Preliminary Population Dynamics Model of Pacific Bluefin Tuna

Steven L. H. Teo[1], Kevin Piner[1]

[1]NOAA Fisheries
Southwest Fisheries Science Center
8604 La Jolla Shores Drive
La Jolla CA. 92037 USA.

Nov. 2012

Working document submitted to the ISC Pacific bluefin tuna Working Group, International Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (ISC), 10-17 November 2012, Honolulu, Hawaii, USA. **Document not to be cited without author's permission.**

Summary

This paper presents a dynamic model of Pacific Bluefin tuna that follows a modeling approach advocated by Francis (2011). Improvement in the representation of the primary tuning indices was the goal of this work. We considered several potential methods to reduce the conflict between indices of abundance and composition data including: statistical down-weighting of composition data, the addition of model process in the form of time varying selection patterns, and a hybrid approach that modeled composition at a fine temporal scale that fixed the selection parameters to an estimate and then did not use the composition in the model total likelihood. Conflict between indices was handled by creation of separate models that represented the trends of the different indices. Two models are forwarded that tune to either the Japanese Coastal Longline CPUE (ModS1) or the Taiwanese Longline CPUE (ModS9). Overall, both models show that spawning biomass has declined in the last decade but the most recent dynamics are different due to the timing of the decline.

Introduction

Pacific Bluefin tuna (PBF) are a highly prized fish stock with a long history of harvest by multiple Pacific Ocean nations (ISC 2012). Bluefin tuna show a strong migratory pattern (Polovina 1996) with spawning in the Western Pacific (WPO) and subsequently, an unknown and likely variable fraction (Piner et al. 2009) migrating to the Eastern Pacific (EPO). All stages of PBF are taken by commercial fleets. Typically, juveniles are taken in more surface waters in both the WNPO and EPO using gears such as troll and purse seine while adults are taken at depth using longline gears. The composition and proportion of fleets has undergone changes in the last 50 years with a trend of increasing catch of juvenile PBF in the recent period. Due to the complexity and variability of the fisheries and migratory patterns of the stock, population models are difficult to develop that represents all data well.

For this paper we generally follow the spirit of the guidelines given by Francis (2011) on data weightings in statistical stock assessment models: 1) do not let other data stop the model from fitting indices of abundance, 2) account for the potential high level of correlations between composition observations, and 3) if an index of abundance is thought to be unrepresentative (in conflict with other indices of abundance), do not down-weight the series but use it in an alternative model. To apply these guidelines applied to the Pacific Bluefin Tuna (PBF) stock assessment requires answering 2 questions. 1) Is there enough conflict between indices of abundance to warrant alternative models? 2) Is there enough conflict between composition data and indices that it is unacceptably impacting the model's fit to the indices of abundance?

What constitutes significant model conflict between indices and composition data or conversely what is an acceptable fit to the indices is both a statistical and subjective decision. Although subjectivity is necessarily problematic and leads to potentially unresolvable differences in opinions, it is also an unavoidable part of the dynamic modeling process. For this paper we employ two measures of "goodness of fit" to the indices of abundance. The first is that the variability of the observed measures of relative abundance should not vary too greatly around the expectation of the model (i.e., root mean square error: RMSE). Second, we subjectively determine that the major trends in the index are represented by the model. On the second point, a more statistical measure of long periods of one-directional residuals could be employed; however many of our series are relatively short and it will be difficult to determine if runs in residuals are problematic.

If model results indicate an unacceptable fit to the indices, the next question is how to improve the model's fit to this data. If the conflict is between two indices of abundance then alternative models containing each index will be developed. When there is apparent conflict between composition data and indices we will employ two procedures (either or both) to resolve conflicts: 1) down-weighting of composition data and 2) adding more process to the model to improve fit to composition data. Additional model process can be the estimation of addition selection parameters, especially via time-varying parameters or more flexible shapes (e.g., cubic splines). A potential third option is two combine both methods and use the data with high weight to derive selection parameters but down-weight the data in the final model run to prevent misfit to size composition from overly-influencing the models ability to track the indices of abundance. This third option explicitly says we have confidence in the data, but do not have the ability to introduce the appropriate model process to match it.

This paper describes our methods to produce a population dynamics model using Stock Synthesis (SS) and the accepted data time series from the May/June PBF workshop that is consistent with the modeling perspective described above.

Materials and Methods

Data

The data for this model is the SS data file (.dat) from the May/June meeting that includes changes to F4 as adopted by the WG during the meeting. Subsequent changes to the data file made for this paper are described as follows. 1) An examination of the input sample sizes showed that some sample sizes of the F3 fishery were abnormally large (e.g. N=377 in 1996). We therefore used an asymptotic curve (Beverton-Holt function) to estimate maximum sample size for F3 based on the observed and effective sample sizes for this fishery. The maximum sample size for F3 was set to 10.3406. Preliminary data analysis and model runs indicated that size compositions in season 2 of F2, and season 4 in F1 and F9 showed seasonality to the modes in the composition data. Seasonal fisheries for season 2 of F2 (F2s2), season 4 of F1 (F1s4) and F9 (F9s4) were therefore created in the data file from the original F2, with its separate catch and size compositions.

General Model structure

This assessment uses SS (V3.23b) as a single sex model estimating dynamics from 1952-2010. Data was compiled and fit quarterly (quarter 1: July-September), with the exception of standardized abundance indices (CPUE), which were treated as annual indices of abundance which represent available biomass midway through a quarter.

Recruitment is modeled assuming a Beverton-Holt relationship with steepness, h=0.999 (ISC 2011). An initial standard deviation of recruitment deviations was specified δ_R=0.6. Additional recruitment parameters of *LN*(*R0*) and *R1* equilibrium recruitment offset were estimated along with deviations from the S/R curve (1952-2009). Because of the paucity and unevenness of data from 1952-1990 we ramp up the full implementation of bias corrections on recruitment estimates linearly from 1952-1990. Key biology and life history assumptions are given in Table 1.

The model attempts to allow significant flexibility in starting conditions so that strong assumptions do not control model results. Initial catch was estimated by not fitting to an equilibrium catch but estimating equilibrium F's for two fisheries F1 and F5 (large and small fish). In addition, 6 initial recruitment deviations were estimated (1946-1951) prior to full dynamics to better fit to the initial period with data.

Likelihood components by fleet

Data available for use in the model are given in Table 2. Parameterization of selectivity patterns by fleet is given in Table 3. For all fleets, catch assumes lognormal error assumption with SE=0.1. Composition assumes multinomial error assumptions with variance (sample size) being the same as the May/June workshop except F3 as noted above. Indices of abundance (CPUE) assume lognormal error distribution described by coefficient of variation (CV) taken from the General Linear Model used in the standardization. The minimum CV for abundance indices is assumed 0.2.

Japanese Longline (F1)

Catch, size composition and CPUE data are available and fitted in the model. F1 is deconstructed into two fleets (F1a and F1b) which contain the data for seasons 1-3; and season 4, respectively. Selection pattern is parameterized as domed and two separate time periods of selection patterns estimated (1952-1992 and 1993-2010) in order to match the assumed changes in catchability of the CPUE time series.

Three CPUE series (S1; S2; S3) treated as indices of relative abundance are included. The selectivity of the indices were assumed to mirror that of F1b because the CPUE series were based data from the spawning season (Ichinokawa and Takeuchi 2012). A separate catchability (*q*) parameter is estimated for each series and assumed to be linear.

Small Pelagic Purse Seine (F2)

Catch and size composition (2001-2010) data are available and fitted in the model. F2 is deconstructed into two fleets (F2a and F2b) which contain the data for seasons 1, 3 and 4; and season 2, respectively. Selectivity pattern is estimated and parameterized as domed shaped. Prior to 2001, the selection pattern is assumed the same as post 1991. No CPUE is available for this fleet.

Tuna Purse Seine Japan Sea (F3)

Catch, size composition (1987-2010), and CPUE (S4) are available for use in the model. Only catch is used in the modeling, while composition and CPUE are not fitted. Selectivity pattern is domed and fixed to an estimate from an earlier run. Prior to 1987, the selection pattern is assumed to be the average of the estimates from 1987-2010.

Tuna Purse Seine Pacific Ocean (F4)

Catch and size composition (1995-2006) is available for use in the modeling. Catch is fitted in the final models but not size composition. Selectivity pattern is domed and fixed at an estimate from an early model run. Prior to 1995, the selection pattern is assumed to be the same as post 1995.

Japan Troll (F5)

Catch, size composition (1993-2010) and 4 CPUEs (S5, S6, S7 and S8) are available for use in the modeling. Catch, composition and East China Sea CPUE (S5) are fitted in the model, while the other CPUEs (S5, S7, and S8) are excluded. Selectivity pattern is estimated and parameterized as domed shaped. Prior to 1993, the selection pattern is assumed the same as post-1993. A catchability (*q*) parameter is estimated and assumed to be linear.

Japan Pole and Line (F6)

Catch and size composition data (1994-2010) are available and fitted in the model. Selectivity pattern is domed and estimated. Selection prior to 1993 is assumed the same as post-1993.

Japan Set Net NOJ weight (F7)

Catch and size composition data (1993-2010, weight) are available for use in the modeling. Catch is included but composition data are not fitted in the final model. Selectivity pattern is domed and fixed at an estimate from an early model run. Selection prior to 1993 is assumed the same as post-1993.

Japan Set Net NOJ length (F8)

Catch and size composition data (1994-2010) are available for use in the modeling. Catch is included but composition data are not fitted in the final model. Selectivity pattern is domed and fixed to an earlier model estimate. Selection prior to 1994 is assumed the same as post-1994.

Japan Set Net OAJ length (F9)

Catch and size composition data (1993-2010) are available for use in the modeling. Catch is included but composition data are not fitted in the final models. F9 is deconstructed into two fleets (F9a and F9b) which contain the data for seasons 1-3; and season 4, respectively. Selectivity pattern is domed and fixed to an earlier model estimate. Selection prior to 1993 is assumed to be the same as post-1993.

Taiwan Longline (F10)

Catch, size composition (1992-2010) and CPUE (1998-2010) is available and used in the modeling. The selection pattern is estimated and assumed asymptotic (Piner et al. 2012). A catchability (*q*) parameter is estimated and assumed linear.

EPO Purse Seine (F11)

Catch, size composition (1952-2010) and 2 CPUE series are available for use in the model. Catches are included but both composition and CPUEs are not used in the model fitting. Selectivity pattern is domed and fixed within 3 time blocks corresponding to the changing composition of the fleet (1952-1982 US PS, 1983-1998 Mixed opportunistic, and 1999-2010 Mexican PS). Fixed selection patterns were the average of year-specific selection patterns estimated in an earlier run and averaged within blocks outside of the model.

F12 EPO Sport (F12)

Catch and size composition (1993-2011) is available for use in the modeling. Catch is included but composition is not fitted in the final models. Selection pattern is assumed to be the same as F11.

Others (F13)

Catch and size composition is available for use in the modeling. Catch is included but composition is excluded. Selection pattern is domed and fixed to average of year-specific selection patterns estimated in earlier model run and averaged outside of the model.

Sensitivity Runs

Table 4 describes a series of sensitivity runs used to evaluate model assumptions and performance. Five sensitivity runs were performed to illustrate the relative influence of the terminal indices (S1, S5 and S9) and size compositions on spawning stock biomass and fits to abundance indices.

Results

Initial modeling indicated that the Japanese CLL (S1) and Taiwanese LL (S9) were indicative of slight, but important differences in population dynamics. Two separate models were created that tuned to S1 (ModS1) and S9 (ModS9) separately. ModS1 did not fit to Japanese LL composition from season 4 (F1b), but fixed the selection pattern to an earlier model estimate. The following results are shown for both models.

