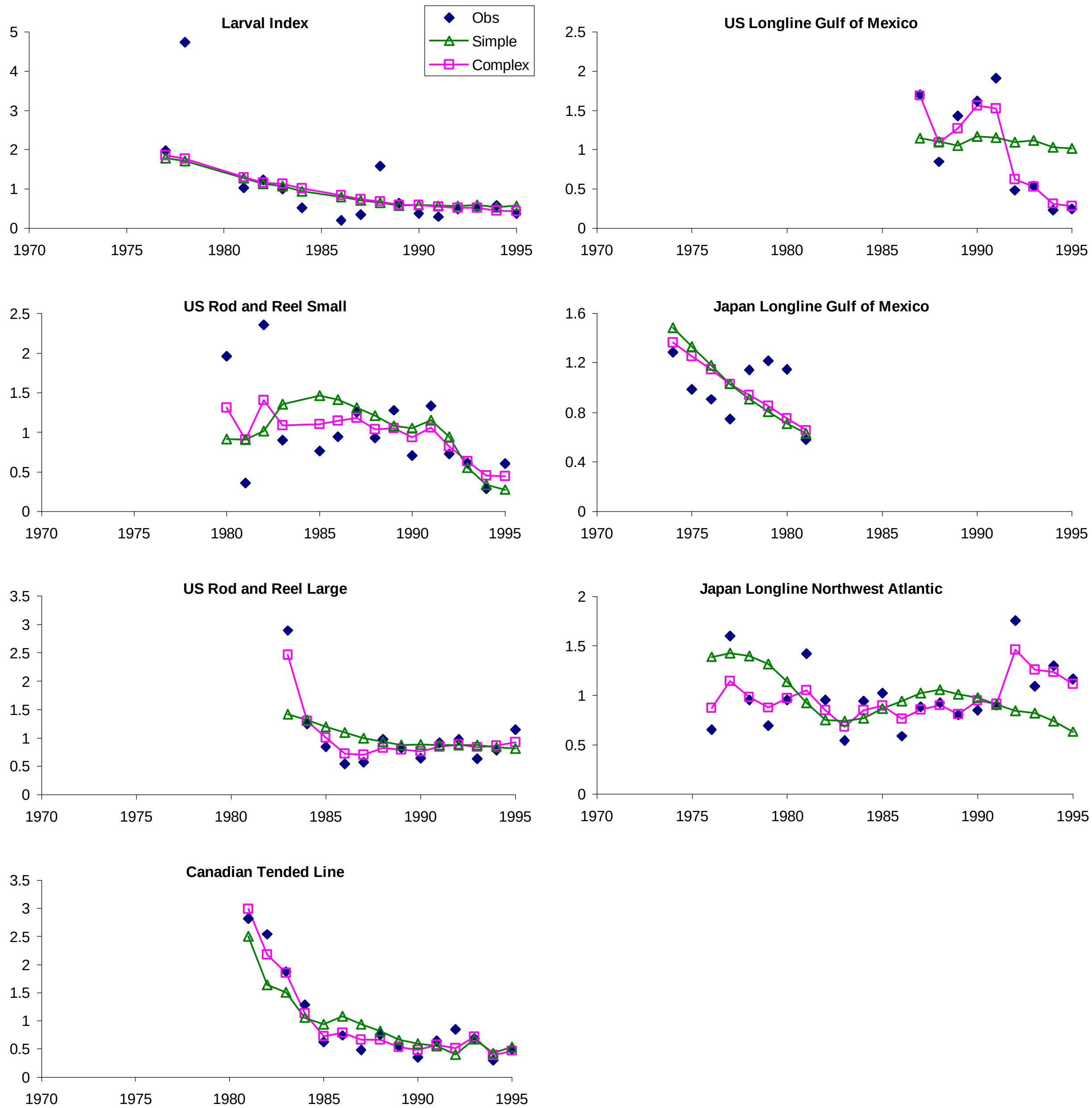

Figure 1. Observed and predicted indices for the simple and complex ASAP analyses.

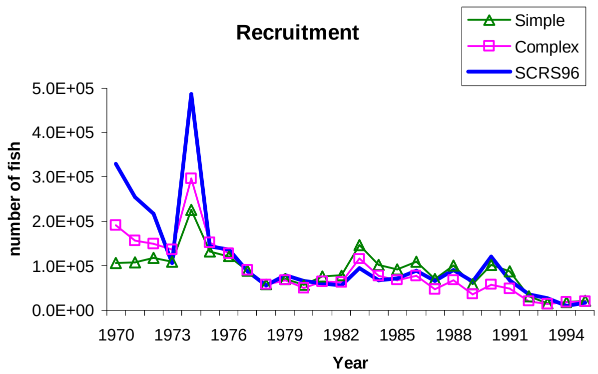

Figure 2. Estimated recruitment from two ASAP analyses and the SCRS 1996 assessment.

