# Preliminary Base-case Models in Stock Synthesis 3.30 for Consideration in the 2021 Pacific Blue Marlin (*Makaira nigricans*) Stock Assessment

Michelle Sculley,* Yi-Jay Chang+

*NOAA National Marine Fisheries Svc
Pacific Islands Fisheries Science Center
+Department of Oceanography
New Taiwan University

Email: michelle.sculley@noaa.gov

## Abstract

Two preliminary base-case models in Stock Synthesis 3.30 for Pacific blue marlin (*Makaira nigricans*) are described for consideration as the 2021 base-case model. The base-case model covers 1971-2016. It includes data from three International Scientific Committee for the Conservation of Tuna and Tuna-like Species (ISC) countries and other countries in aggregate from the Western Central Pacific Fisheries Council (WCPFC) and Inter-American Tropical Tuna Commission (IATTC). This paper describes the data available for inclusion in the base-case model, a model using the biological parameters from the 2016 base-case stock assessment model (old growth) and a model using an updated growth curve presented to the ISC Billfish Working Group at the 2021 Data Preparatory meeting (new growth). Both models converge and appear to fit the data well. Initial diagnostics do not indicate major problems. Preliminary results from both models suggest the Pacific blue marlin stock is being fished below $F_{MSY}$ and spawning stock biomass is above $SSB_{MSY}$.

## Introduction

The International Scientific Committee for the Conservation of Tuna and Tuna-like Species (ISC) Billfish Working Group (BILLWG) has proposed to run a benchmark assessment on Indo-Pacific blue marlin (*Makaira nigricans,* BUM). Data were compiled from the International Scientific Committee for North Pacific Tuna and Tuna-like Species (ISC) member countries, other Western and Central Pacific Fisheries Commission (WCPFC) countries, and other Inter-American Tropical Tuna (IATTC) countries. Countries were asked to contribute catch, CPUE, and size-frequency data. It was decided to run the assessment using a two-sex, single-stock model in Stock Synthesis version 3.30 (Methot and Wetzel, 2013). Biological parameters were discussed by the billfish working group (BILLWG) at the data preparatory meeting in November 2020, where a new growth curve was presented (Chang *et al.,* 2020). The WG agreed that both growth curves were plausible, although significantly different in terms of function (von Bertalanffy vs two-stanza growth) and $L_{inf}$, a parameter that has been shown to be very sensitive in assessment models. Due to this, the WG agreed to explore two possible base-case models for BUM, one using the life history parameters from the 2016 assessment (hereafter, the old growth model) and one using the new growth curve and updated natural mortality based upon the new growth curve (the new growth model). The available data and the preliminary model results and diagnostics for both of these models will be presented in this document for consideration at the ISC BILLWG BUM stock assessment meeting.

# Methods

## *Spatiotemporal structure*

The Indo-Pacific blue marlin (*Makaira nigricans*) is assessed as a Pacific-wide stock. Blue marlin (BUM) are found in tropical and sub-tropical waters throughout the entire region. The working group agreed to run the model from 1971 to 2019 when catch, CPUE, and size-frequency data are all available, although there has been fishing on the stock historically, with industrial catch records as early as the 1950s.

## *Definition of fisheries*

Twenty different fleets are available for inclusion in the base-case model, 16 catch time series, 4 CPUE indices, 7 fleets with length composition data, and one fleet with weight composition data. The fleet names and numbers are detailed in Table 1. The data available for each fleet is in Figure *1*. The acronyms in the fleet names are defined as follows: WCPFC is Western and Central Pacific Fisheries Commission; EPO is Eastern Pacific Ocean; LL is longline; CLL is coastal longline; early is the early time period; late is the late time period, DRIFT is high-seas and coastal driftnet gear; Oth is other fishing gear (e.g. troll, handline, net, harpoon, and others); PS is purse seine gear; Bait is bait fishing.

## *Catch*

Three ISC countries contributed catch time series, Japan, Taiwan, and the US (for both Hawaii and American Samoa). Also, catch from countries reporting to the WCPFC and IATTC were obtained from each RFMO, respectively (Figure *2*). All catch reported was in biomass (mt). The CV for catch was set to 0.05 for all fleets. Catch for fleets with only annual data were divided equally into each quarter. Six Japanese catch time series were provided from 1971-2019, the catch from longline fleets F1-2 was reported quarterly, F3 was reported annually until 1993 and quarterly from 1994-2019, and all other fleets, F4-6, reported annually (Ijima, 2021b). The catch series for F1 and F2 were divided in 1993 because of significant changes in the logbook reporting system, which requires the CPUE data to be standardized in different time periods.