Diagnostics

Model diagnostics showed no evidence of a lack of model convergence. The model hessian was positive-definite and the variance-covariance estimated. Examination of the correlations between parameters did not reveal problematic high correlations and random perturbations of starting values and phasing did not result in a better fitting model nor substantially change the estimated spawning biomass trends (Figure 1). A table of total and component likelihoods for models ModS1 and ModS9 are given in Table 5. A profile over fixed levels of R0 indicates that the primary data components strongly influencing the global scaling are the penalty to recruitment deviates (on the low side) and the indices of abundance and the F10 composition (on the high side) (Figure 2). The influence of the recruitment deviate penalty is due to the inflation of recruitment deviates to compensate for the lower R0 in models with low fixed R0. The influence of indices of abundance on scale is desirable and the influence of Taiwanese composition information on global scaling is the result of the strong assumption of asymptotic selection as described by Piner (2012).

Overall fit to size composition was generally reasonable from both fleets with estimated selection patterns with composition data included in the total likelihood and fleets with fixed selection and composition not fitted (Figure 3 and Table 6). However, it should be noted that although the aggregated fit is reasonable, the year-specific prediction from the constrained selection parameterization leads to significant misfit in some years. Some fleet composition data showed patterns that cannot be resolved without adding fine scale time varying process (eg. F4), which could have been included in the model runs but for the sake of parsimony was not included. For these fleets, an “average” selection pattern was used and the inevitable misfit of composition excluded from the likelihood function. Fits to the indices for both ModS1 and ModS9 were generally very good, with RMSE ~0.2 for all indices and major trends in the indices being well represented by the models (Figure 4 and Table 6).

Model Results

Time varying selection patterns did not always indicate a major shift in the proportions at length selected in the recent period (Figure 5). The Japanese longline in seasons 1-3 did not show a major shift in selection (F1a) after 1993 but season 4 (F1b) did shift to larger fish making it more similar to the Taiwanese longline fleet (F10). Similarly the EPO purse seine fleet selection (F11), which was estimated from an earlier model run prior to fixing, also indicated a shift to larger fish as fleet composition changed.

Estimated biomass trends were similar from both models ModS1 and ModS9 for the majority of the dynamic period (Figure 6). However the Taiwanese longline CPUE (ModS9) indicated that the stock is at a larger current spawning biomass level than the Japanese Coastal LL index (ModS1). Both models indicate that current spawning biomass is at low levels relative to the virgin state.

Estimated recruitment has varied without strong long-term trends (Figure 7). Although it appears that the scale of recruitment may have increased in the last 15 years, a ranking of recruitment levels indicates that recruitment is at approximately the same levels as in the 1960s or 1970s.

Estimated trends in F showed a general pattern of increasing F on age 0 and adults, and relatively high mortality rates on juveniles (Figure 8).

Sensitivity Results

Sensitivity run A, which is not fitted to any size composition data, show the best fit possible when both S1 and S9 are fitted to the model (Figure 9). In contrast, sensitivity run B, which includes the size composition data used in the base case models (ModS1 and ModS9), indicate that when both S1 and S9 are included, fit to S9 is relatively poor (Figure 9). Spawning biomass estimates from the five sensitivity runs are shown in Figure 10. Interestingly, sensitivity run D, which does not fit any terminal abundance index, shows a similar spawning biomass trend to the base case models. This suggests that the size composition data in the base case models are relatively consistent with the abundance indices in the base case models, albeit with higher uncertainty. Sensitivity run E, which does not fit the last 3 years of the Japanese coastal longline (S1) index, shows that excluding the last 3 years of the S1 index slightly improves the fit to Taiwanese longline index (S9) but not substantially so (Figure 11).

Discussion

The models presented in this paper satisfy the criteria listed in the introduction for an acceptable model result. The fit to the primary tuning indices of abundance is good. The general good agreement between model estimates and observed CPUE is a result of efforts to minimize misfit to composition or its addition to the likelihood function.

We note that two troll gear indices of abundance (primarily age 0) were used in the May workshop. However initial analysis showed that S5 and S6 were inconsistent with each other and S5 was much more consistent with other data. Thus in this model S6 was removed and only S5 was included. An alternative model could be constructed with the S6 index, but we judged it as likely to be unrepresentative of the age 0 abundance and therefore unnecessary to be included as a separate run.

This paper focused on the technical question regarding how to use size data that are not well described by a single time-invariant process of selection. For some fisheries, we do not have the ability to introduce the appropriate process to fit the data such as F11 (movement) or F13 (multiple gears with potential multiple selection processes). Other fleets require year-specific selection (F3) that in practice is difficult to estimate. One approach to deal with this situation is to down-weight the data such that the misfit will not contribute much to the total likelihood. However, it is difficult to choose the appropriate weight that insures appropriate estimates of average selection but does not allow the residual misfit of composition to an under-parameterized model to distract the model from matching the indices. Another approach is to live with the residual misfit, as long as it does not cause significant deterioration to the fit to the indices. However the residual misfit implies that the composition data are not good representations of the capture process, which can be interpreted as not believing the data. In this paper, we have chosen a third approach for some fleets that uses the data with full weight to get the appropriate shape of the selection pattern, but then removes the residual misfit of the composition by not allowing it to contribute to the likelihood function. It can be argued that this approach, especially for fleet with hypothesized time varying process makes the most use of the data. Most importantly, it is consistent with the approach of preferential fitting to indices of abundance (Francis 2011).

The sensitivity analyses indicated substantial agreement among CPUE and size data sources in the base models regarding the scale and relative biomass levels. The reduced composition information (sensitivity run D) used in our base case model is consistent with the Japanese coastal longline index (S1) in that it indicates a decline in spawning biomass since 2005. The full size composition model (run C) is less so, indicating that given a constant selection, the spawning biomass has been relatively constant since 1990. The sensitivity runs (C and D) that fitted to only size composition data for the most recent period show large differences in their uncertainty in the spawning biomass estimates in the recent period. Sensitivity run D (reduced size composition data set) show a pattern of very high uncertainty but sensitivity run C (full composition data set) show very low uncertainty. We regard the low uncertainty to be a reflection of both increased amount of information (more fleets composition included in the total likelihood function) and also due to constrained selection parameterization imposing an overly rigid model structure.

Several concerns remain in our modeling, including poor estimates to the F4 size composition and therefore uncertain selection pattern. This is a concern as the fleet was a very substantial proportion of the total catch prior to 1990. We have little information on the size composition of the catch in the early years and data in the later years is noisy and does not appear to reflect a well behaved selection process. Without addition data analyses this issue cannot be addressed. Additional sensitivity analyses needs to be done to the effects on biomass trends of mis-specifying the selection process especially in the early years of the model. We also note that the selectivity pattern to F1b in the ModS1 model shows a high correlation (~0.95) for the ascending limb and peak parameters. However, we do not consider this a problematic level of correlation for these parameters. The implementation of the bias adjustment for recruitment estimates is also an area for further consideration. The current bias adjustment begins in 1946 to correspond with the start of the main recruitment period but could be moved earlier so that the level of bias adjustment corresponds better to the amount of adjustment indicated by the model.

Our goals in this paper are to explore methods of using complicated data sets that provide acceptable fits to important data series and as such, was not focused on commenting on the estimated stock levels. However, it is important to note that reasonable matching of the main tuning indices results in a pessimistic model with very low terminal spawning biomasses and very high levels of depletion. Results indicate that catch of PBF has shown a long term shift to younger ages and that population levels have undergone a long-term decline. The authors note that the ratios of estimated recent spawning biomass levels to the estimated unfished spawning biomass are extremely low, likely too low, compared to other highly migratory species, which may indicate some unknown model mis-specification, data bias or both. Nevertheless, all models indicate a population that is at low relative levels with an exploitation pattern that is far from optimal.

Literature Cited

Francis, R.I.C.C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences* 68, 1124–1138.

ISC (2012). Report of the billfish working group workshop. Annex 7 in Report of the twelfth meeting of the International Scientific Committee for tuna and tuna-like species in the north Pacific Ocean. 18-23 July, 2012. Sapporo, Hokkaido Japan.

Ichinokawa, M. and Takeuchi, Y. 2012. Standardized CPUE of north Pacific bluefin tuna caught by Japanese coastal longliners: updates until 2011. ISC12-1/PBFWG/8

Piner, K. (2012) Selection of an asymptotic selectivity pattern. ISC12-2/PBFWG/15

Piner, K., H.H., Lee, I., Taylor, A. Aires-da-Silva, and S., Teo (2011). Introduction of a spatially structured model of Pacific bluefin tuna. ISC/11-1/PBFWG/06

Polovina, J.J. (1996). Decadal variation in the trans-Pacific migration of northern bluefin tuna (*Thunnus thynnus*) coherent with climate induced changes in prey abundance. *Fisheries Oceanography* 5(2):114-119.

Table 1. Key life history used in the modeling. In the "estimate" column, values represent assumed parameter value and "estimated" indicates estimated in the model.

parameter	estimate	Time period
	S-R	
Ln (R0)	estimated	
h	0.999	
SigmaR	0.6	
R1 offset (equib rec)	estimated	
deviations	estimated	1952-2009
	Growth and Mortality	
Von Bertalanffy		
Length (age 0)	21.5	
Length (age 3)	109.194	
K	0.157474	
CV young (age 0)	estimated	
CV old (age 3)	0.08	
$W(kg)=aL(cm)^b$	$Wt (kg) = 1.7117E-5 L(cm)^{3.0382}$	
$Natural Mortality^{-yr}$		
Age 0, 1, 2+	1.6, 0.386, 0.25	
Maturity (%)		
Age 0-2,3,4,5+	0, 20, 50, 100	

Table 2. Data available for modeling. Not all data are fitted in the model. See "Methods" section for detailed description of data used. Fishery objects abbreviations (e.g., F1 or S4) are given in this table.

Fleet	Catch	Size comp	CPUE	Comments
JPN LL (F1)	1952-2010	1952-1968, 1994-2010	1952-1973 (S2), 1974-1992 (S3), 1993-2010 (S1)	Split into 2 fleets: season 1-3 and season 4
Pelagic PS (F2)	1952-2010	2001-2010		Split into 2 fleets: season 1,3,4 and season 2
PS Japan Sea (F3)	1952-2010	1987-2010	1987-2010 (S4)	
PS Pacific O. (F4)	1952-2010	1995-2006		Many observations deleted in May
JPN Troll (F5)	1952-2010	1993-2010	1980-2010 (S5), 1994-2010 (S6), 1981-2010 (S7), 1994-2010 (S8)	Only use S5
JPN Pole Line (F6)	1952-2010	1994-2010		
JPN Set Net NOJ weight (F7)	1952-2010	1993-2010		
JPN Set Net NOJ length (F8)	1952-2010	1994-2010		
Set Net OAJ length (F9)	1952-2010	1993-2010		Split into 2 fleets: season 1-3 and season 4
Taiwan LL (F10)	1952-2010	1992-2010	1998-2010 (S9)	
EPO PS (F11)	1952-2010	1952-2010	1960-1982 (S10), 1999-2010 (S11)	
EPO Sport (F12)	1952-2010	1993-2011		
Others (F13)	1952-2010	1994-2010		

Table 3. Selectivity pattern descriptions by SS object. Designation column indicates if the object is a fleet (F#) with catch or an index of abundance (S#). Mirror column indicates if the object assumes the same selection pattern as another object. Number of parameters indicates the number of parameters estimated for the pattern in the final model. Fixed indicates that parameters estimated in early model run and then not estimated in final model run. If number of parameters is 0 this indicates that the object borrowed the pattern.

