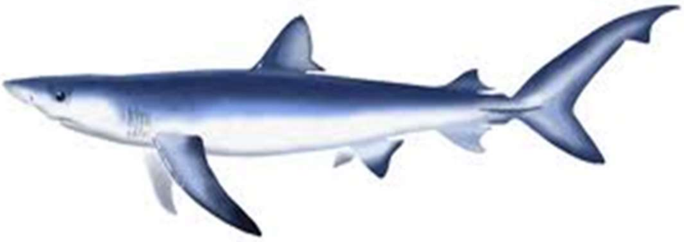
ISC/22/SHARKWG-2/6

Main outcomes and model diagnostics of the Stock Synthesis base-case model candidate in the stock assessment for blue sharks in the North Pacific [1]

Mikihiko Kai[2], Steven L. H. Teo[3], Nicholas Ducharme-Barth[4], Felipe Carvalho[4]

[2] Fisheries Resources Institute, Japan Fisheries and Education Agency,
5-7-1 Orido, Shimizu, Shizuoka, 424-8633, Japan
[3] NOAA Fisheries, Southwest Fisheries Science Center, La Jolla, California USA
[4] NOAA Fisheries, Pacific Islands Fisheries Science Center, Honolulu, Hawaii USA

Email: kaim@affrc.go.jp



---

[1] Working document submitted to the ISC Shark Working Group Workshop, 19-22, 26-28 March 2022, Web-meeting. **Document not to be cited without author's permission.**

Summary

This working paper summarizes main outcomes of the Stock Synthesis (SS3) base-case model candidate and its model diagnostics in the stock assessment for the blue sharks in the North Pacific Ocean. The main outcomes include: 1) Estimates of key management quantities; 2) Time series of spawning output, age-0 recruits (1000s) and fishing mortality; 3) Kobe plot. The model diagnostics include: 1) R0-profiles of the log-likelihoods; 2) Age-structured Production Model (ASPM); 3) Retrospective analysis; 4) Jitter analysis. The base-case model candidate was determined based on the entire model fits to the data that indicated better fits for the representative abundance indices and all length composition data, and the hessian matrix was positive definite with sufficient small value of maximum gradient (1.31E-05). The annual spawning biomass constantly decreased until 1992 and slightly increased until recent years. The annual fishing mortality gradually increased in the late 1970s and 1980s and suddenly dropped around 1990, after that it has been slightly decreasing. The annual age-0 recruits were stable over the assessment period and the number was maintained between 10 and 20 million. The Kobe matrix indicated that the current stock status is not overfished, and overfishing has not been occurred in recent years.

Introduction

Blue shark (*Prionace glauca*) is widely distributed from tropical to temperate waters around the globe and is the most abundant species of oceanic pelagic shark (Nakano and Steven, 2008). The stock assessment of blue shark in the North Pacific Ocean was conducted in 2017 by the ISC SHARK working group (WG). The Stock Synthesis (SS3) model platform (Method and Wetzel, 2013) was used to determine the stock status of the North Pacific stock of blue shark. In the 2017 assessment of this stock, the female spawning biomass in 2015 (SB_{2015}) was estimated to be 71% higher than at the maximum sustainable yield (MSY) level (SB_{MSY}), while the recent annual fishing mortality ($F_{2012-2014}$) was estimated to be well below F_{MSY}, at approximately 37% of F_{MSY}. Therefore, the results of the reference case, with input parameter values considered most probable, suggested that this stock was not overfished, and overfishing was not occurring, relative to MSY (ISC, 2017).

The objective of this working paper is to summarize the main outcomes and model diagnostics of the base-case model candidate. Details of the base-case model candidate can be found in Kai et al. (2022a, b).

Materials and Methods

Decisions of the SS3 base-case model candidate

The base-case model candidate was selected based primarily on the fits to the representative CPUEs (i.e., S5 and S6), and all the length composition data. In addition, the hessian matrix of the model was required to be positive definite with a sufficiently small maximum absolute gradient (i.e., < 0.0001). The model fits to those data (e.g., residual patterns of CPUE and length composition data) were visually checked using the residual plots for CPUEs and length composition data. In addition, the root mean square errors (RMSEs) of the model fits to relative abundance indices were calculated and compared. The RMSEs were expected to be < 30 % of the indices. During model development, some parameters were estimated to be on the parameter bounds and some parameter pairs were highly correlated (> 0.95). These issues were resolved by adjusting the initial values of selectivity parameters, by changing the structure of the selectivity (e.g., time block, time varying selectivity), or by fixing the less-effective parameters.

Model diagnostics

Procedures and details of contemporary model diagnostics in SS3 were concisely summarized in a cookbook (Carvalho et al., 2021). Four major model diagnostics (R_0-likelihood profile (R_0-profile), Age-structured production model (ASPM), Retrospective analysis and Jitter analysis) were used to identify problems with model specifications and to remove/minimize such problems. A likelihood profile of R_0 was conducted by sequentially fixing

R_0 to a range of values ($8.25 < \log(R_0) < 10.25$ with 0.1 intervals) to examine the change in the total and data-component likelihoods. The R_0-profile can detect relatively informative data sources as well as conflicts in the data and/or model misspecification.

The ASPM was conducted by fixing all the selectivity parameters to the estimated values and turning off estimation of all parameters except R_0 and the parameters representing initial conditions (e.g., initial fishing mortality parameters). The recruitment deviates were also set to zero to allow recruitment to follow the stock recruitment curve. Further, the penalty for recruitment deviates estimation in the likelihood function was turned off. The model was fit only to the catch and CPUE data. If the changes in catch can explain the trends in the indices of the ASPM, it suggests that a production function is apparent in the relationship between the catch and index data. If the general scale and trend of the population in the ASPM is consistent with the full model, it suggests that the data sources for both models are providing consistent information on the overall productivity of the stock.

Within-model retrospective analysis was conducted by sequentially removing the terminal year of observations (i.e., catch, CPUE and length composition data) from successive models (1 to 3 years), while keeping the same model structure between models. Retrospective analysis can detect potential biases in the estimated spawning stock biomass and fishing mortality, especially in the terminal period, by comparing the relative difference between model estimates from the full time series and the truncated time-series. The retrospective bias was quantified using an annualized version of Mohn's rho (Mohn, 1999) (ρ_{M}), which was proposed by Hurtado-Ferro et al. (2015). If ρ_{M} falls outside of -0.15 to 0.20, it indicates an undesirable retrospective pattern for longer-lived blue sharks (Carvalho et al., 2017). Additionally, a large positive ρ_{M} for spawning biomass indicates a systematic overestimation that may result in overly optimistic stock status from the assessment.

The jitter analysis was conducted by adding a small random jitter (10 % of the parameter range) to the initial parameter values with 100 iterations. The jitter analysis can detect whether the model has failed to converge to a global solution. In addition, we can compare the management quantities from the models in the jitter analysis to check that, even if the model had not converged to a global minimum, whether the results of the models would be highly similar.

The model diagnostics were conducted using the modified R-code on the website of ss3diags (https://pifscstockassessments.github.io/ss3diags/) and the entire model outputs such as time series of spawning biomass and fishing mortality were visually checked using the "SS Output" produced by R-package "r4ss" (Taylor et al. 2021).

Results and Discussions

Model convergence to a local or global minimum was indicated by a positive definite hessian matrix and a sufficiently small value of maximum gradient (1.31E-05). In addition, there were no parameters with correlations above 0.95 and no parameters were on bounds. The model fits to the annual CPUEs for S5 (JPN_EARLY) and S6 (JPN_LATE) were generally good, while the model fits to the length composition data for several fleets (F1, F3, F4, F15 and F18) had an apparent lack of randomness in their residual distributions (**Figs. 1, 2**). However, the RMSEs for both CPUEs (i.e., S5 and S6) and length composition data used in the preliminary base-case model were lower than 30% (**Tables 1, 2**), which indicated reasonably good fits overall to the CPUEs (i.e., S5 and S6) and length composition data.

R_0-profile diagnostics indicated that the gradient of the likelihood profile for the penalty on the recruitment deviations was greater than the other two data sources (**Fig. 3**). The second and third strongest gradient was observed for length composition data and CPUE data, respectively. The minimum value along the R_0-profile for the CPUE data was close to that of the length composition data, while that for the penalty on the recruitment deviates was much higher than any other data (**Tables 3)**. These results suggested that there is no conflict between the likelihood for the CPUE and the length composition data, but major conflict was observed between the likelihood for the penalty on the recruitment deviates and the other two data sources. Overall, the maximum likelihood estimate of R_0 is balancing conflicting signals from these three data sources.