The US catch data were provided from 1971-2019 (Ito and Sculley, 2021). The longline fleets F7 and F8 were provided quarterly and the other fleet, F9, annually. The US longline catch was divided into two fleets. F7 is catch from the Hawaii-based longline fleet targeting swordfish and tunas, this fleet fishes primarily in the Northern Pacific. F8 is catch from the American Samoa-based longline fleet targeting tunas and is primarily based in the Southern Pacific. Fleet F9 is troll and handline gear.

Taiwanese catch data were provided from 1971-2016 and all fleets were reported annually (Chang, Y.J. pers. comm.). Two time series were provided. Fleet F10 is Taiwanese longline catch data and fleet F11 is all other Taiwanese data. French Polynesia annual catch data from 1990 to 2019 from their longline fleet. In addition, four other catch time series were included in the base-case model, all reported annually. Fleet F12 is all other longline gear from the WCPFC and IATTC. Fleets F14 and F15 are purse seine catch data from the WCPFC and EPO, respectively. All other blue marline catch from the WCPFC are included in F16, which includes gear types such as handline, troll, and harpoon.

### *Relative Abundance Indices*

The four CPUE indices available for inclusion in the model were assigned to a quarter based upon the recommendations of the country providing the index and are assumed to represent the quarter in which the highest catches take place for each fishery. Japanese longline fleets (S1 and S2) were all assigned to quarter 1; the US longline fleet (S3) was assigned to quarter 4; the Taiwanese longline fleet (S4) was assigned to quarter 1. The Taiwanese CPUE was standardized as three separate indices due to changes in the logbook reporting for the fleet, with breaks in 1979 and 1999. The index was treated as a single fleet in the base-case model with time-varying catchability estimated for each time period.

An a priori analysis of the CPUE indices was conducted to evaluate the potential for conflict within the indices. The analysis was performed using the diags component of the FLCore package (Version 2.6.6, Kell *et al.* 2007) in R (version 3.4.0, R Core Team, 2021). These packages provide a standardized method to plot and summarize CPUE data so that modelers can better evaluate their input data into assessment models. Each CPUE index was fit using a Loess smoother with only year as an explanatory variable, and the residuals from that smoother were examined graphically. A pairwise correlation analysis was used to evaluate similarities and discrepancies in the trends of each pair of indices. A hierarchical clustering analysis using a set of dissimilarities was conducted to identify significant clusters of indices. Finally, a cross-correlation analysis was performed to evaluate strong year class trends, which may appear in fleets if they are targeting different age classes. The CPUE indices were assumed to be linearly proportional to biomass where catchability (q) was assumed to be constant and occur in the first month of the quarter assigned.

The CVs for each CPUE index were assumed to be equal to the SE on the log scale. The minimum CV was scaled to a minimum of 0.2 or the square root of the residual variance (RSME) of what we would expect the assessment model to fit the CPUE index at best by adding a constant to each CV value. This was calculated as the square root of the residual variance of a loess smoother fit to each index (Francis 2011, Lee et al., 2014).

$$RSME_{smoother} = \sqrt{(\frac{1}{N})\sum_{t=1}^{N}(Y_t - \hat{Y}_t)^2}$$

where $Y_t$ is the observed CPUE in year *t* on the log scale, $\hat{Y}_t$ is the predicted CPUE in year *t* from the smoother fit to the data on the log scale, and N is the number of CPUE observations. RSME values for each index are listed in Table 4.If the input SE was greater than these values, it was left unchanged.