Designation	Shape	Parameterization	mirror	Time varying	Number parameters
F1	Domed	Double normal	N/A	1952-1992, 1993-2010	16
F2	Domed	Double normal	N/A		6
F3	Domed	Double normal	N/A		fixed
F4	Domed	Double normal	N/A		fixed
F5	Domed	Double normal	N/A		3
F6	Domed	Double normal	N/A		3
F7	Domed	Double normal	N/A		fixed
F8	Domed	Double normal	N/A		fixed
F9	Domed	Double normal	N/A		fixed
F10	Asymptotic	logistic	N/A		2
F11	Domed	Double normal	N/A	1952-1982,1983-1997, 1998-2010	fixed
F12	Domed		F11		0
F13	Domed	Double normal	N/A		fixed
S1,S2,S3	Domed		F1		0
S4	Domed		F3		0
S5,S6,S7,S8	Domed		F5		0
S9	Asymptotic		F10		0
S10,S11	Domed		F11		0

Table 4. List of sensitivity runs.

	Indices Fitted	Size Compositions Fitted	Comments
A	S1, S2, S3, S5, S9	None	All selectivities fixed
B	S1, S2, S3, S5, S9	F1s123, F1s4, F2s134, F2s2, F5, F6, and F10	
C	S2, S3	All size compositions except F11, F12, and F13	
D	S2, S3	F1s123, F1s4, F2s134, F2s2, F5, F6, and F10	
E	S1, S2, S3, S5, S9	F1s123, F1s4, F2s134, F2s2, F5, F6, and F10	Last 3 years of S1 not fitted

Table 5. Negative log-likelihoods for model components in models: ModS1 (model tuned to S1) and ModS9 (model tuned to S9). Bold numbers indicate components that were not included in the total likelihood for that specific model but are included here for comparison purposes.

Likelihood Component	ModS1	ModS9
Total	1114.28	1040.05
Catch	0.00044256	0.0003471
Recruitment	1.29584	-1.3446
Surveys		
S1_JpCLL	-19.0949	**11.7726**
S2_JpDWLLto74	-22.397	-24.5844
S3_JpDWLLfrom75	-26.4293	-26.3501
S5_JpnTrollChinaSea	-38.7412	-38.6837
S9_TWLL	**21.1094**	-11.8485
Size Composition		
F1s123_JLL	195.514	199.202
F1s4_JLL	83.351	**103.791**
F2s134_SPelPS	172.332	170.492
F2s2_SPelPS	58.3343	58.4289
F5_JpnTroll	379.153	378.938
F6_JpnPL	285.259	286.364
F10_TWLL	45.6901	49.4225

Table 6. Model fit diagnostics for components in models: ModS1 (model tuned to S1) and ModS9 (model tuned to S9).

		ModS1		ModS9	
Surveys	**N**	**Q**	**rmse**	**Q**	**Rmse**
S1	18	0.0130253	0.209496		
S2	22	4.04E-05	0.217512	4.04E-05	0.198386
S3	19	5.39E-05	0.132188	5.39E-05	0.133444
S5	31	0.000179846	0.16964	0.000175933	0.170077
S9	13			0.0136173	0.236307
Size Composition Fitted	**N**	**Mean_effN**	**Mean_InputN**	**Mean_effN**	**Mean_InputN**
F1a	50	35.8221	11.6937	33.7783	11.6937
F1b	32	61.7578	12.8052		
F2a	28	17.1182	12.1056	17.183	12.1056
F2b	9	23.1762	12.1056	22.6891	12.1056
F5	66	28.7683	12.1056	28.6934	12.1056
F6	32	14.085	12.1056	13.9223	12.1056
F10	20	77.5603	12.1056	82.9006	12.1056
Size Composition Not Fit	**N**	**Mean_effN**	**Mean_InputN**	**Mean_effN**	**Mean_InputN**
F1b	32			52.8309	12.8052
F3	23	12.9442	8.54487	16.0608	8.54487
F4	11	42.9772	9.81307	40.6434	9.81307
F7	53	18.8408	12.1056	20.0213	12.1056
F8a	8	20.4601	10.543	21.3378	10.543
F8b (mirrored to F8a)	16	22.6135	12.8868	22.7959	12.8868
F9a	52	29.2008	12.1056	29.4812	12.1056
F9b	18	34.4447	12.1056	35.3822	12.1056
F11	122	10.8809	7.91803	11.024	7.91803
F13	38	16.1416	12.1056	18.9131	12.1056

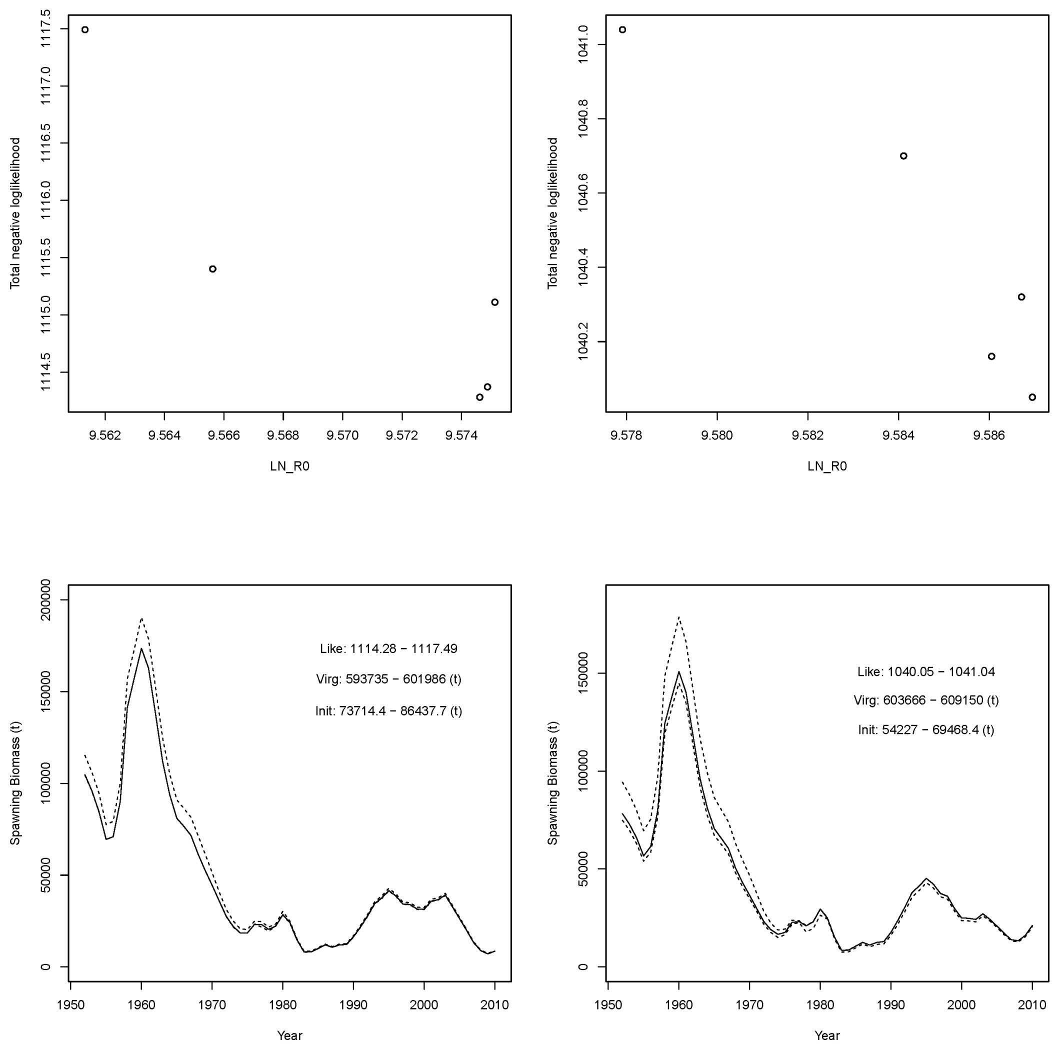
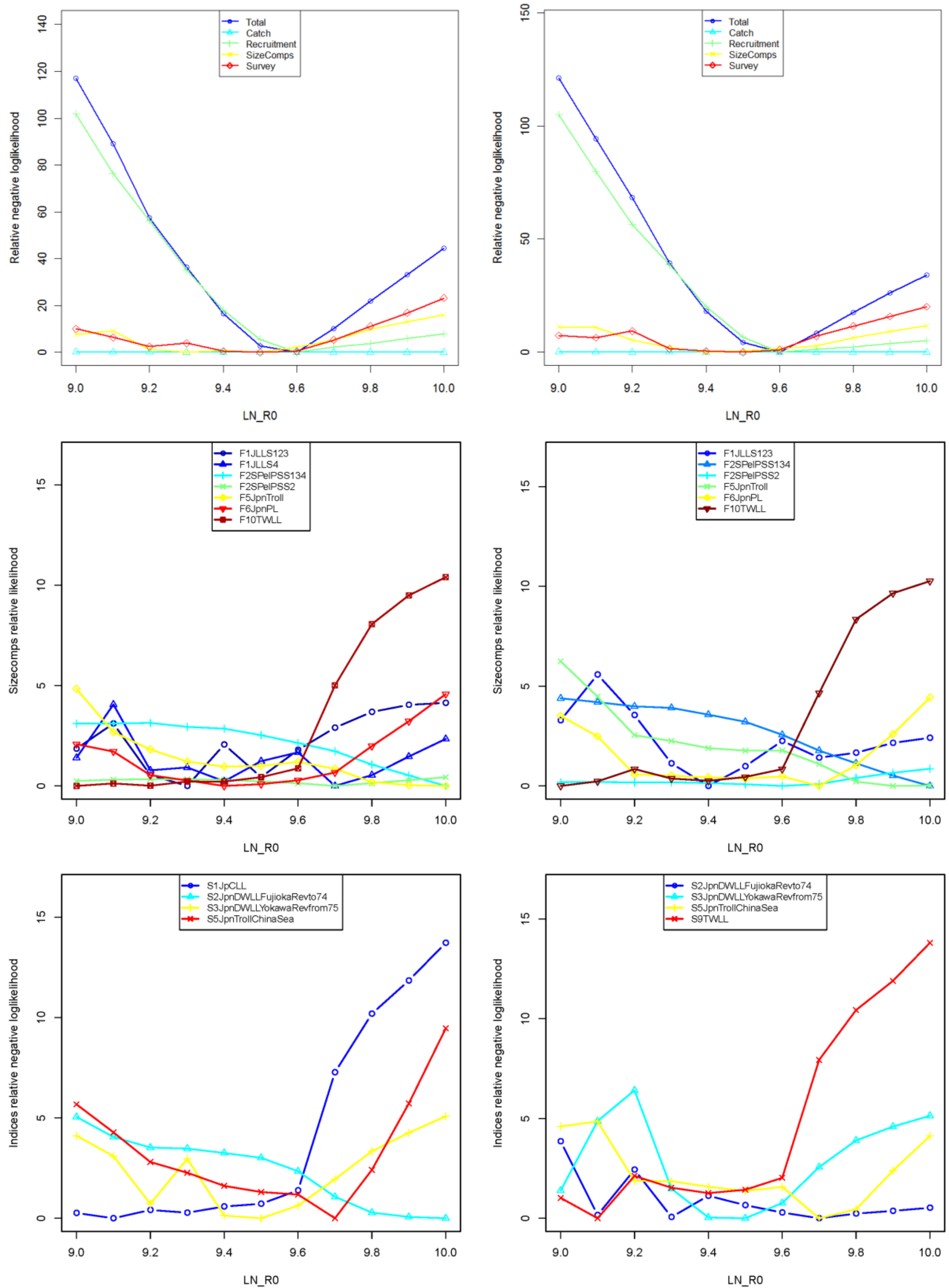
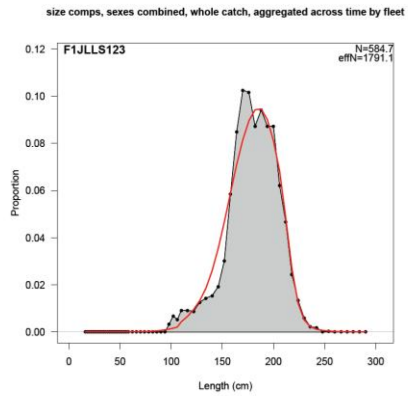


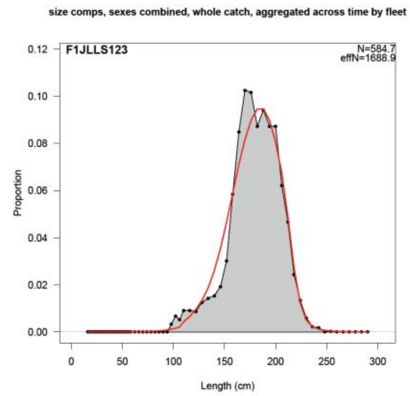
Figure 1. Results of random perturbation of starting values and phases (jittering) for models a) ModS1 (left column) and b) ModS9 (right column). Best fitting model lowest on plot and the symbol represents multiple runs (>10).