ASPM diagnostics indicated that the CPUEs and spawning biomass showed somewhat similar trends to those of base-case model, while the spawning stock biomass of ASPM was higher than that of base-case model over the assessment period (**Fig. 4**). The fit to the indices in the ASPM was adequate but could be better. These results suggests that a production function is apparent, albeit weakly, in the relationship between the catch and index data. However, more work is needed to improve the fits to the indices in the ASPM as well as understand the sources for the difference in scale between the ASPM and base-case models.

Retrospective results indicated a consistent negative bias for the spawning stock biomass but no clear bias for the fishing mortality and annual CPUE (**Fig. 5**). It should be noted that the ρ_M=-0.19 for spawning stock biomass and ρ_M= 0.26 for fishing mortality rate are outside of the acceptable range (-0.15 to 0.20) for long-lived species. In addition, the reduction of the absolute spawning stock biomass as the number peel years increased were due to the reduction in the estimated age-0 recruitment over the assessment period. Here, the reason of the three year's retrospective analysis is that the F5_JPN_KK_DP has no size data prior to 2017 that may result in the difficulty of estimating selectivity parameters for the fishery.

Jitter analysis indicated that all of the 100 model runs were converged. Some runs showed larger negative log-likelihood but the likelihoods for most of the runs were close to that of the base-case model candidate (**Fig. 6**). These results suggested that the base-case model candidate was sensitive to initial values of the parameters but there is low possibility that the likelihood of the base-case model candidate is far from the global minimum.

The annual spawning biomass (SB) constantly decreased until 1992 and slightly increased until recent years (**Fig. 7**). The female spawning biomass in 2020 (SB_{2020}) was 54% higher than at maximum sustainable yield (MSY) level (SB_{MSY}) (**Tables 4**). The annual fishing mortality (F) gradually increased in the late 1970s and 1980s and suddenly dropped around 1990, after that it has been slightly decreasing (**Fig. 8**). The recent annual fishing mortality ($F_{2017-2019}$) was estimated to be well below F_{MSY}, at approximately 46% of F_{MSY}. Therefore, relative to MSY, the base-case model candidate with input parameter values considered most probable suggested that this stock is not overfished, and overfishing is not occurring.

The annual age-0 recruits were stable over the assessment period and the number was maintained between 10 and 20 million (**Fig. 9**). The estimated stock-recruitment relationships had no clear correlation (0.06) and the expected recruitment slightly increased as the spawning biomass increased (**Fig. 10**). The trends in the annual recruitment deviations showed positive values in 1980s and 1990s and negative values after 2000 (**Fig. 11**). The annual variance of the stock-recruitment relationships was shown in **Fig. 12** and the bias adjustment used in the preliminary base-case model fitted well to the variances. Overall, the fits to the sex-specific length composition data by fleet were good (**Fig. 13**) and the estimated selectivity curves seemed reasonable in light of the size ranges of the catch for each fleet (**Fig. A1**). The historical trajectories of stock status revealed that North Pacific blue shark had experienced some level of depletion and overfishing in previous years showing that the trajectories moved through the orange (overfishing) zone in the Kobe plots (**Fig. 14**). However, in the last three decades, the stock condition returned into the Kobe green zone. By the standard terminology, this would indicate that the stock is not in an overfished state, and that overfishing is not occurring.

References

Carvalho et al., 2021. A cookbook for using model diagnostics in integrated stock assessments. Fish. Res. 240. 105959.

Carvalho, F., Punt, A.E., Chang, Y.-J., Maunder, M.N., Piner, K.R., 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments? Fish. Res. 192, 28–40.

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A., Whitten, A.R., Punt, A.E., 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. Ices J. Mar. Sci. 72, 99–110.

ISC. 2017. Stock assessment and future projections of blue shark in the North Pacific Ocean through 2015. ISC Plenary report and documents. Annex 13.

Kai, M., Teo, S.L.H., Ducharme-Barth, N., Semba,Y., Carvalho, S. 2022a. Stock Synthesis settings of data file for the stock assessment of blue shark in the North Pacific. ISC/22/SHARKWG-2/4.

Kai, M., Teo, S.L.H., Ducharme-Barth, N., Semba,Y., Carvalho, S. 2022b. Stock Synthesis settings of control file for the stock assessment of blue shark in the North Pacific. ISC/22/SHARKWG-2/5.

Method, R.D., Taylor, I.G. 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Can. J. Fish. Aquat. Sci. 68, 1744–1760.

Method, R.D., Wetzel, C.R., Taylor, I.G., Doering, K. 2020. Stock Synthesis User Manual Version 3.30.16. NOAA Fisheries, Seattle, WA.

Mohn, R. 1999. The retrospective problem in sequential population snalysis: An investigation using cod fishery and simulated data. ICES J. Mar. Sci. 56, 473–488.

Nakano, H., Stevens, J. D. 2008. The biology and ecology of the blue shark, *Prionace glauca*. In: Camhi, M. D., Pikitch, E. K., Babcook, E. A., editors. Sharks of the open ocean: biology, fisheries and conservation. Oxford: Blackwell Scientific. p. 140–51.

Taylor, I.G., Doering, K.L., Johnson, K.F., Wetzel, C.R, Stewart, I.J. 2021. Beyond visualizing catch-at-age models: Lessons learned from the r4ss package about software to support stock assessments, Fish. Res. 239, 105924. https://doi.org/10.1016/j.fishres.2021.105924

Table 1. Root mean squared error of abundance indices for each fleet. Two representative CPUEs for S5 and S6 were used in the model fits to the data in the base-case model.

Fishery number	Reference Code	Number of observation	RMSE (%)
S1	US_HW_DP	19	18.9
S3	TAIW_LG	17	53.7
S5	JPN_EARLY	18	7.7
S6	JPN_LATE	27	8.5
S7	JPN_RTV	27	35.3
S9	SPC_OBS_TROPIC	27	28.9
S10	MEX	17	25.2

Table 2. Root mean squared error of length composition data for each fleet.

Fishery number	Reference Code	Number of observation	RMSE (%)
F1	MEX	15	5
F3	CHINA	26	8.3
F4	JPN_KK_SH	13	9.8
F5	JPN_KK_DP	5	4.6
F7	JPN_ENY_DP	29	3.9
F8	JPN_LG_MESH_EARLY	5	8.9
F9	JPN_LG_MESH_LATE	12	9.6
F14	NON_ISC	19	18.5
F15	US_GIILL	31	11.7
F17	US_HW_DP	18	6.3
F18	US_HW_SH	17	17
F19	TAIW_LG	17	1.8
F20	TAIW_SM	9	3.1

Table 3. Log-likelihood profiles for R0 for all data likelihood components. Bold denotes the minimum value of each component.

Log (R0)	Total likelihood	Survey	Length composition	Recruitment
8.75	137.96	-71.75	130.29	67.40
8.85	110.96	-73.10	126.52	48.32
8.95	89.19	-75.59	128.17	30.22
9.05	66.71	-75.67	123.19	14.68
9.15	51.05	**-76.40**	120.62	3.36
9.25	39.23	-76.25	116.15	-2.65
9.35	34.92	-76.09	**114.65**	-5.29
9.45	37.49	-76.28	116.11	-6.97
9.55	32.78	-76.24	114.84	-7.85
9.65	33.61	-76.23	116.10	-8.61
Base case (9.66)	**32.38**	-76.23	114.86	-8.59
9.75	32.42	-76.21	115.04	-9.09
9.85	32.72	-76.16	115.21	-9.38
9.95	33.15	-76.10	115.34	-9.55
10.05	33.65	-76.04	115.44	-9.64
10.15	34.21	-75.96	115.52	**-9.65**
10.25	34.81	-75.88	115.59	-9.62

Table 4. Estimates of key management quantities for the base-case model.