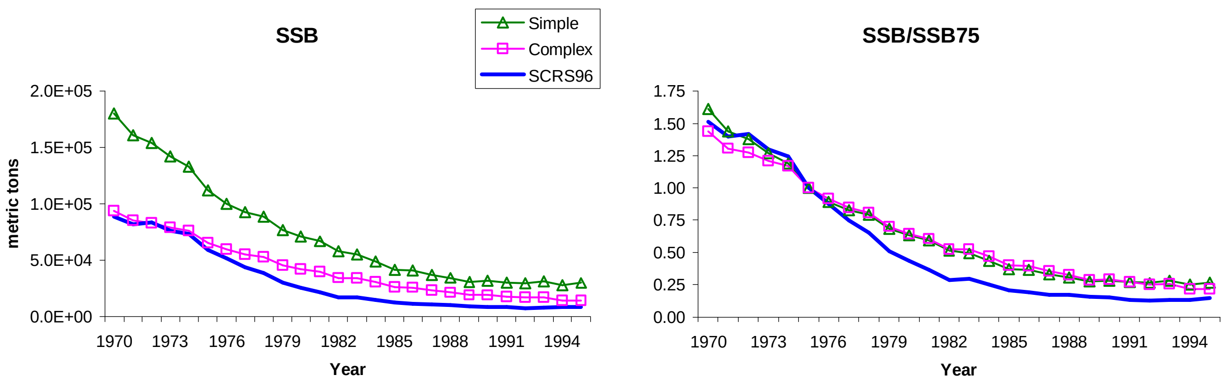

Figure 3. Spawning stock biomass (SSB) from two ASAP analyses and SCRS 1996.

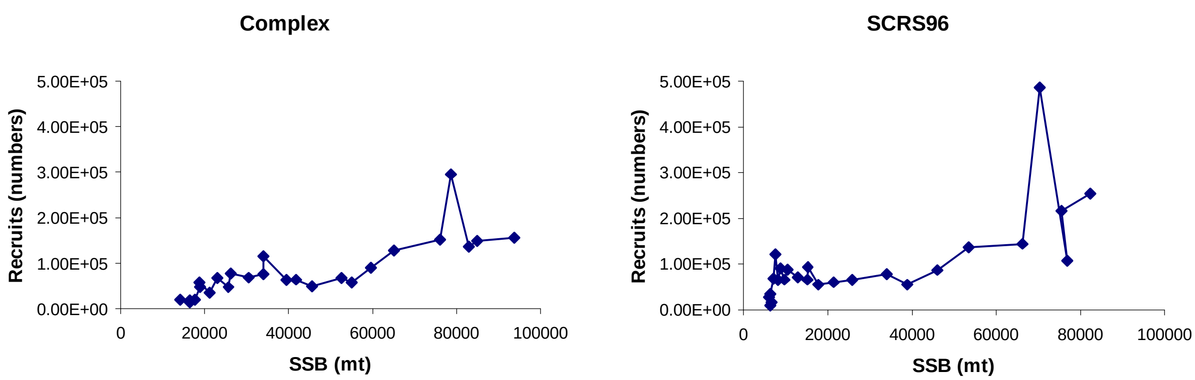

Figure 4. Complex ASAP analysis and SCRS 1996 stock-recruitment relationships.

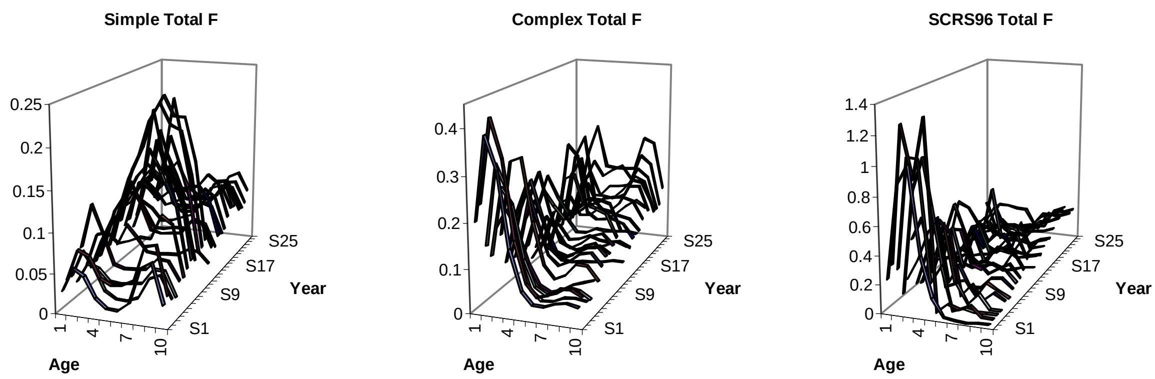

Figure 5. Estimated fishing mortality rates by age and year for two ASAP analyses and SCRS 1996.