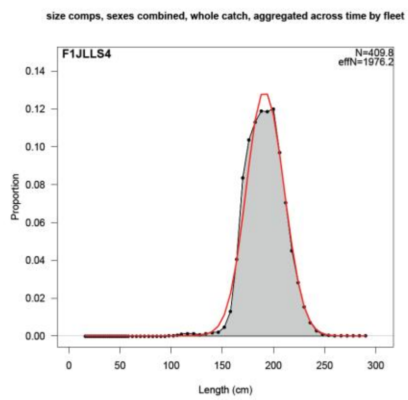


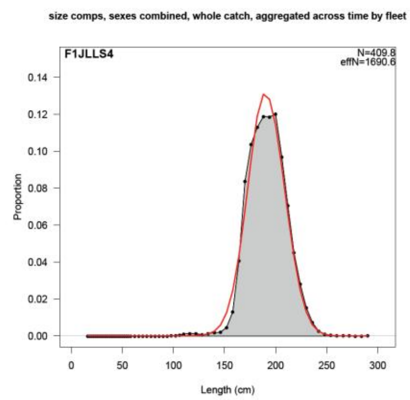
ModS1 **ModS9**

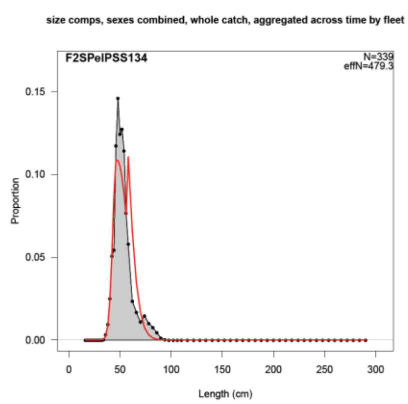
Figure 2. Plot of likelihood profiles for models ModS1 (left column) and ModS9 (right column). Plots are: 1) total and major component likelihoods, b) size composition likelihood by component, and c) index likelihood by component.

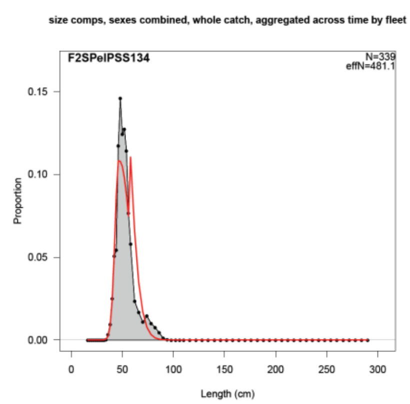

size comps, sexes combined, whole catch, aggregated across time by fleet
F1JLLS123
N=584.7
effN=1791.1
Proportion
Length (cm)

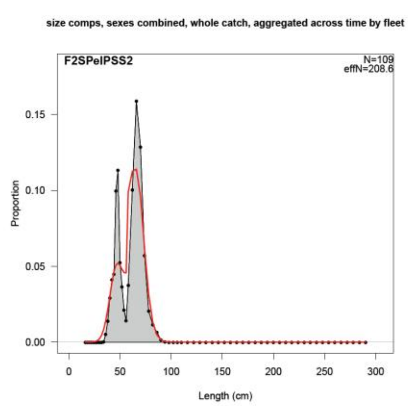

size comps, sexes combined, whole catch, aggregated across time by fleet
F1JLLS123
N=584.7
effN=1688.9
Proportion
Length (cm)

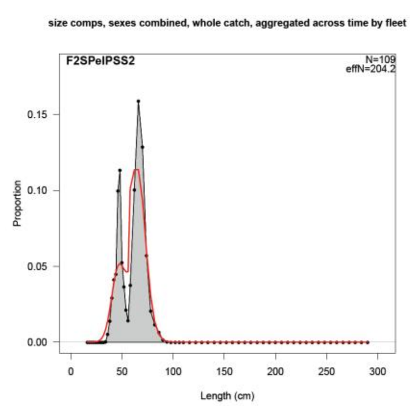

size comps, sexes combined, whole catch, aggregated across time by fleet
F1JLLS4
N=409.8
effN=1976.2
Proportion
Length (cm)

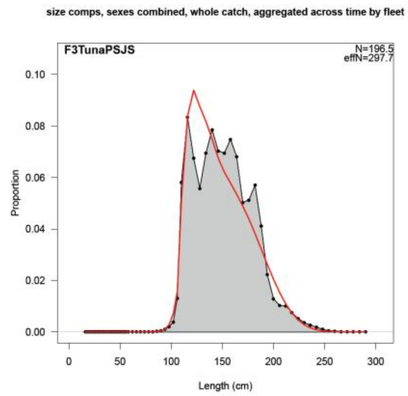

size comps, sexes combined, whole catch, aggregated across time by fleet
F1JLLS4
N=409.8
effN=1690.6
Proportion
Length (cm)

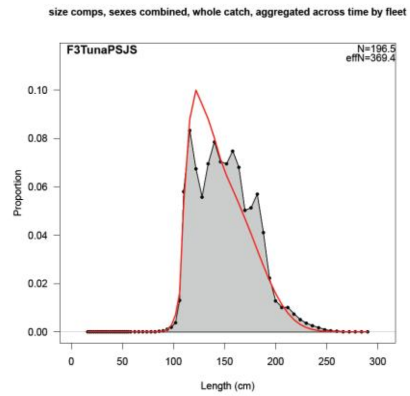

size comps, sexes combined, whole catch, aggregated across time by fleet
F2SPelPSS134
N=339
effN=479.3
Proportion
Length (cm)

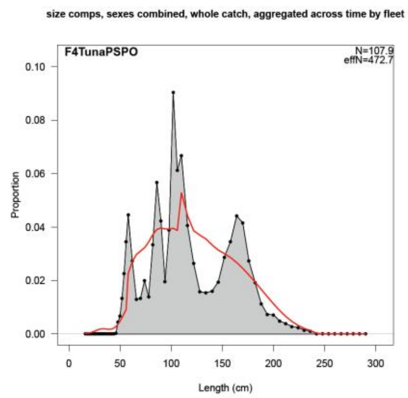

size comps, sexes combined, whole catch, aggregated across time by fleet
F2SPelPSS134
N=339
effN=481.1
Proportion
Length (cm)

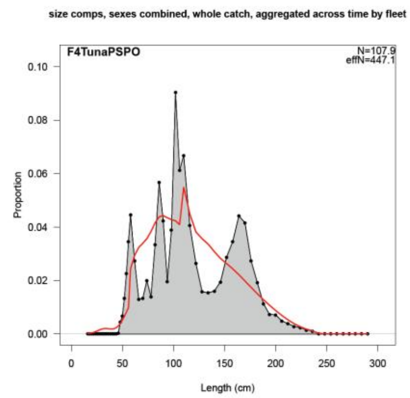

size comps, sexes combined, whole catch, aggregated across time by fleet
F2SPelPSS2
N=109
effN=208.6
Proportion
Length (cm)

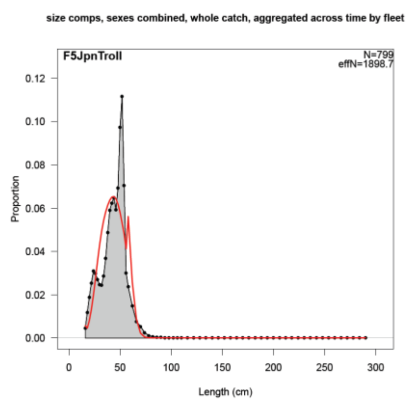

size comps, sexes combined, whole catch, aggregated across time by fleet
F2SPelPSS2
N=109
effN=204.2
Proportion
Length (cm)

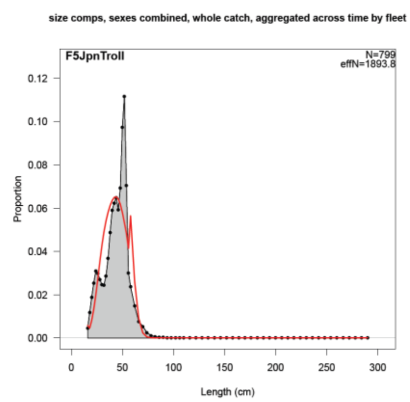

size comps, sexes combined, whole catch, aggregated across time by fleet
F3TunaPSJS
N=196.5
effN=297.7
Proportion
Length (cm)

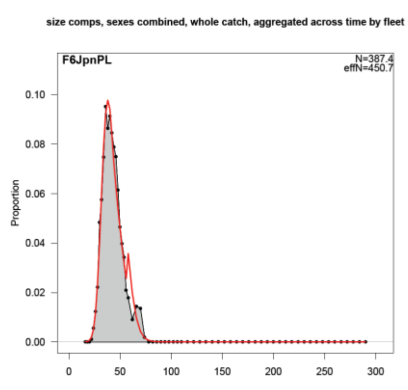

size comps, sexes combined, whole catch, aggregated across time by fleet
F3TunaPSJS
N=196.5
effN=369.4
Proportion
Length (cm)

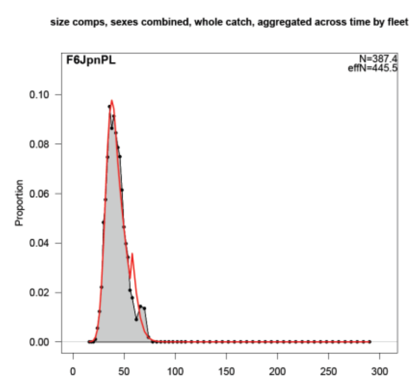

size comps, sexes combined, whole catch, aggregated across time by fleet
F4TunaPSPO
N=107.9
effN=472.7
Proportion
Length (cm)

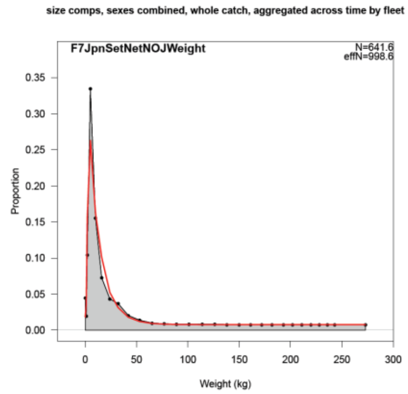

size comps, sexes combined, whole catch, aggregated across time by fleet
F4TunaPSPO
N=107.9
effN=447.1
Proportion
Length (cm)

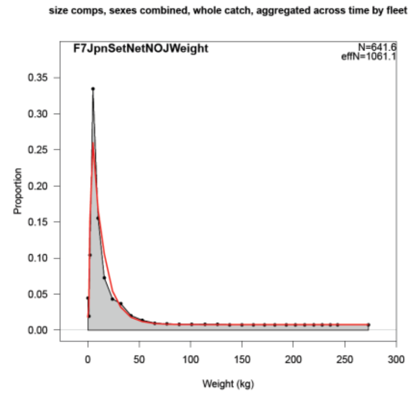

size comps, sexes combined, whole catch, aggregated across time by fleet
F5JpnTroll
N=799
effN=1898.7
Proportion
Length (cm)