No	Manegment Quantity	Unit	Base-case model
1	SB_{1971}	MT	177,072
2	SB_{2020}	MT	138,987
3	SB_{MSY}	MT	90,166
4	MSY	MT	54,183
5	F_{1971}	$year^{-1}$	0.331
6	$F_{2017-2019}$	$year^{-1}$	0.332
7	F_{MSY}	$year^{-1}$	0.729
8	SB_{2020}/SB_{MSY}		1.541
9	$F_{2017-2019}/F_{MSY}$		0.457

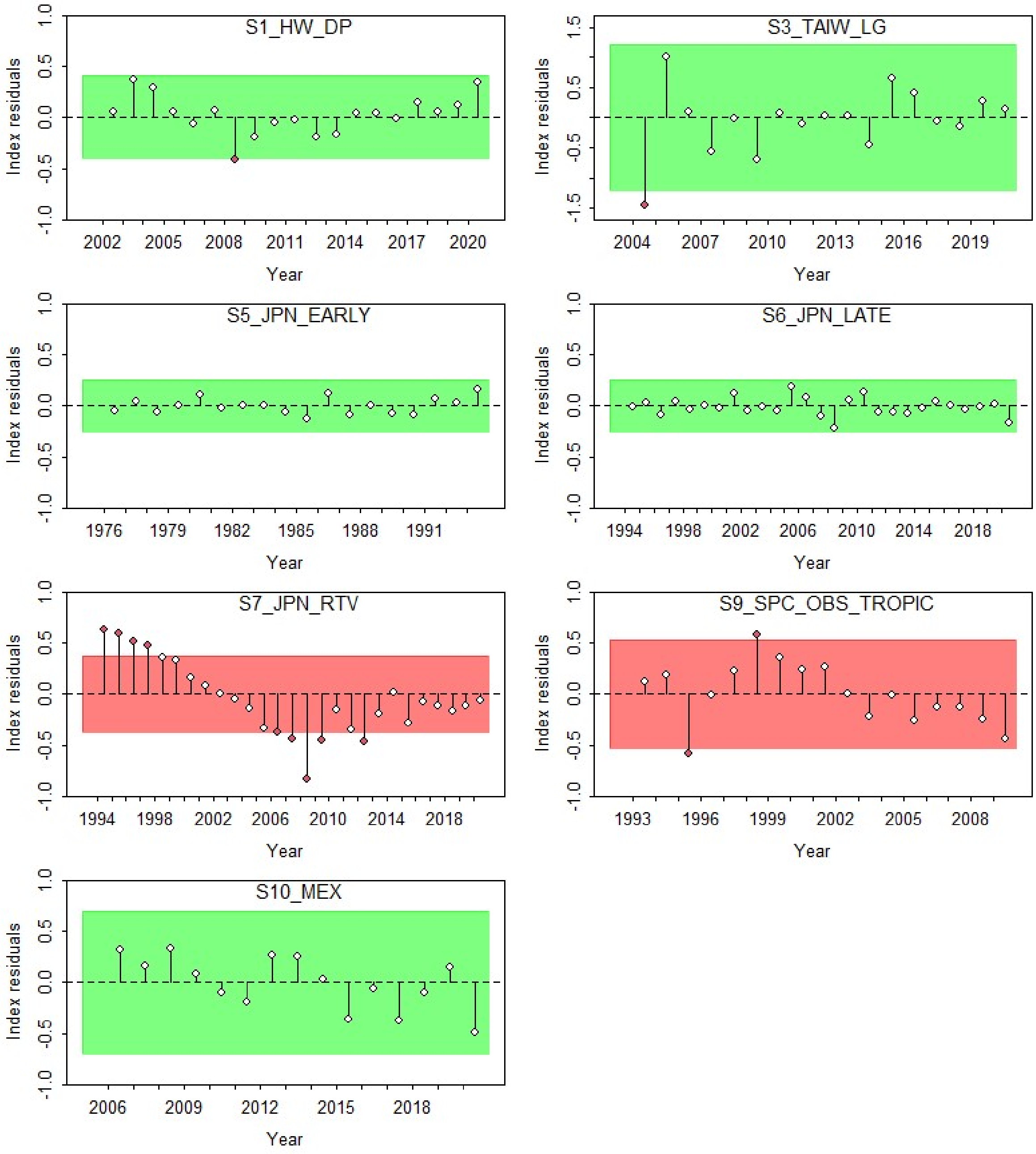


Figure 1. Annual index residuals by fleets from the runs test and "three-sigma limit" values. Green shaded area indicates the residuals are randomly distributed (p-value ≥ 0.05) and red shaded area indicates the residuals are not randomly distributed and there is some misspecification. Two representative CPUEs for S5 and S6 were used in the model fits to the data in the base-case model.

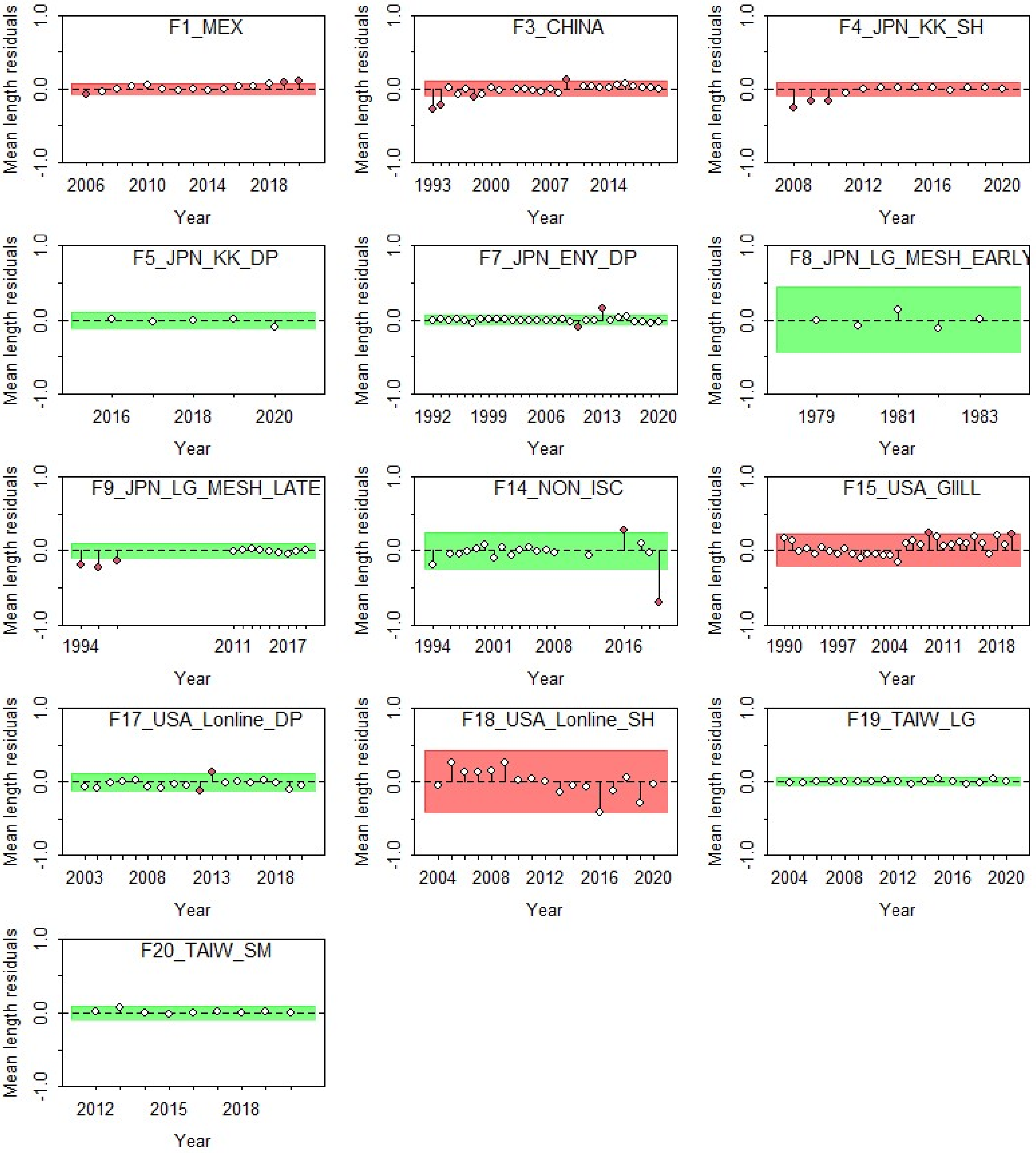


Figure 2. Annual mean length residuals by fleets from the runs test and “three-sigma limit” values. Green shaded area indicates the residuals are randomly distributed (p-value ≥ 0.05) and red shaded area indicates the residuals are not randomly distributed and there is some misspecification.