### *Size Composition Data*

Eight size composition time series were provided for consideration in the 2021 SS assessment model; seven as length composition data and one as weight composition data. Japanese length composition data were provided for the offshore and distant-water longline (EarlyLL and LateLL) fleets in 5 cm bins from 1971-1993 (early) and 1994-2019 (late, Figure *3*). Weight composition data were provided for the high-seas and coastal drift net fisheries in 5kg size bins from 1977 to 2003 (Ijima, 2021b, Figure *4*). US length composition data in 5 cm bin size were provided for the Hawaiian longline fleet from 1994 to 2019 (Ito and Sculley, 2021). Taiwan length composition data were provided in 5 cm bin size from 1981 to 2019 for the distant-water longline fleet (Chang*,* Y.J. Pers. Comm.). However, due to difficulties in estimating the selectivity of the early length composition data and the poor data quality of samples measured in this period, Taiwanese length composition data prior to 2005 were excluded from the base-case model. The EPO purse seine fishery also provided length composition data from 1990 to 2019 in 5 cm bin size. Two fleets provided length composition data in 10 cm bin size from 1992 to 2019: French Polynesia longline fleet and all other longline fleets (Ijima, 2021c).

Length composition data were available in quarterly time steps (Figure 3 - Figure *5*).

Quarters with fewer than 25 total samples were removed from the time series due to limited sample size, as agreed upon in the BILLWG data preparatory meeting. In addition, the length composition data for F13 were excluded as only represent a small component of the fishery and required 12 parameters to fit adequately the data (double normal selectivity with 3 time blocks). Data were fit using a multinomial error structure. Length composition data were weighted using the 2-stage process based upon the Francis (2011) method. In the first stage, the effective sample size was scaled to a mean of 25 by multiplying each number of samples by a constant. The second stage weighting was attempted based upon the T.A1.8 equation (Francis 2011) as calculated by the model using r4ss, an R package for plotting SS results (R version 3.4.0, R Core Team, 2017, r4ss version 1.28.0, Taylor et al., 2017).

## *Base-case model description*

The assessment was conducted with Stock Synthesis (SS) version 3.30.16.00-SAFE released 09/03/2020 using Otter Research ADMB 12.2 by Richard Methot (Methot and Wetzel, 2013). The model was set up as a single area model with two sexes and four seasons (quarters). Spawning was assumed to occur in May (month 5) while recruitment was assumed to occur in July (month 7). Age at recruitment was calculated based upon the model estimated average selectivity at age based upon the quarterly selectivity at length. Sex-specific biological parameters were used, with sex- and age-specific natural mortality (Table *2* エラー！参照元が見つかりません。) as agreed upon in the BILLWG Data Preparatory Meeting (ISC Report 2021). In the old growth model, the maximum age of BUM was set to 26, the age at length $L_1$ is set to age 1, the CV of the growth curve was set to 0.14 for young fish and 0.15 for old fish for females and 0.14 for young fish and 0.1 for old fish for males, and the sex ratio at birth was assumed to be 1:1. The growth curve used a von Bertalanffy growth curve for ages 1-26 with a K = 0.107 for females and 0.211 for males, and an $L_{inf}$ = 304cm EFL and 226 cm EFL for females and males with the size at age 1 = 144 cm EFL for both sexes. In the new growth model, the maximum age of BUM was set to 20, the age at length $L_1$ is set to age 0.5 to better match the two-stanza growth curve, the CV of the growth curve was set to 0.13 for young fish and 0.15 for old fish for females and 0.2 for young fish and 0.1 for old fish for males, and the sex ratio at birth was assumed to be 1:1. The growth curve used a Richards growth curve for ages 0.5-20 with a K = 0.31 for females and 0.18 for males, and an $L_{inf}$ = 249 cm EFL and 206 cm EFL for females and males with the size at age 0.5 = 136 cm EFL for both sexes. Both models used a Beverton-Holt spawner-recruit relationship with steepness (h) estimated with a strong normal prior with a mean of 0.87 and SD of 0.05

and sigmaR ($\sigma_r$) rescaled to 0.4.

Nineteen fleets were included in the model, 16 catch fleets and 3 survey fleets. Initial fishing mortality was estimated for F1. Main recruitment deviations were estimated from 1975-2019. The recruitment deviations were bias-adjusted based upon the estimates from Methot and Taylor (2011). Early recruitment deviations were estimated from 1965 to 1975 as the population was not at equilibrium prior to the start of the model. 1965-1971 was the "ramp-up" period where the bias adjustment of $\sigma_r$ was 0 at the beginning of the period and increased linearly to the maximum bias adjustment of 0.83 in 1971. Full bias adjustment was from 1983-2018.