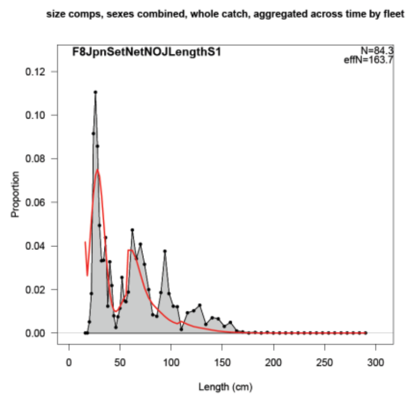

size comps, sexes combined, whole catch, aggregated across time by fleet
F5JpnTroll
N=799
effN=1893.8
Proportion
Length (cm)

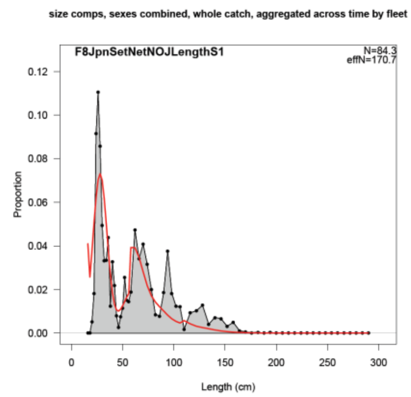

size comps, sexes combined, whole catch, aggregated across time by fleet
F6JpnPL
N=387.4
effN=450.7
Proportion
Length (cm)

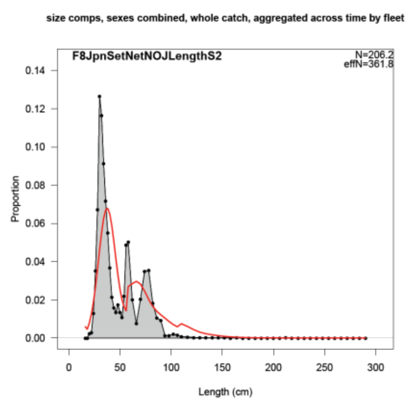

size comps, sexes combined, whole catch, aggregated across time by fleet
F6JpnPL
N=387.4
effN=445.5
Proportion
Length (cm)

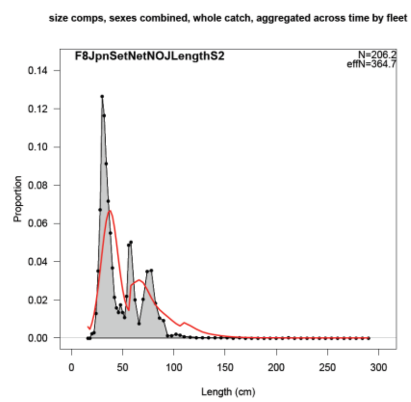

size comps, sexes combined, whole catch, aggregated across time by fleet
F7JpnSetNetNOJWeight
N=641.6
effN=998.6
Proportion
Weight (kg)

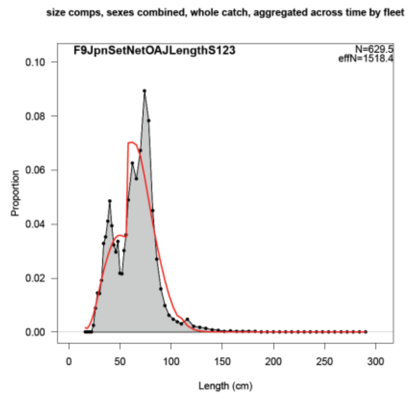

size comps, sexes combined, whole catch, aggregated across time by fleet
F7JpnSetNetNOJWeight
N=641.6
effN=1061.1
Proportion
Weight (kg)

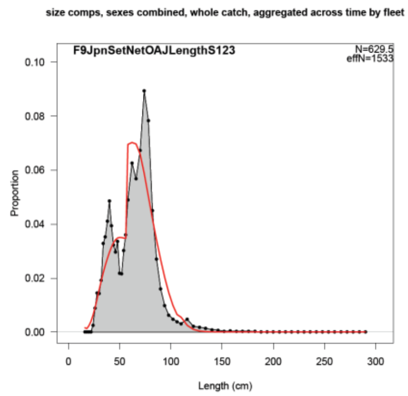

size comps, sexes combined, whole catch, aggregated across time by fleet
F8JpnSetNetNOJLengthS1
N=84.3
effN=163.7
Proportion
Length (cm)

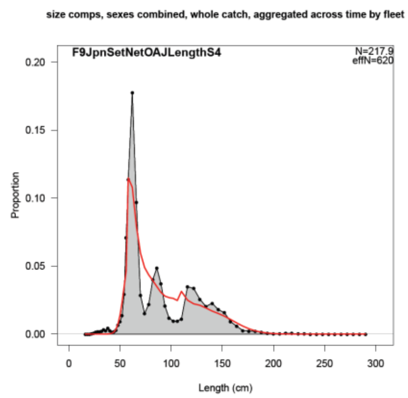

size comps, sexes combined, whole catch, aggregated across time by fleet
F8JpnSetNetNOJLengthS1
N=84.3
effN=170.7
Proportion
Length (cm)

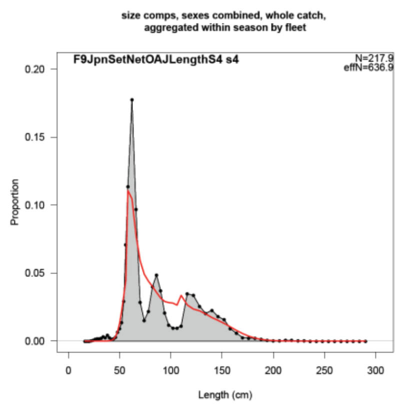

size comps, sexes combined, whole catch, aggregated across time by fleet
F8JpnSetNetNOJLengthS2
N=206.2
effN=361.8
Proportion
Length (cm)

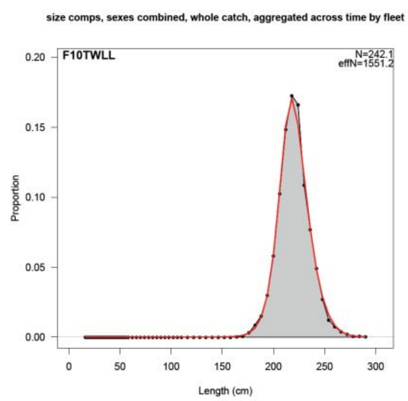

size comps, sexes combined, whole catch, aggregated across time by fleet
F8JpnSetNetNOJLengthS2
N=206.2
effN=364.7
Proportion
Length (cm)

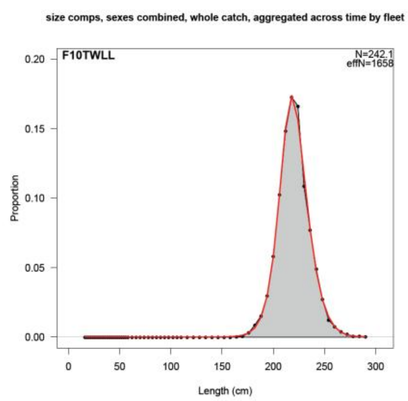

size comps, sexes combined, whole catch, aggregated across time by fleet
F9JpnSetNetOAJLengthS123
N=629.5
effN=1518.4
Proportion
Length (cm)


size comps, sexes combined, whole catch, aggregated across time by fleet
F9JpnSetNetOAJLengthS123
N=629.5
effN=1533
Proportion
Length (cm)


size comps, sexes combined, whole catch, aggregated across time by fleet
F9JpnSetNetOAJLengthS4
N=217.9
effN=620
Proportion
Length (cm)


size comps, sexes combined, whole catch,
aggregated within season by fleet
F9JpnSetNetOAJLengthS4 s4
N=217.9
effN=636.9
Proportion
Length (cm)


size comps, sexes combined, whole catch, aggregated across time by fleet
F10TWLL
N=242.1
effN=1551.2
Proportion
Length (cm)


size comps, sexes combined, whole catch, aggregated across time by fleet
F10TWLL
N=242.1
effN=1658
Proportion
Length (cm)

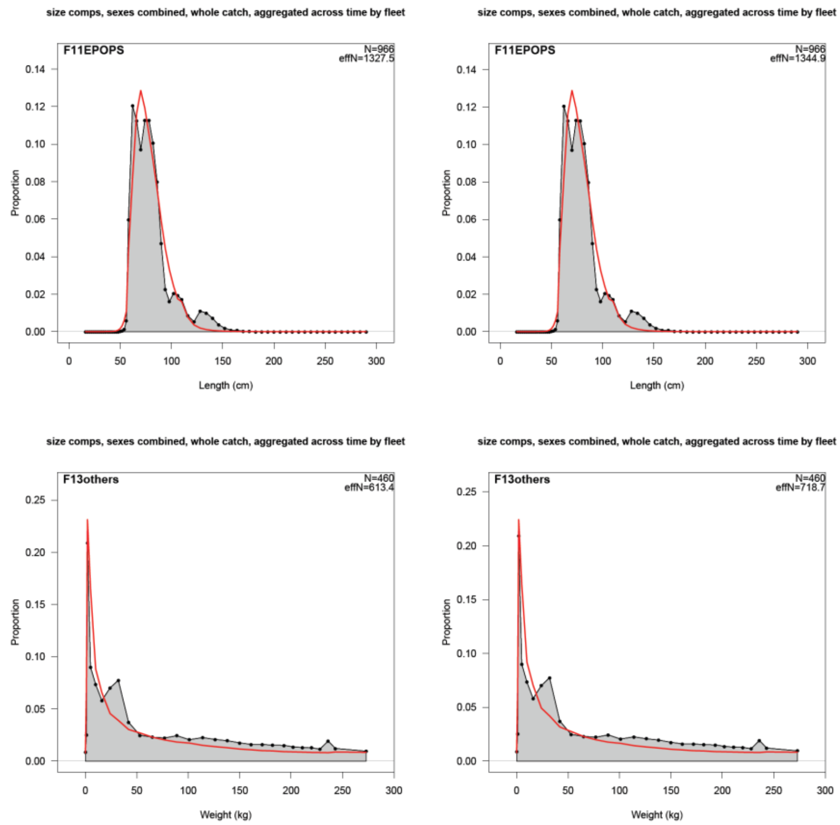


Figure 3. Plot of observed and expected proportions at length from ModS1 (first column) and ModS9 (second column). Solid areas are observed proportion at size and the red line is the expected proportions summed across all seasons and years. Summed inputted sample size is denoted by (N) and model estimated sample size (effN).