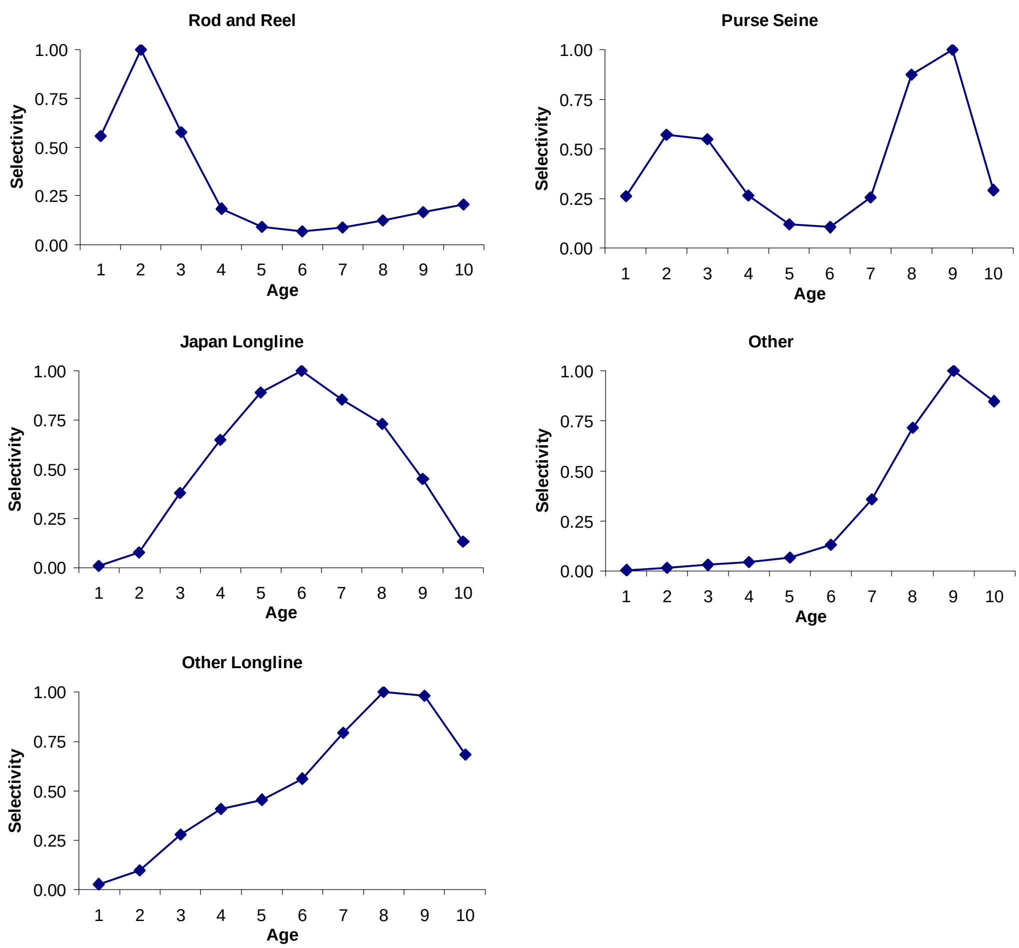

Figure 6a. Selectivity at age for the simple ASAP analysis, constant over all years for each fleet.