The population model and the fishery length data had 55 five cm length bins from 30-320+ cm. The population had 21 or 27 annual ages from age 0 to 20 or 26 for the old and new growth models, respectively. There were no age data. Fishery size data were used to estimate selectivity patterns, which controlled the size distribution of the fishery removals. Three different selectivity patterns were used based upon the best fit to the size composition data and CPUE indices. Japanese longline F2 and US Hawaii LL F7 used three-parameter cubic spline selectivity patterns. Japanese longline F1 used a four-parameter cubic spline pattern. The EPO purse seine fleet F14 used an asymptotic logistic selectivity pattern, using a more flexible double normal selectivity pattern resulted in the logistic shape, and therefore the simpler pattern was used for the fleet. All other fleets with size data were estimated as six-parameter double normal (dome-shaped) selectivity patterns. Survey selectivity patterns mirrored their respective catch fleets (Table 3).

Model estimated time series of total biomass (B in metric tons, mt = 1000 kg), age 1+ total biomass ($B_{1+}$ mt), female spawning biomass (SSB mt), and recruitment (R in 1000s of fish) were tabulated on an annual basis. The annual exploitation rate was calculated as Catch/$B_{1+}$. Stock status indicators were calculated based upon MSY and not a target reference level.

### *Convergence Criteria and Diagnostics*

The model was assumed to have converged if the standard error of the estimated parameters could be derived from the inverse of the negative Hessian matrix. Various convergence diagnostics were also evaluated. Excessive CVs (>50%) on estimated parameters would suggest uncertainty in the parameter estimates or model structure. A gradient of >0.001 would suggest poorly fit parameter estimates. The correlation matrix was also evaluated to identify highly correlated (>95%) and non-informative (<0.01) parameters. Parameter estimates hitting bounds of the prior was also indicative of poor model fit.

Several diagnostics were run to evaluate the fit of the model to the data. An Age-Structure Population Model (APSM) was used to evaluate the influence of the length composition data on the population trends (Carvalho *et al.*, 2017). Profiling the likelihood on $R_0$, where the $R_0$ is fixed at a range of values around the maximum likelihood estimate and then the likelihood is estimated, was used to identify influential data components (Lee, *et al.,* 2014). A runs test was used to evaluate randomness in the residuals of the CPUE data (Carvalho, *et al.,* 2021). Finally, residual plots and plots of the observed vs expected data were examined to evaluate goodness-of-fit.

## CPUE Analysis Results

The CPUE time series are plotted in Figure *6*, to compare trends by stock. In general, the US Hawaii longline index appears to have the potential of conflicting with the other CPUE indices, especially the Taiwanese longline fleet. The US index appears to be a one-way trip with the declining trend flattening in the last 4-5 years. Japanese CPUE increases in the early time period (1975-1993) and gradually declines in the late period. The Taiwanese CPUE increases from 1971 to the 1980s, declines again until the late 1990s and has been increasing ever since.

To look at deviations from the overall trends, the residuals from the fits are compared in Figure *7*. This allows for conflicts between indices to be highlighted by patterns in the residuals, autocorrelation within indices identified which may be due to year-class effects, or the identification of other potentially important factors not included in the standardization of the CPUE. Other than large residuals in the US Hawaiian LL, there appear to be no significant patterns in the residuals.

Figure *8* illustrates the correlation between indices; the lower-left triangle displays the pairwise scatter plots of one index plotted against another with a linear smoother, the upper right triangle displays the correlation coefficients, and the diagonal displays the range of observations. A single influential point may cause a strong spurious correlation therefore, it is important to look at the plots as well as the correlation coefficients. Most of the indices have moderate to strong positive correlations. The Japanese early LL appears to be positively correlated with the Taiwanese LL fleet (corr = 0.297) and the Japanese late LL appears to be positively correlated with the Hawaii LL (corr =0.271). The US Hawaiian LL has a high negative and statistically significant correlation with Taiwan (corr = -0.597). Which could indicate potential conflict between these two indices.