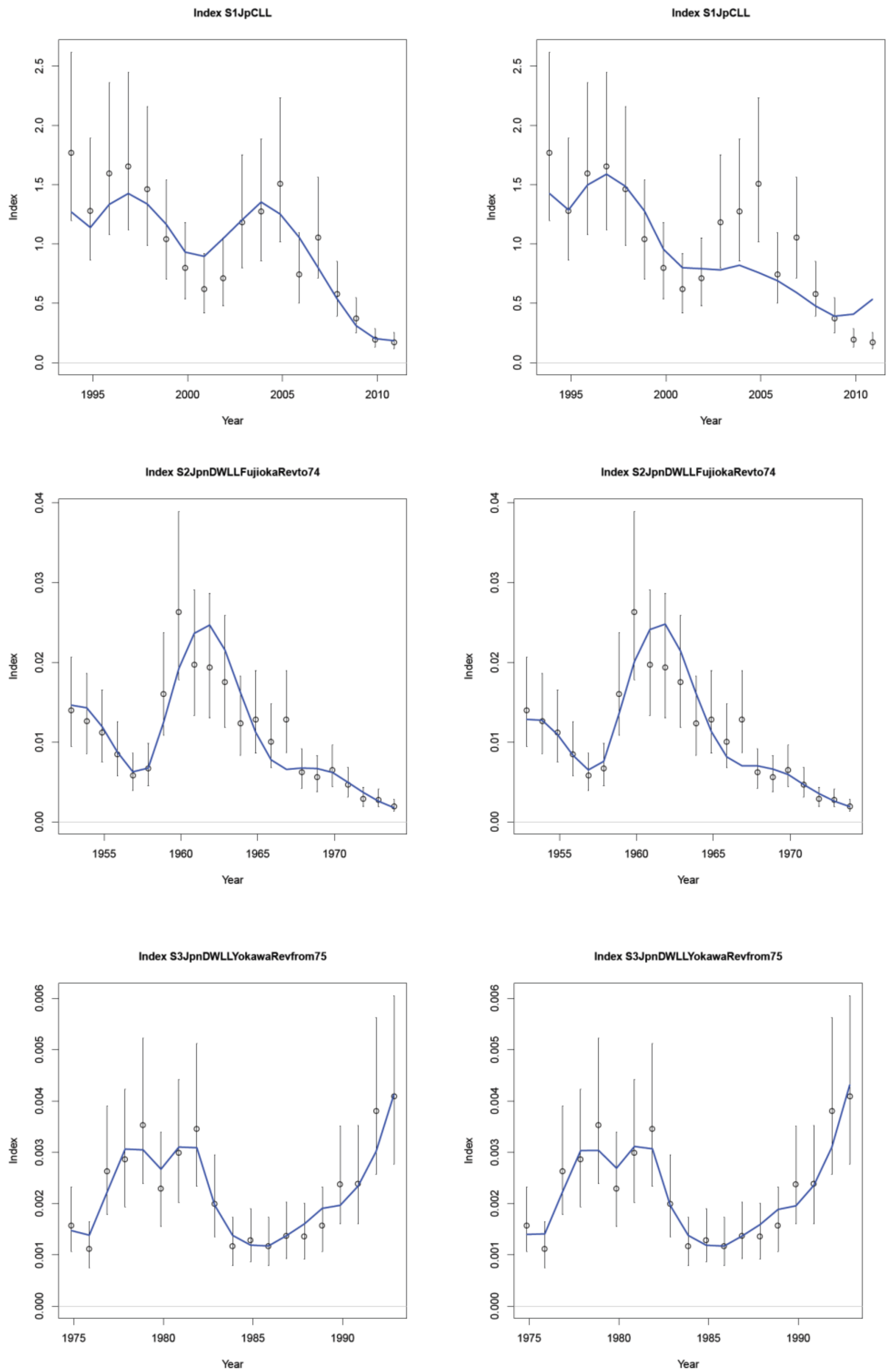


Figure 4. Fit to the primary indices of abundance for models ModS1 (first column) and ModS9 (second column). Plots are for a) Japanese CLL (S1), b) Japanese DWLL early (S2), c) Japanese DWLL middle (S3), d) Japanese Troll (S5) and e) Taiwanese CLL (S9). Note for ModS1 the Taiwanese CLL is not included in the total likelihood and for ModS9 the Japanese CLL in not included, but results are plotted for comparison.

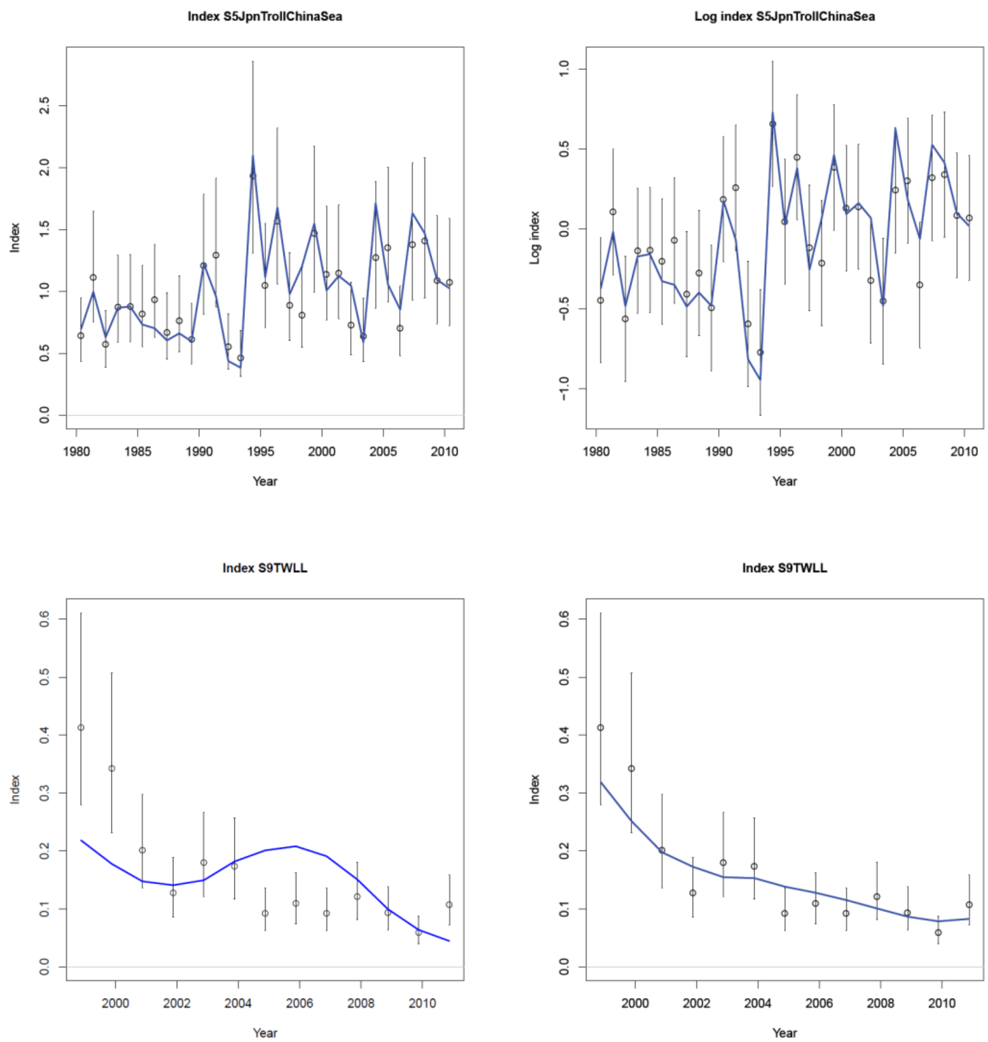


Figure 4 continued. Fit to the primary indices of abundance for models ModS1 (first column) and ModS9 (second column). Plots are for a) Japanese CLL (S1), b) Japanese DWLL early (S2), c) Japanese DWLL middle (S3), d) Japanese Troll (S5) and e) Taiwanese CLL (S9). Note for ModS1 the Taiwanese CLL is not included in the total likelihood and for ModS9 the Japanese CLL in not included, but results are plotted for comparison.

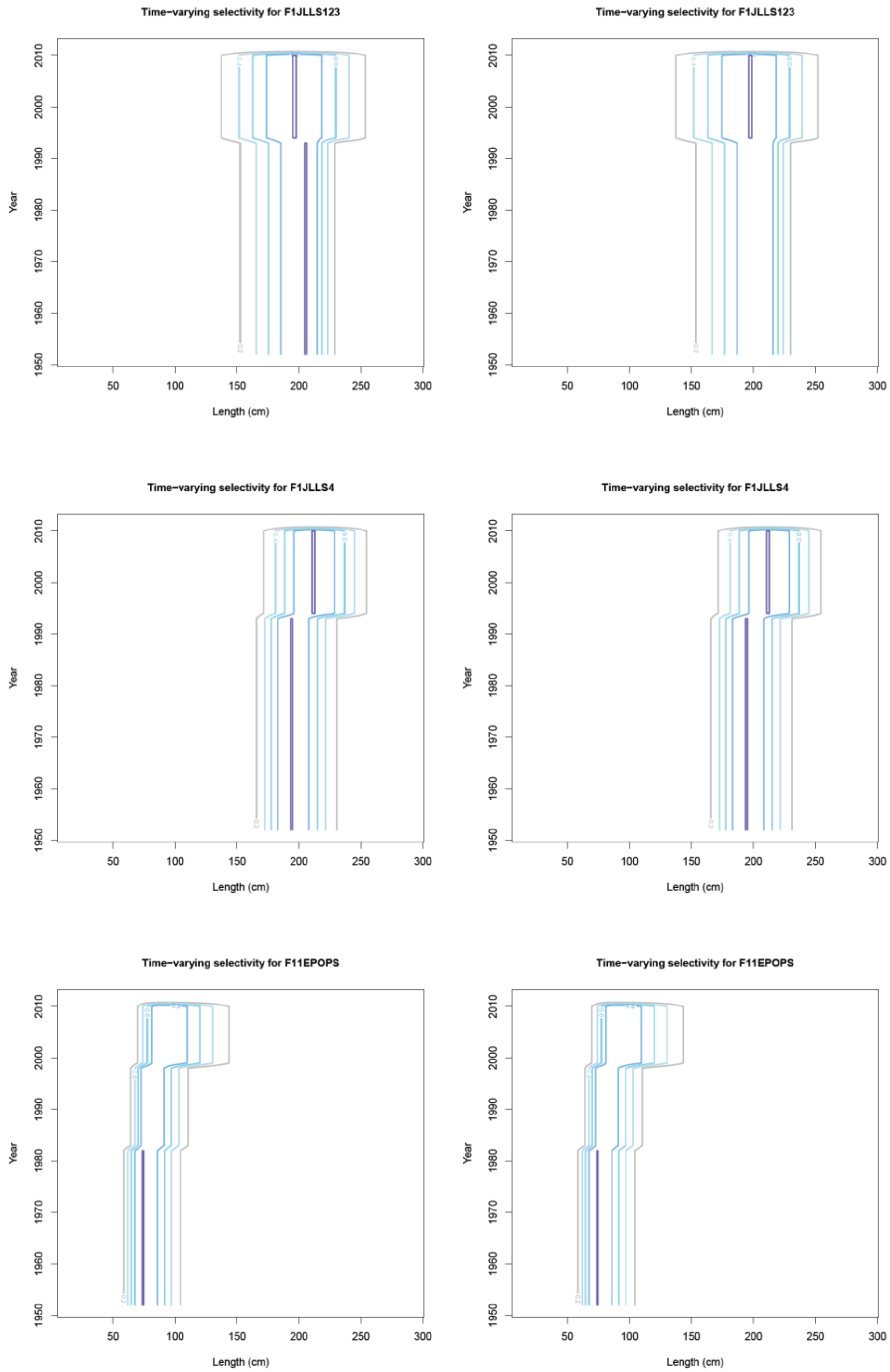


Figure 5. Plots of time varying selections patterns estimated or fixed in the models ModS1 (first column) and ModS9 (second column). Contours represent the selection probalility for fleets: a) Japanese longline seasons 1-3 (F1a), b) Japanese longline season 4 (F1b) and c) EPO purse seine (F11).