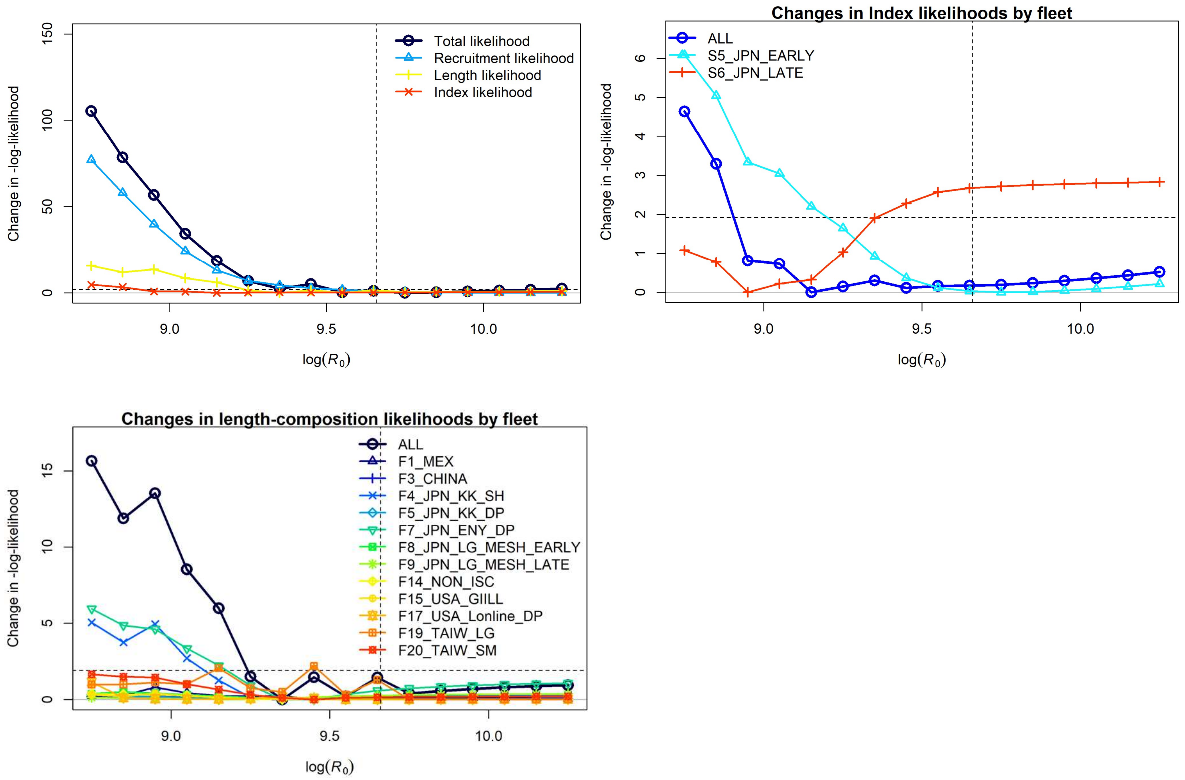


Figure 3. Log-likelihood profiles for R0 for the various data components; all data likelihood components (upper left panel), among abundance indices (upper right panel) and among length composition data (lower panel). Vertical dotted line denotes the value of the base-case model.

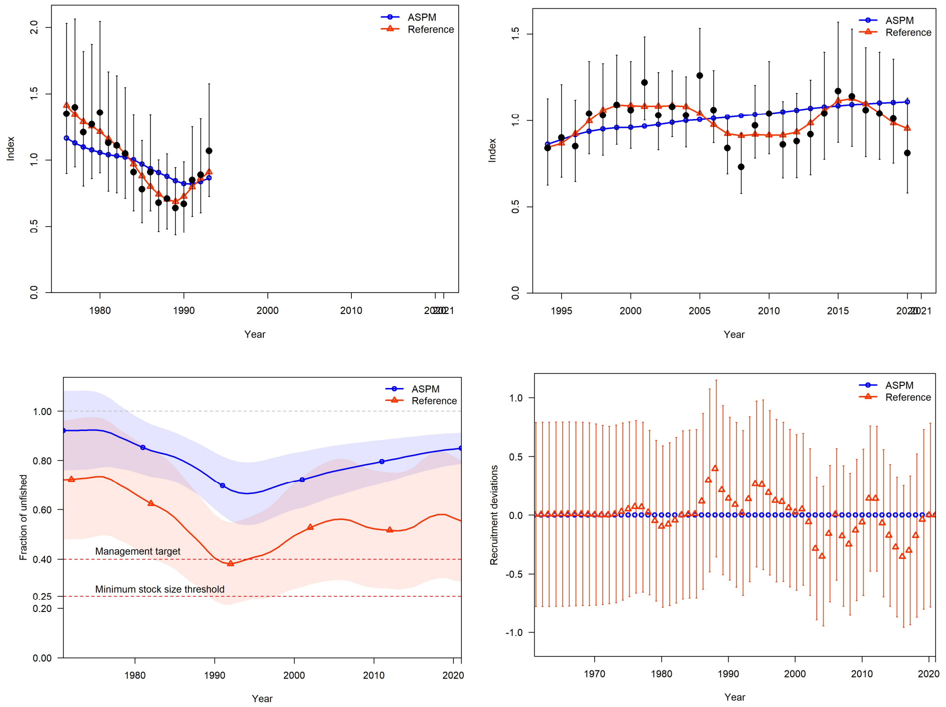


Figure 4. Comparisons of estimated time series for the CPUEs of S5 (upper left) and S6 (upper right), spawning stock biomass relative to unfished level (lower left) and recruitment deviation (lower right) between the base-case model (Reference) and Age-Structured-Production Model (ASPM).

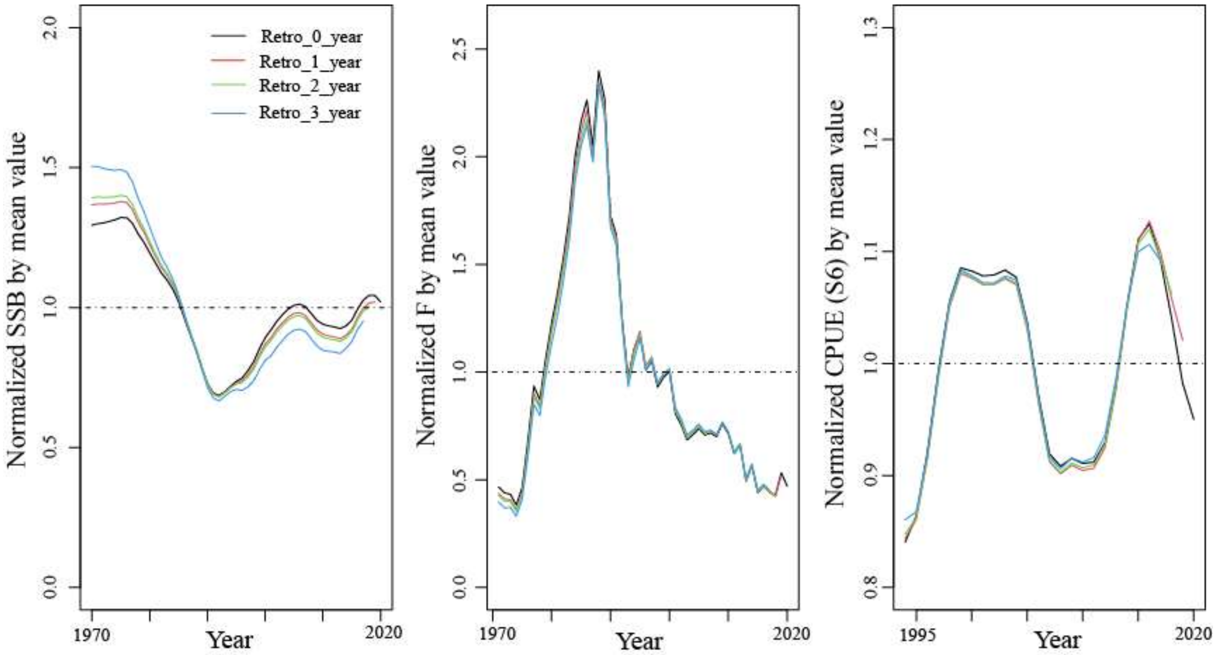


Figure 5. The three-year's retrospective results for the time series of spawning stock biomass, fishing mortality rate and late CPUE (S6). The time series were normalized by the mean value.

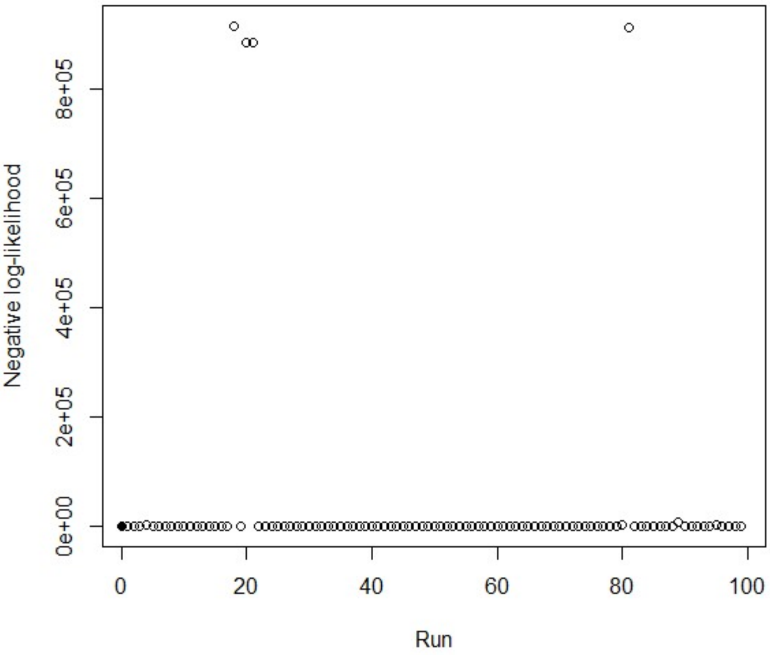


Figure 6. Negative log-likelihood values from the 100 iteration of Jitter analysis. Black circle denotes the total negative-log likelihood of the base-case model candidate.