If indices represent the same stock components then it is reasonable to expect them to be correlated. If indices were not correlated or negatively correlated, which would indicate that they show conflicting trends, this may result in poor fits to the data and bias in the estimates. Therefore, the correlations can be used to select groups that represent a common hypothesis about the evolution of the stock (Kell *et al.*, 2007). This allows for multiple models to be proposed that may reflect different possible states of nature suggested by the CPUE indices. Figure *9* shows the results from a hierarchical cluster analysis using a set of dissimilarities. Blue indicates positive correlation and red indicates negative correlations. The width of the oval indicates the scale of the correlation. Most series appear to be similar, with the HI LL index and Taiwan DWLL index least similar. A single cluster with Japanese early LL and Taiwanese LL was identified.

Overall, the results of the CPUE analysis suggest that there may be the potential for conflict between the candidate CPUE indices with the HI LL index the most likely candidate for exclusion from the model. Due to this analysis and the strong patterning in the residuals of the CPUE index when included in the base-case model, the US HI LL index was ultimately excluded from the base-case model.

# Old-Growth Results

## *Model fit*

The old growth base-case model ran in about 40 minutes, estimated 96 parameters, and had a total likelihood of 5184.73. The inverse Hessian was positive definite, which allowed for the estimation of parameter standard deviations and suggests that the model converged, and the maximum gradient component was 0.0061 which is greater than 0.001. However, fixing steepness at 0.87, the prior mean, results in a maximum gradient component <0.001. None of the parameter estimates hit a bound, no parameters had correlations above 0.95 and three selectivity parameters had correlations below 0.01. Thirteen of fifteen early recruitment deviations (1960-1974) and 34 of 44 (78%) of the main recruitment deviations had CVs > 50%. Five of 31 selectivity parameters had CVs

>50%. These parameters were from the dome-shaped selectivity functions and three were the width of the plateau. All of the parameters below the threshold for uncorrelated parameters also had CVs > 50%.

Fits to the abundance indices were relatively good, with no substantial divergences between the expected and estimated CPUEs except for the Japan Early LL index (Figure *10* - Figure *12*). However, the Japan Late LL index was the only index to pass the runs test (Figure *13*), which indicates that the residuals are not random.

Estimated selectivity for each fleet are in Figure *14* and Figure *15*. Fits to the length composition data were also relatively good (Figure *16* -Figure *21*), although several problems are evident in the fitting to the Japanese early LL length composition data (F1). The residual pattern for F1 show large positive and negative residuals for the whole time series, and the annual mean length is estimated to be larger than observed (Figure *18*). Large positive residuals are also observed at the beginning of the Japanese late LL fleet (F2), where a block has already been used to estimate a separate selectivity from the rest of the time series. Also, in most of the fisheries with size composition data, a pulse of small fish appear in the fisheries in the early 2000s and again after 2015, which suggests strong recruitment at those times.

Model estimates of age 1+ biomass show an initial decrease in biomass from 1971 to 1987, then biomass increased to 1991, declined to its lowest level in 2006, and has varied since. Biomass is on an increasing trend in the last three years of the assessment model (Figure *22*). Initial female spawning stock biomass was estimated to be approximately 35,000 mt and virgin SSB was around 147,000 mt (Figure *23*). Annual fishing mortality is reported as the average for fish ages 1-10 (Figure *24*). Fishing mortality was below MSY for all except 8 years and has been well below $F_{MSY}$ in the last three years. Recruitment was variable but the log of the deviations were generally between 0.4 and -0.4 (Figure *25*). Current depletion, as estimated as the age 1+ biomass in 2016 compared to the virgin age 1+ biomass was estimated to be 0.16.

### *Diagnostics*

Profiling on $R_0$ showed that the recruitment estimates were highly influential in the model results, especially below the MLE estimate (Figure 26 -Figure *28*). Results from the ASPM model showed the same population trend as the full model, especially after 1990, but estimated a slightly smaller initial spawning stock biomass, 114,000 mt compared to 147,000 mt in the full model (Figure *29*). This suggested that the length composition data did not have substantial conflict with the abundance indices but did scale the population size slightly and the abundance indices are driving the population trend.