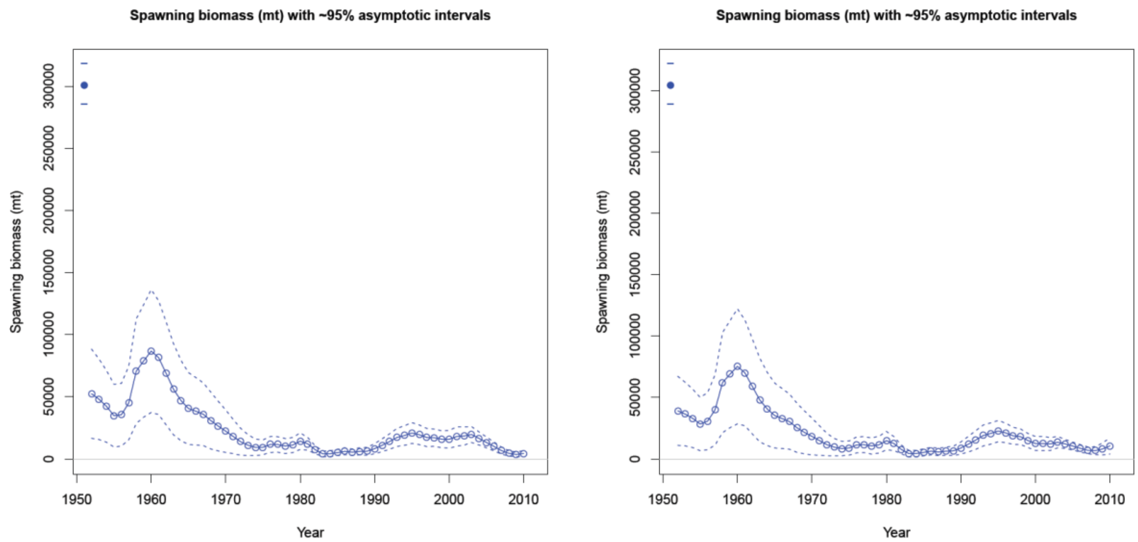


Figure 6. Spawning biomass estimates from ModS1 (first column) and ModS9 (second column). Dotted lines represent 95% CI based on the asymptotic error.

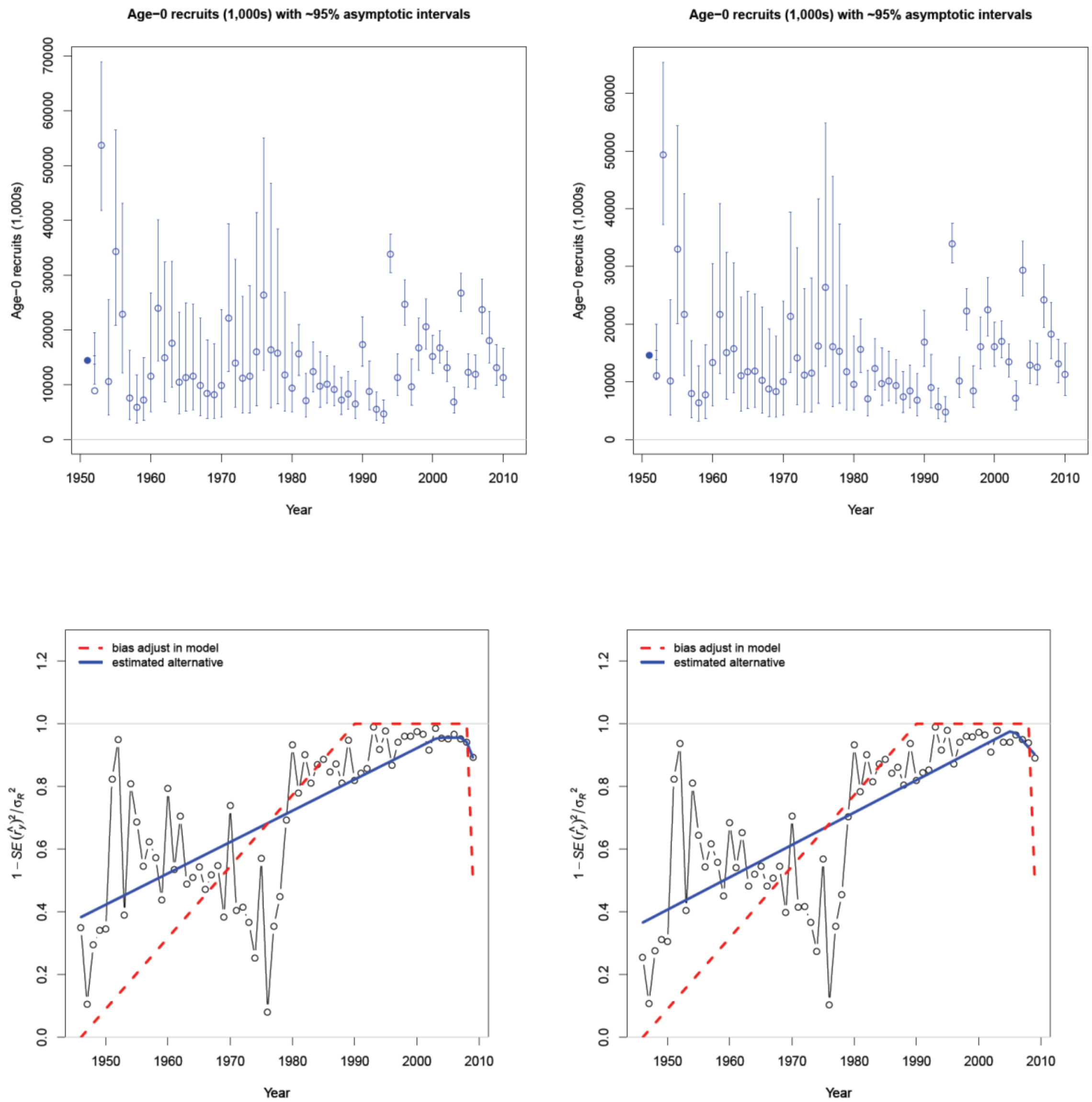


Figure 7. Plot of a) age 0 recruitment estimates from and b) the recruitment bias adjustment used in ModS1 (first column) and ModS9 (second column). In a) the error bars represent 95% CI based on the asymptotic error. In b) the solid line is the phase in of the log bias adjustment and the dashed line is a estimated linear trend in the predicted bias adjustment.

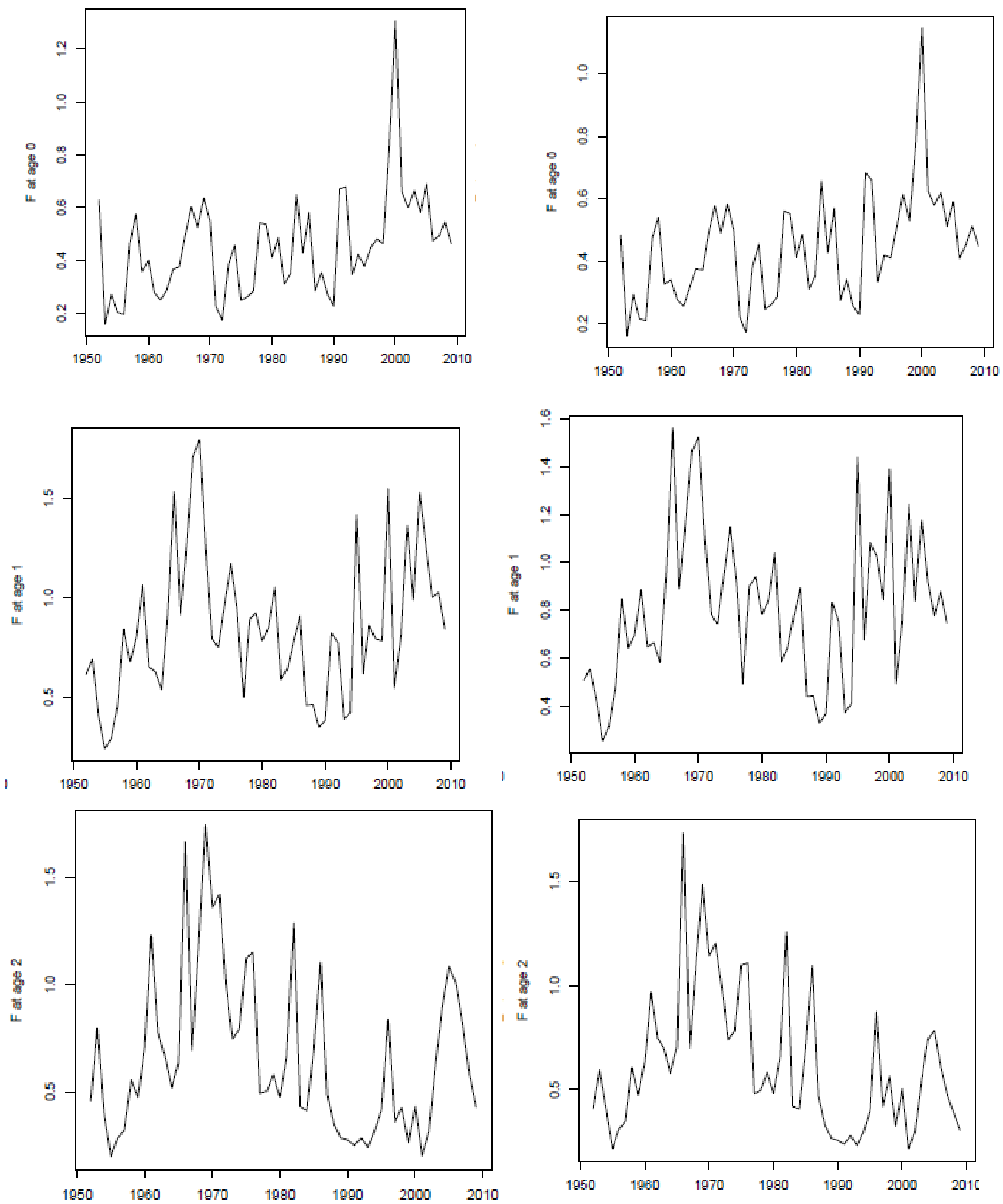


Figure 8. Estimated F at-age from models ModS1 (first column) and ModS9 (second column). F is given for a) age 0, b) age 1, and c) age 2.