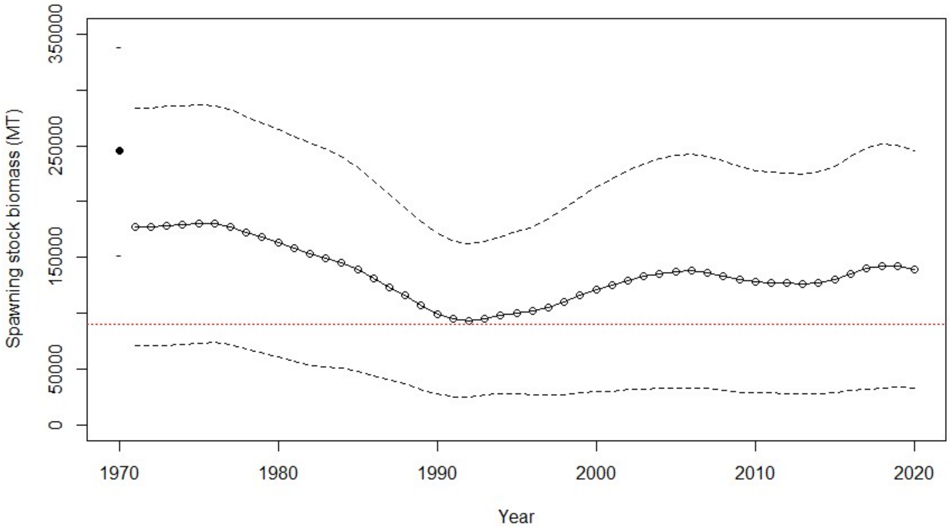


Figure 7. Annual spawning stock biomass (metric tons). The filled circle in 1970 denotes the unfished spawning stock biomass. Red line denotes the estimator of maximum sustainable yield (MSY) and the dotted lines are 95% confidence intervals.

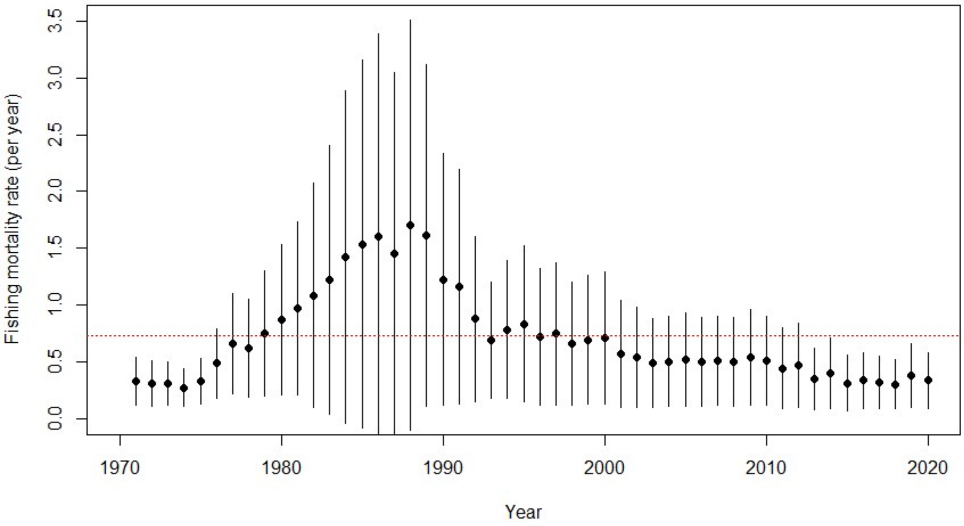


Figure 8. Annual fishing mortality (sum of them across all fleets). Red line denotes the estimator of maximum sustainable yield (MSY) and the vertical lines denote 95% confidence intervals.

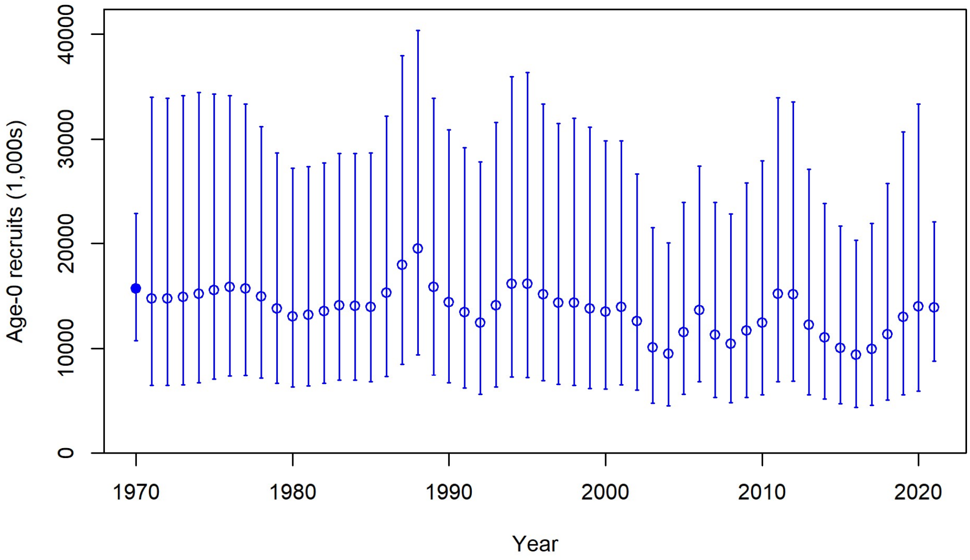


Figure 9. Annual age-0 recruits (1000s) with 95% confidence intervals. The filled circle in 1970 denotes the unfished age-0 recruits.

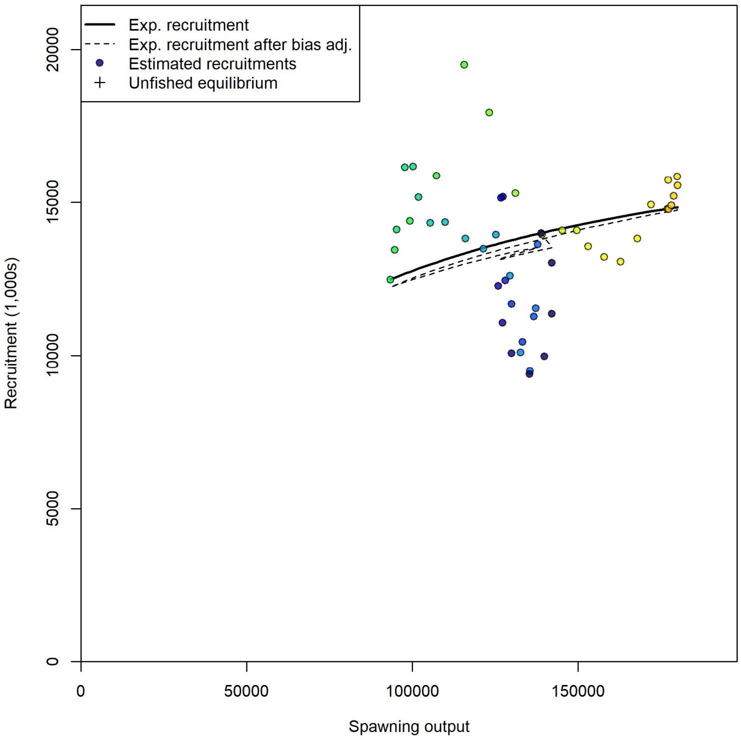


Figure 10. Estimated stock recruitment relationships (filled circle) and expected curves of stock recruitment with bias adjustment (broken line) and without bias adjustment (solid line).