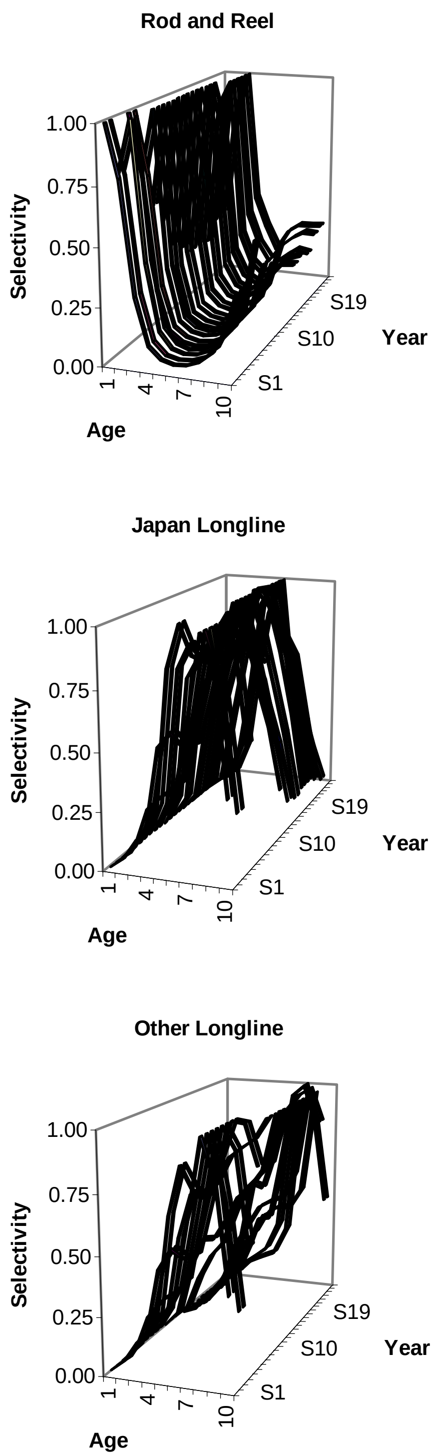

Figure 6b. Selectivity at age for the complex ASAP analysis.

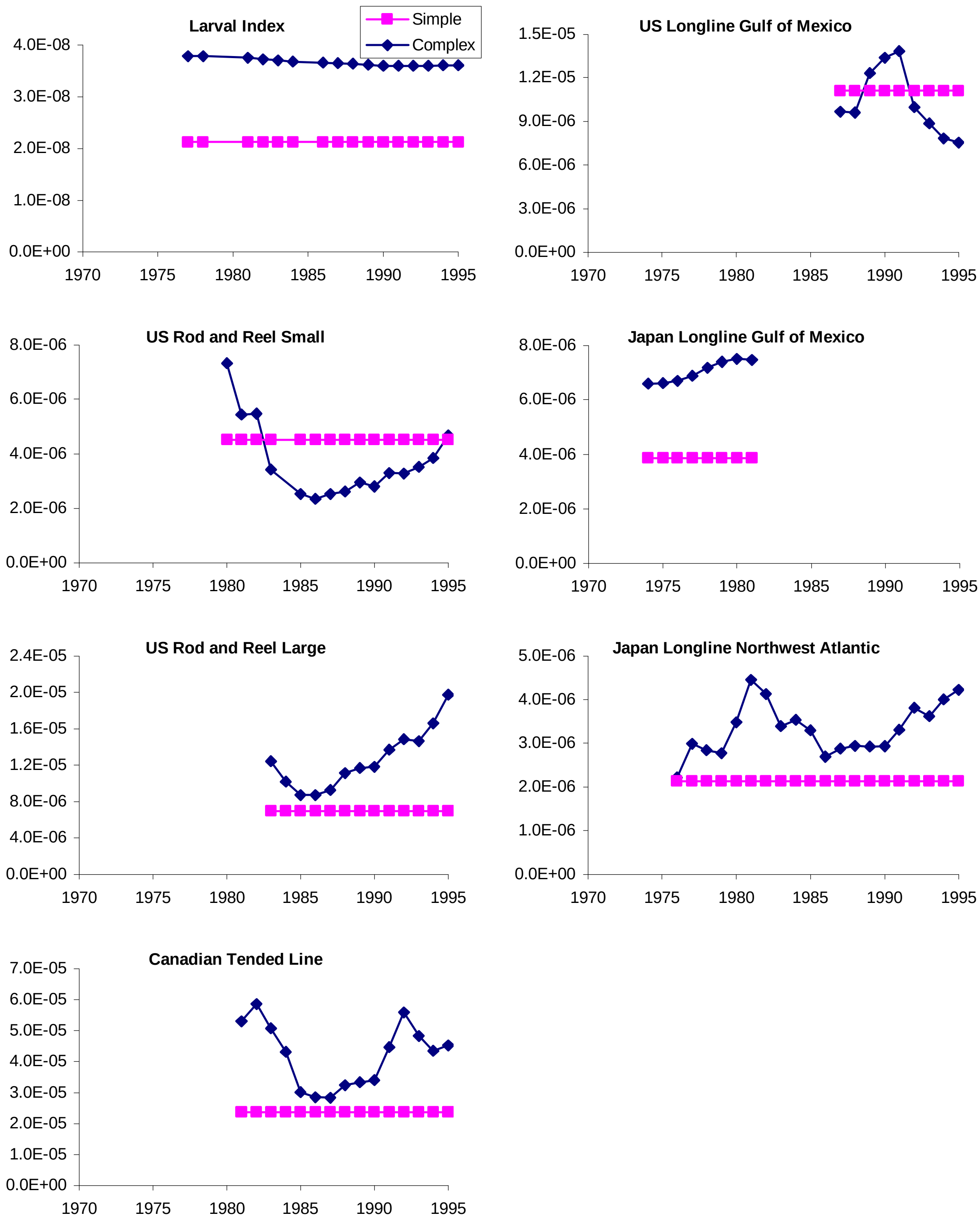

Figure 7. Catchability for each tuning index from the two ASAP analyses.