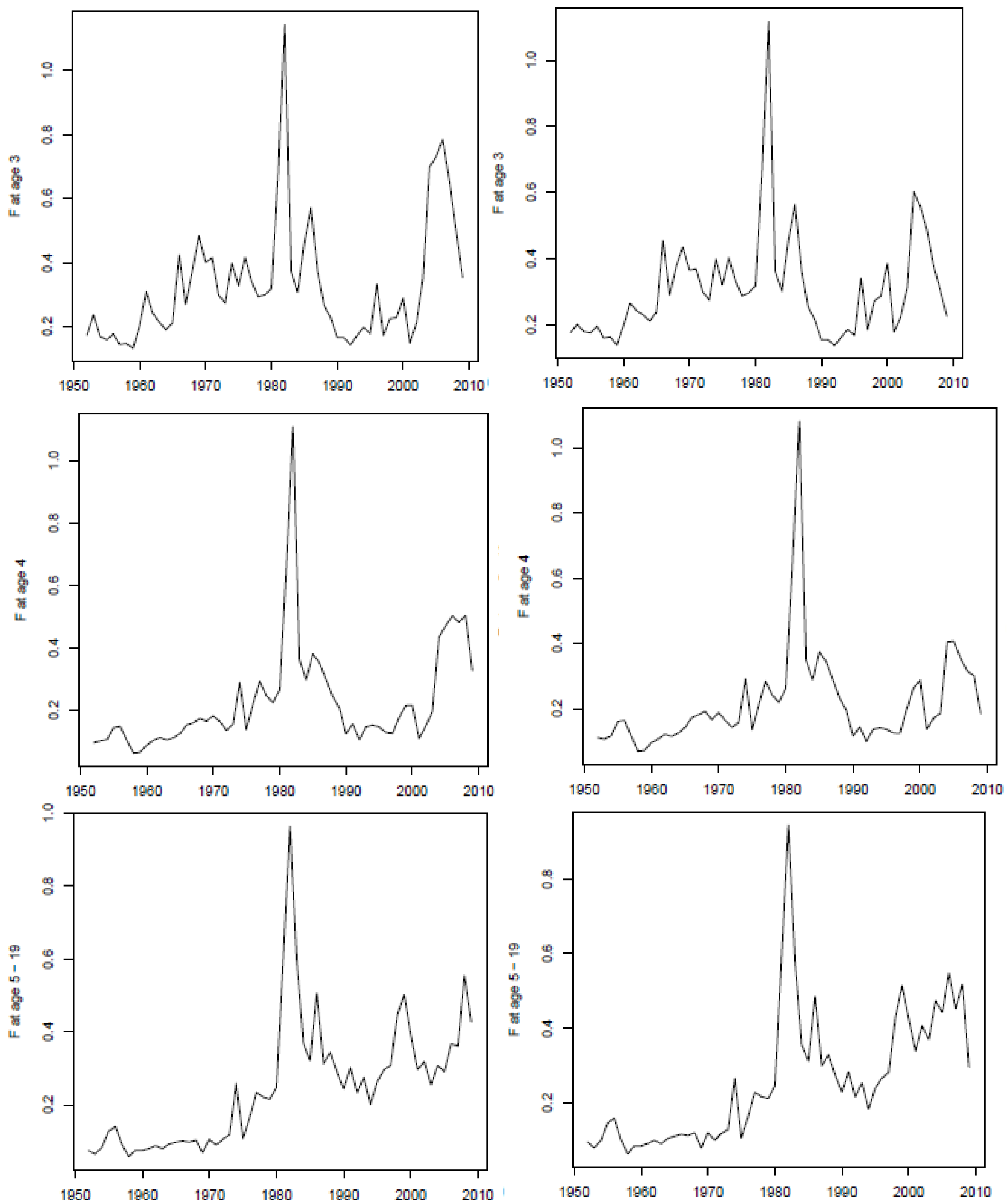


Figure 8 continued. Estimated F at-age from models ModS1 (first column) and ModS9 (second column). F is given for age 3, e) age 4, and f) age 5+.

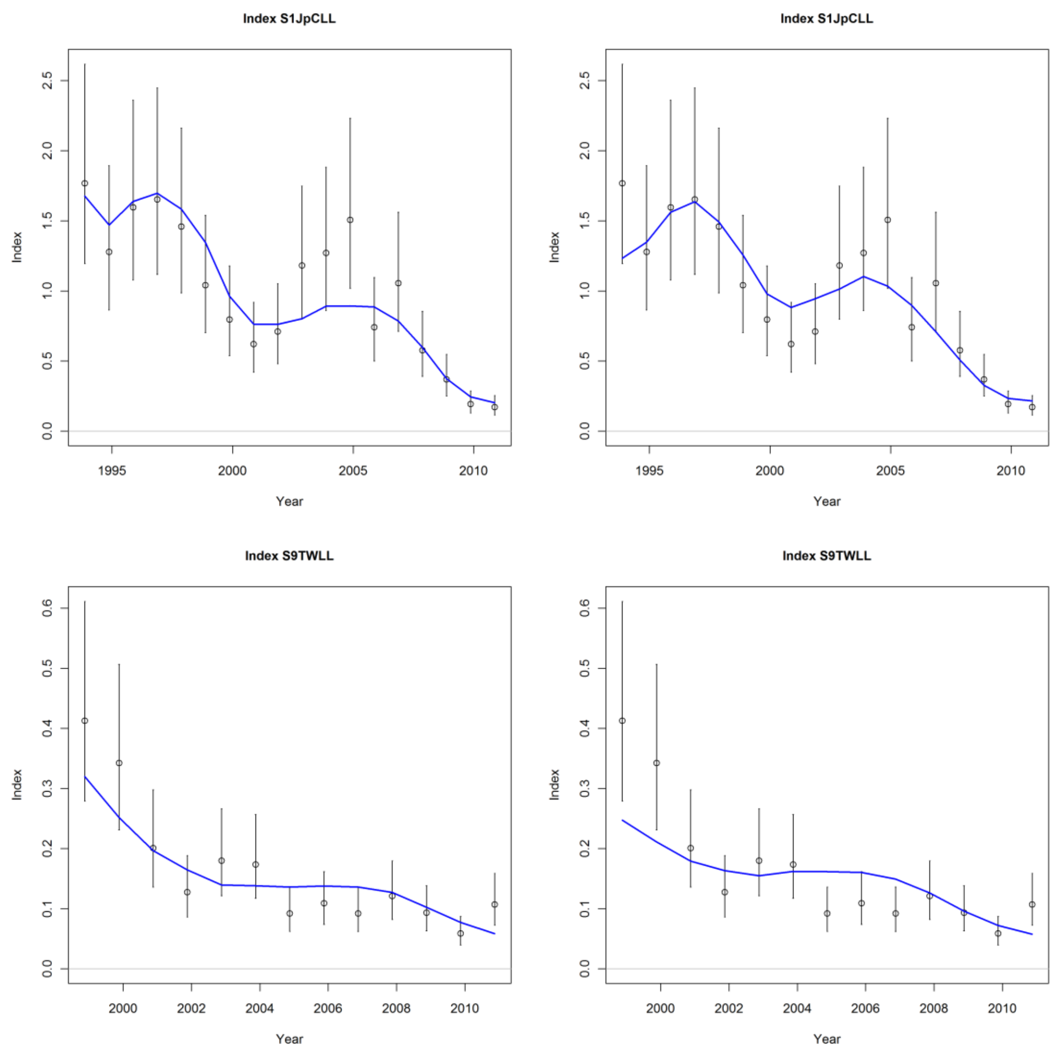


Figure 9. Fits to the terminal longline indices of abundance for sensitivity runs A (first column) and B (second column). Plots are for a) Japanese CLL (S1), and b) Taiwanese CLL (S9). See Table 4 for details of sensitivity runs.

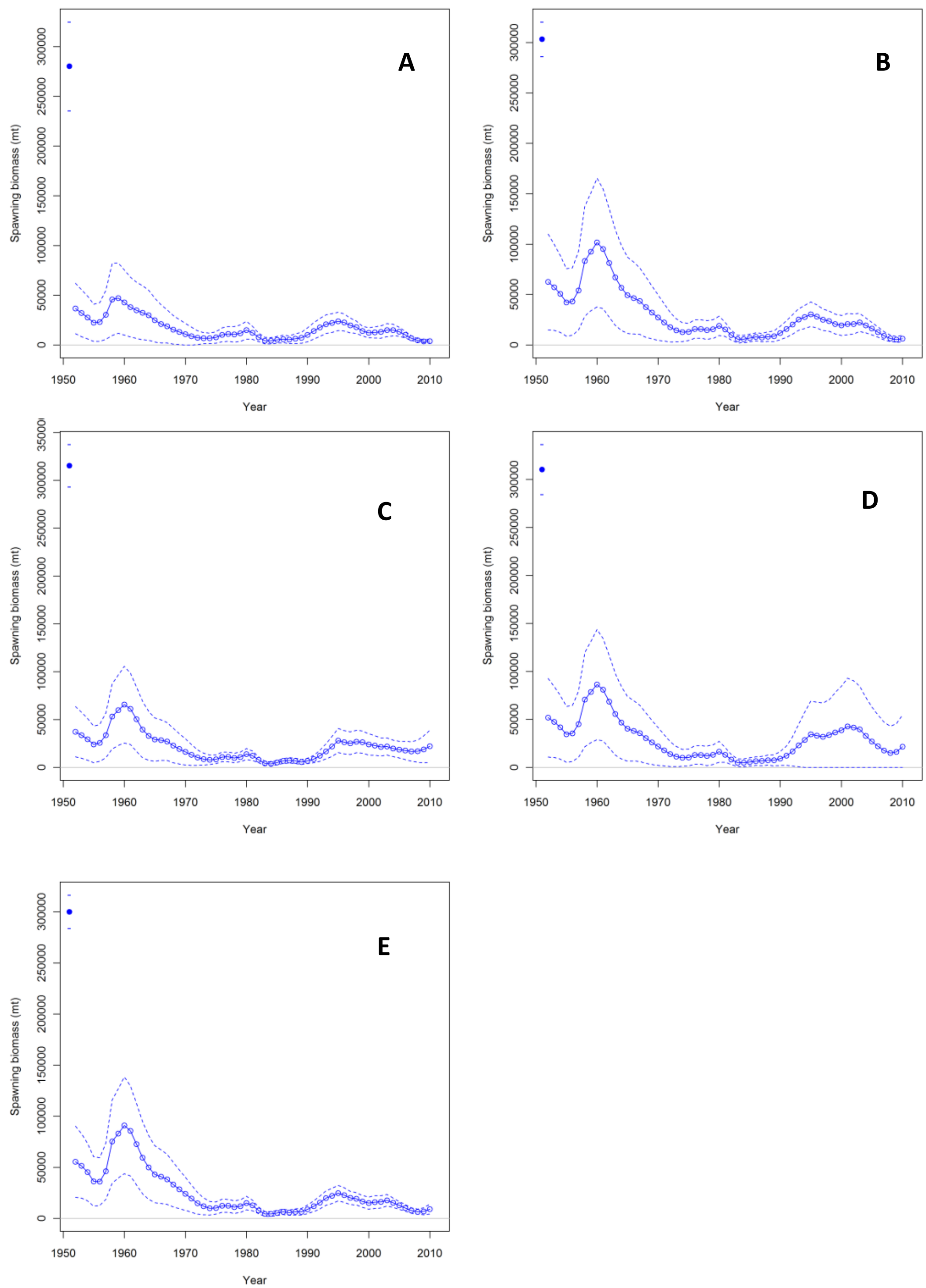


Figure 10. Spawning biomass estimates from sensitivity runs A-E. Dotted lines represent 95% CI based on the asymptotic error. See Table 4 for details of sensitivity runs.

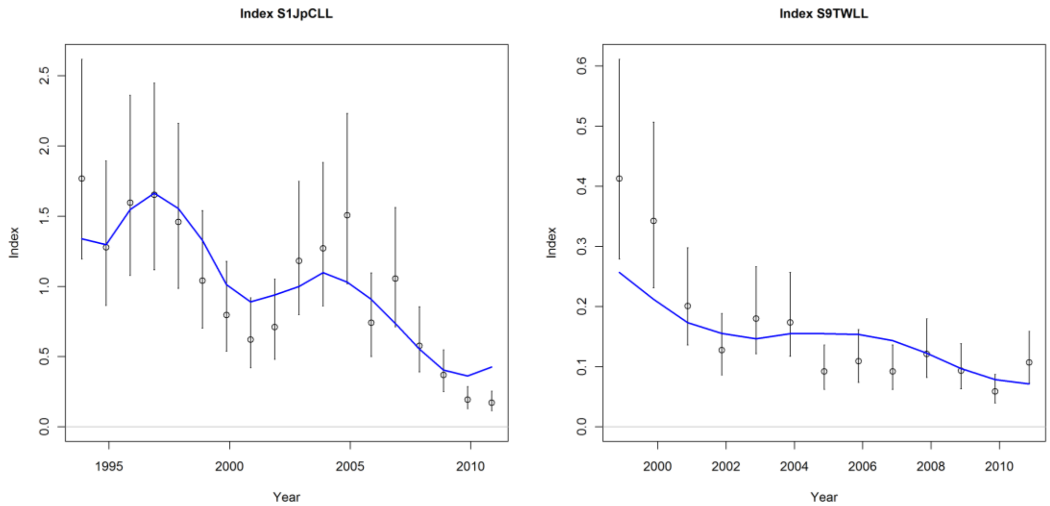


Figure 9. Fits to the terminal longline indices of abundance for sensitivity run E. Plots are for a) Japanese CLL (S1), and b) Taiwanese CLL (S9). Note last 3 years of S1 index are not fitted in this sensitivity run. See Table 4 for details of sensitivity runs.