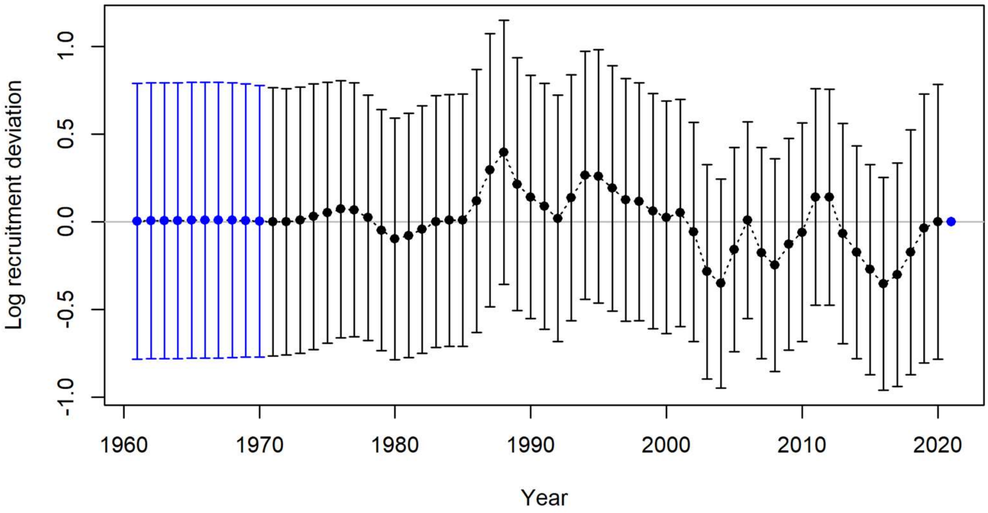


Figure 11. Time series of recruitment deviation on the log-scale with 95% confidence intervals.

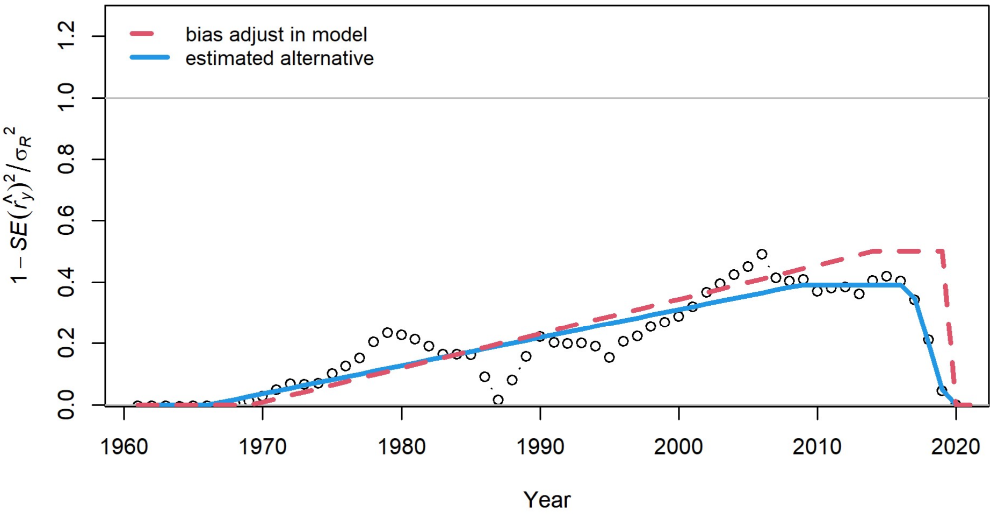


Figure 12. Points are transformed variances of recruitment deviations. Red line indicates current settings for bias adjustment and blue line indicates least squares estimate of alternative bias relationship for recruitment deviations. See the paper by Method and Taylor (2011) for the details.

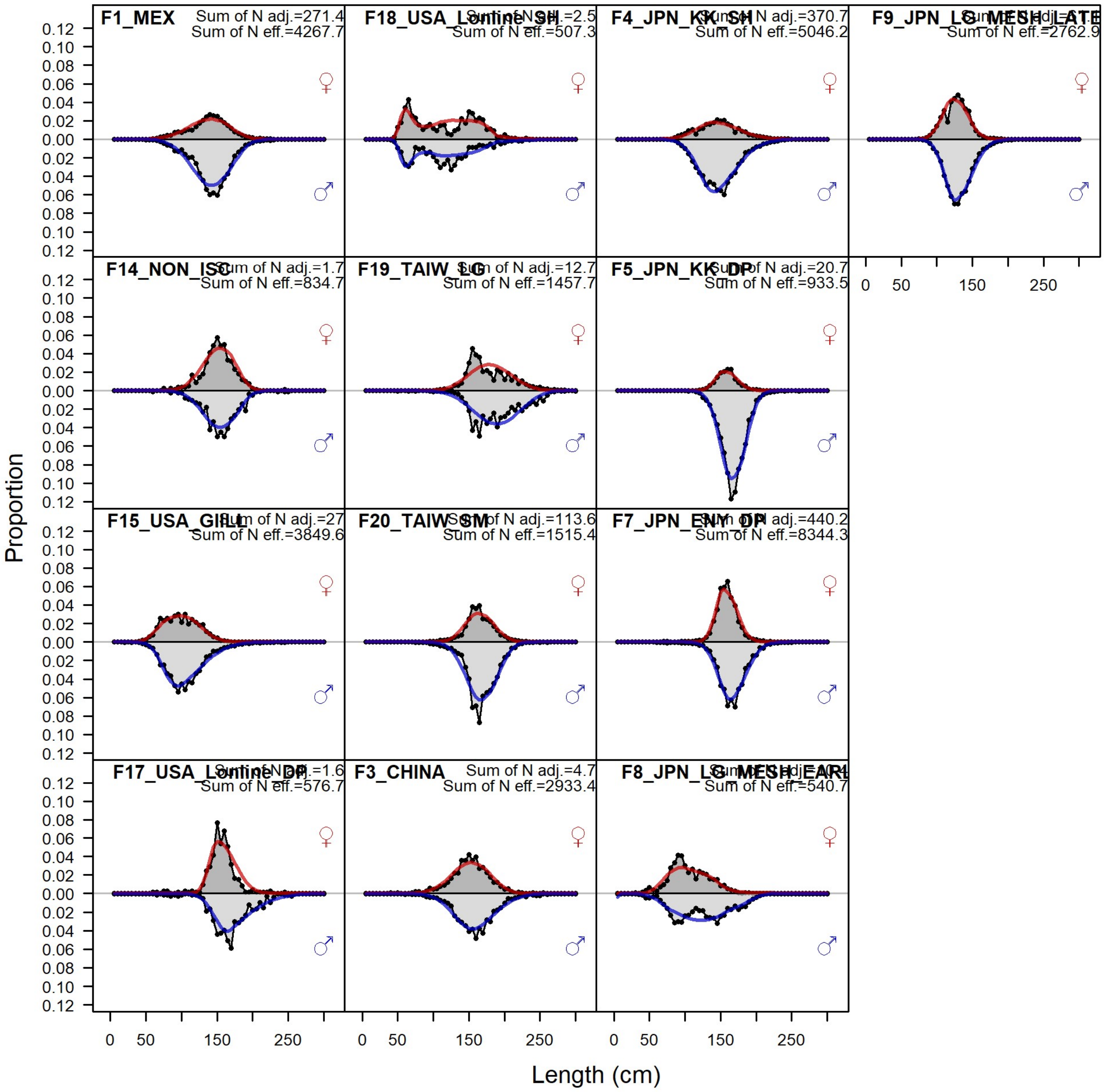


Figure 13. Fits to the sex-specific length composition data across time by fleet.

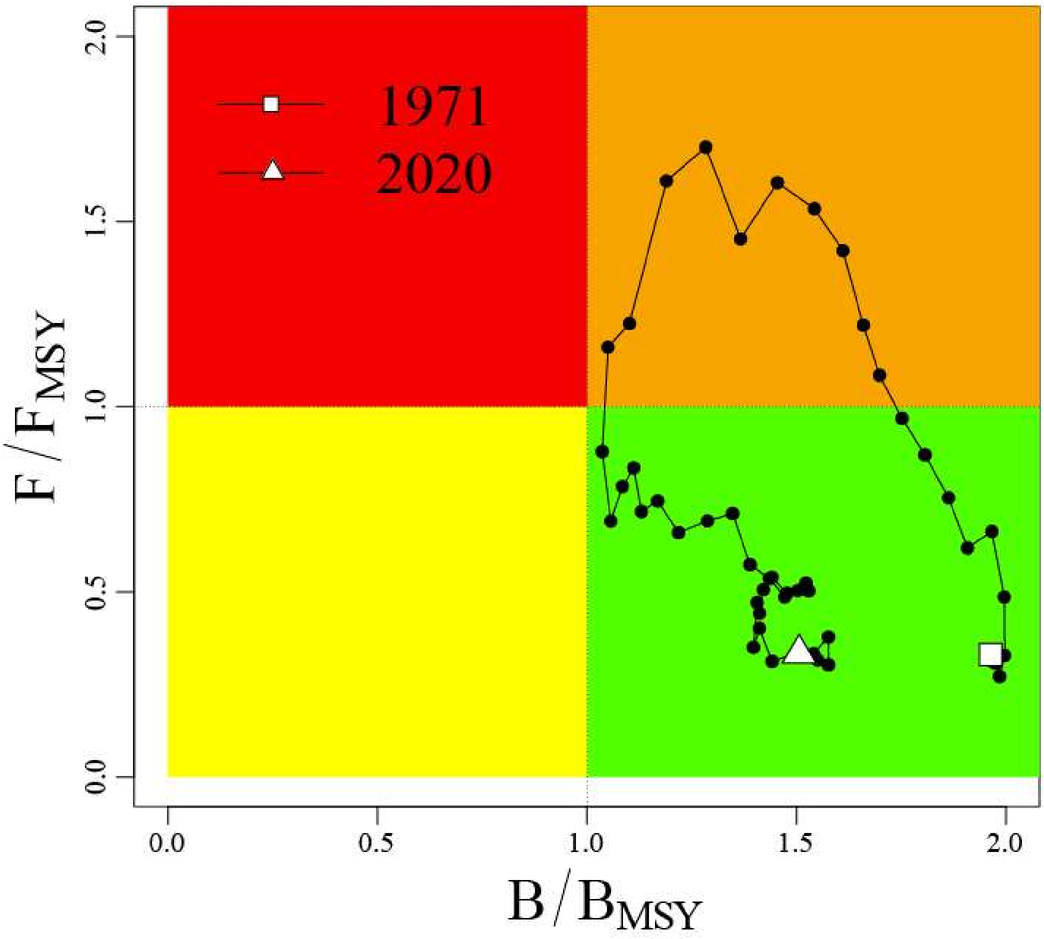


Figure 14. Kobe plot of the trends in estimates of fishing mortality and spawning stock biomass relative to maximum sustainable yield (MSY) levels for 1971-2020 for the base-case model.

Appendix figures

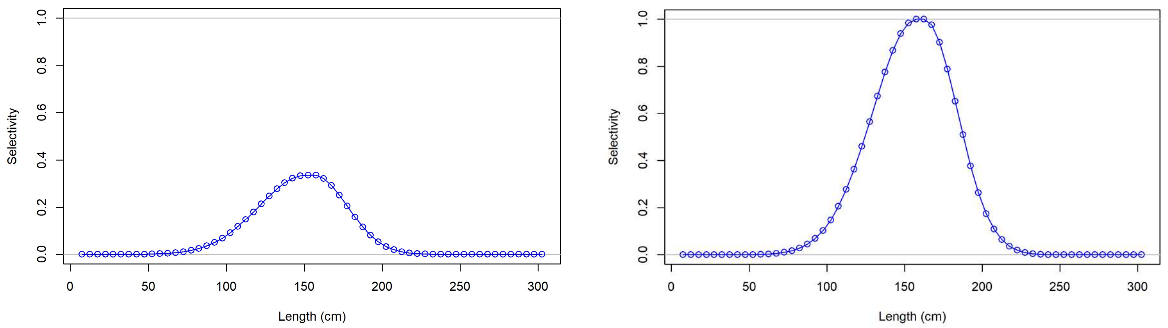


Figure A1. Estimated selectivity curves for female (left panel) and male (right panel) for F1: MEX, F2: CAN and F12: IATTC.

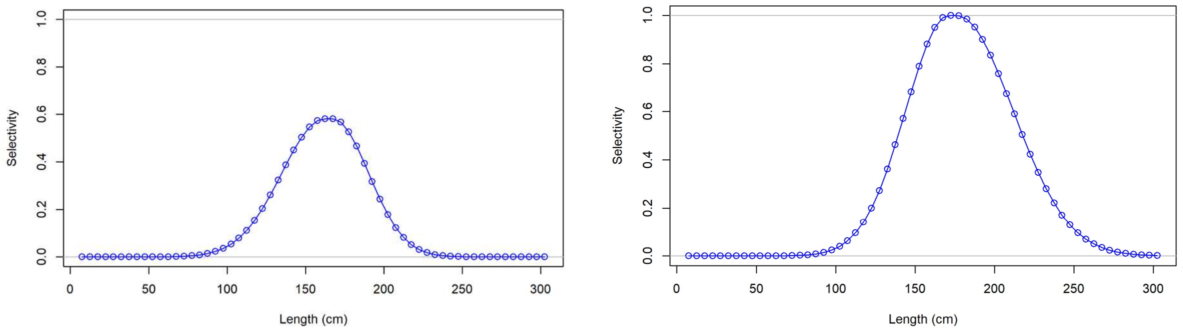


Figure A2. Estimated selectivity curves for female (left panel) and male (right panel) for F3: CHINA and KOREA.

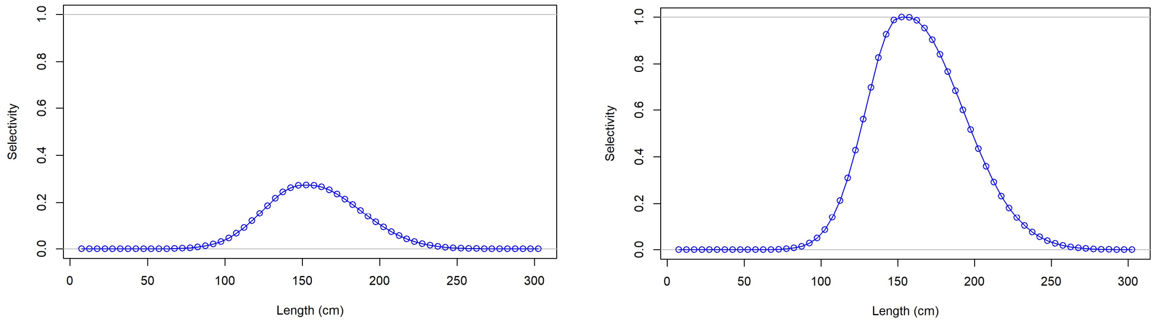


Figure A3. Estimated selectivity curves for female (left panel) and male (right panel) for F4: JPN_KK_SH and F6: JPN_ENY_SH.

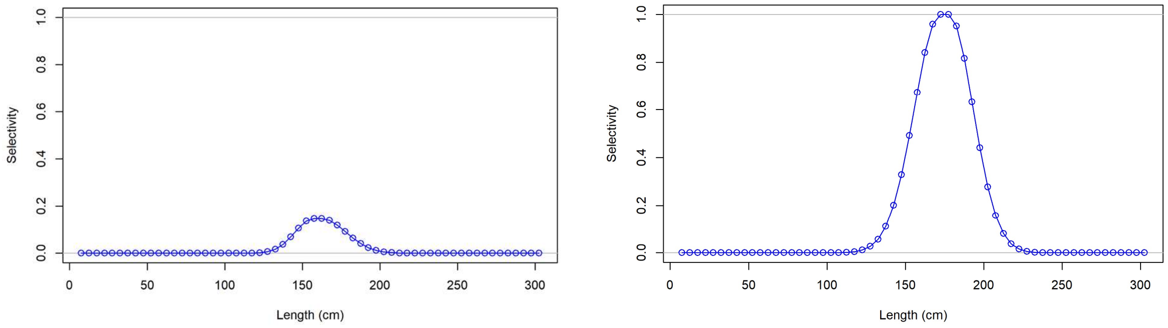


Figure A4. Estimated selectivity curves for female (left panel) and male (right panel) for F5: JPN_KK_DP.

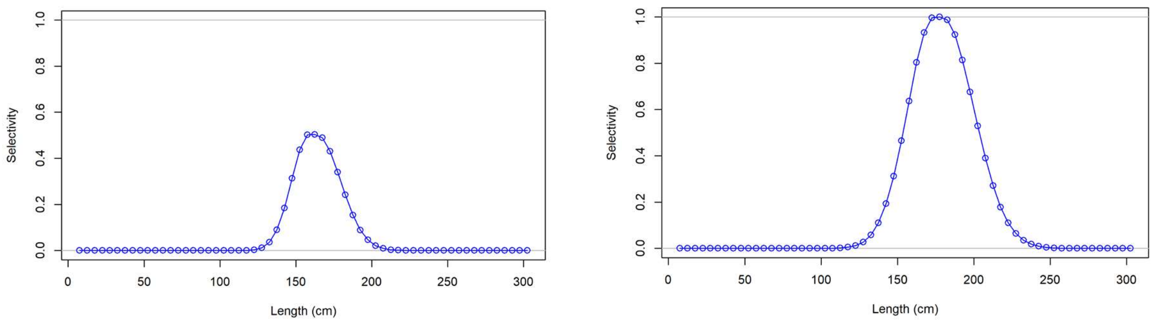


Figure A5. Estimated selectivity curves for female (left panel) and male (right panel) for F7: JPN_ENY_DP.

Figure A6. Estimated selectivity curves for female (left panel) and male (right panel) for F8: JPN_LG_MESH_EARLY.

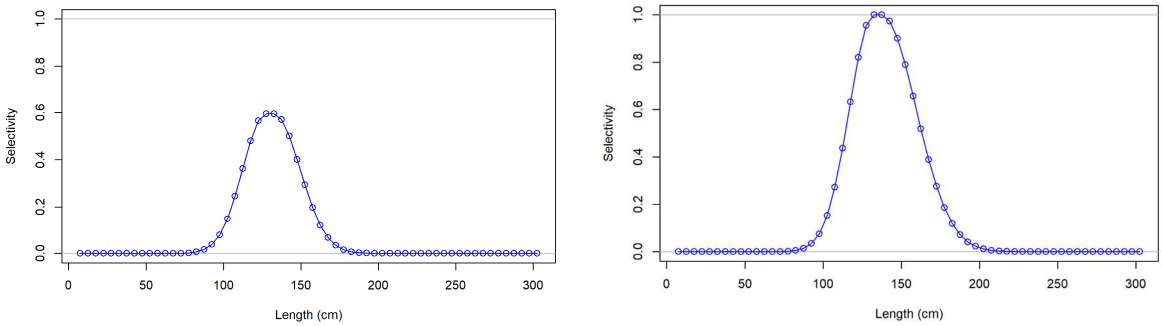


Figure A7. Estimated selectivity curves for female (left panel) and male (right panel) for F9: JPN_LG_MESH_LATE and F10: JPN_CST_OTH.

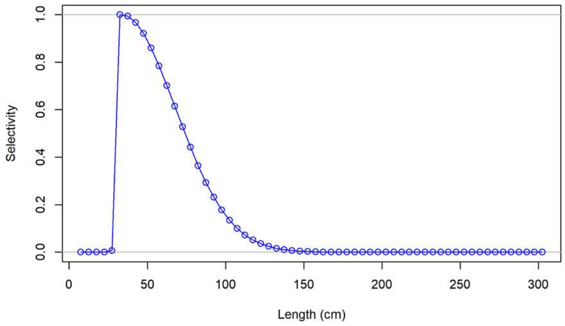


Figure A8. Estimated sex-combined selectivity curves for F11: SM_MESH.

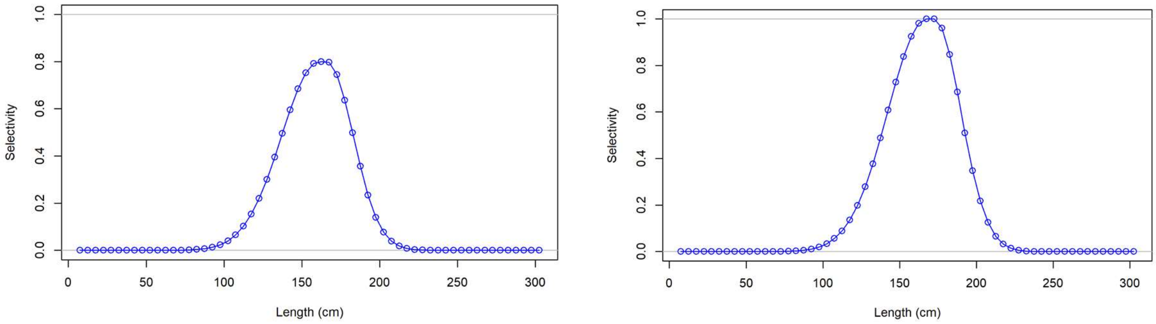


Figure A9. Estimated selectivity curves for female (left panel) and male (right panel) for F14: NON_ISC.

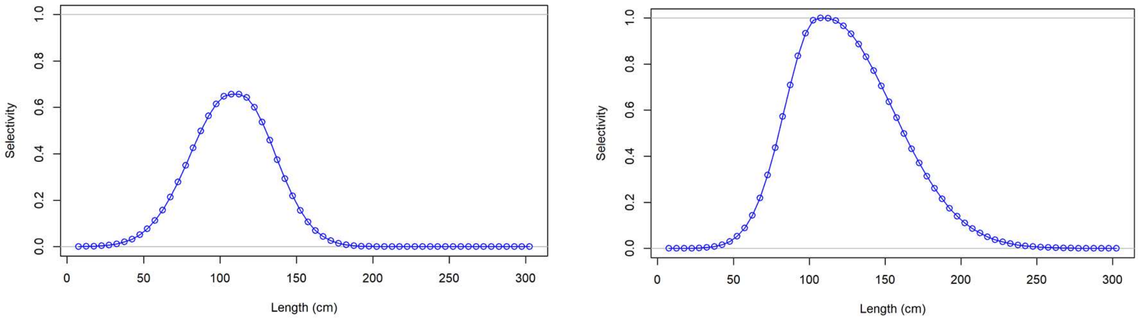


Figure A10. Estimated selectivity curves for female (left panel) and male (right panel) for F15: US_GILL and US_SPORT.

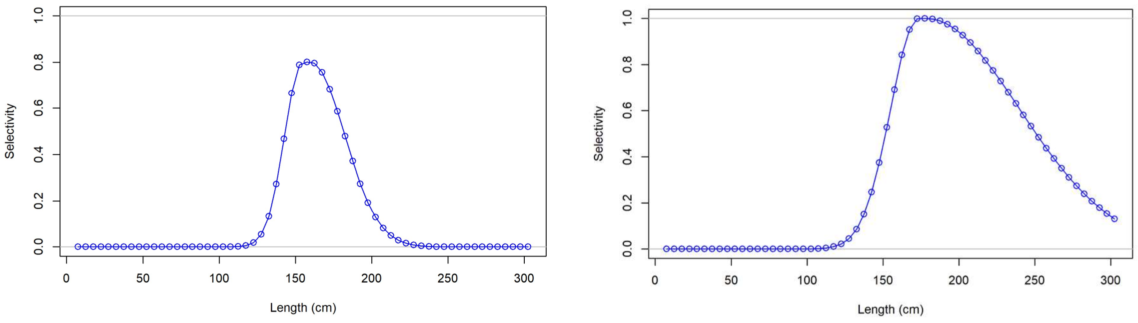


Figure A11. Estimated selectivity curves for female (left panel) and male (right panel) for F17: US_HW_DP.

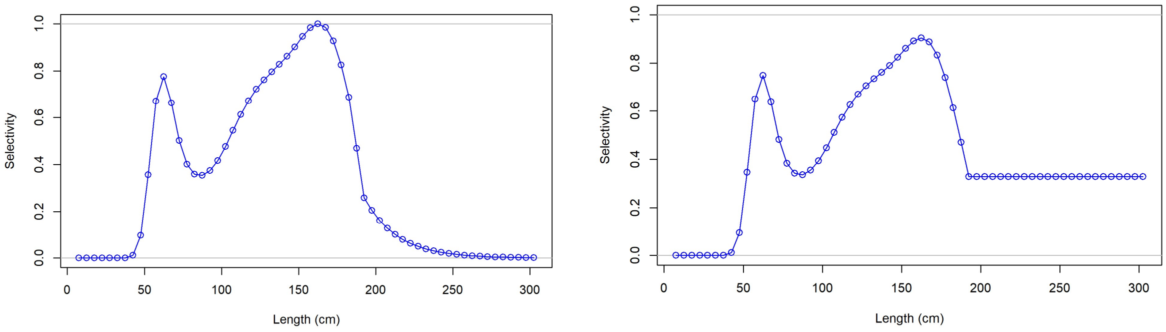


Figure A12. Estimated selectivity curves for female (left panel) and male (right panel) for F18: US_HW_SH.

Figure A13. Estimated selectivity curves for female (left panel) and male (right panel) for F19: TAIW_LG.

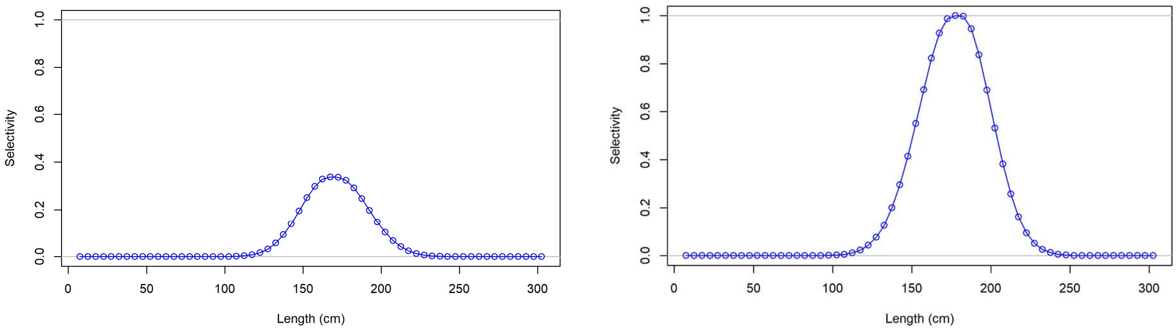


Figure A14. Estimated selectivity curves for female (left panel) and male (right panel) for F20: TAIW_SM.