## New-Growth Results

### *Model fits*

The new growth base-case model ran in about 30 minutes, estimated 95 parameters, and had a total likelihood of 4485.74. The inverse Hessian was positive definite, which allowed for the estimation of parameter standard deviations and suggests that the model converged, and the maximum gradient component was 0.00073 which is greater than 0.001, however, this due to estimating steepness within the model. For model runs with a steepness fixed at 0.87 (the prior mean), the maximum gradient is <0.0001. None of the parameter estimates hit a bound, no parameters had correlations above 0.95, and four selectivity parameters had correlations below 0.01. Six of ten early recruitment deviations (1965-1974) and 21 of 54 (54%) of the main recruitment deviations had CVs > 50%. Five of 31 selectivity parameters had CVs >50%. Four parameters were from the dome-shaped selectivity functions and were the width of the plateau. All of the parameters below the threshold for uncorrelated parameters also had CVs > 50%.

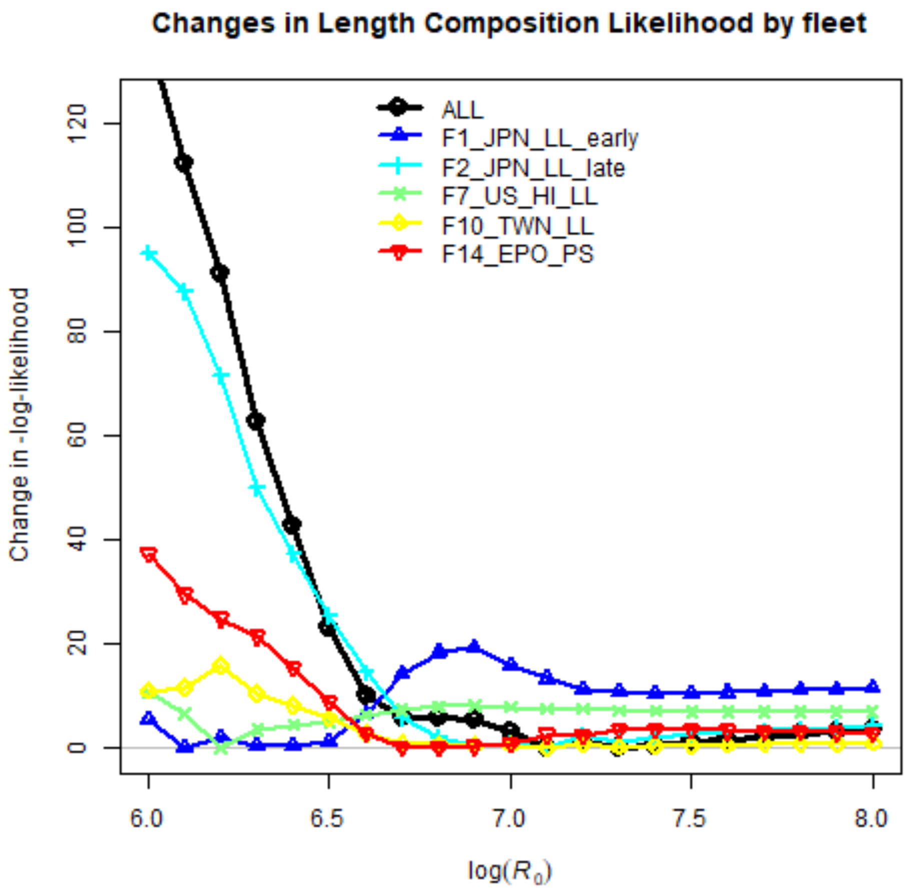
Fits to the abundance indices were relatively good, with no substantial divergences between the expected and estimated CPUEs except for the Japan Early LL index (Figure *30* -Figure *32*). However, the Japan Late LL index was the only index to pass the runs test (Figure *33*), which indicates that the residuals are not random.

Estimated selectivity for each fleet are in Figure *34* and Figure *35*. Fits to the length composition data were also relatively good (Figure *36* -Figure *41*). The residual patterns for F1 show more small fish caught than expected from 1971-1975 but fit the rest of the data well (Figure *38*). The residual pattern for F2 shows more large fish caught in the early 2000s when expected, a time period that is already estimated as a separate selectivity time block. Most of the fleets show a pattern of larger than expected number of small fish entering the fishery in the early 2000s and to a lesser extent, after 2015. This suggests that there are periodic strong recruitments.

Model estimates of age 1+ biomass show an initial decrease in biomass from 1971 to 1988, then biomass increased to 1992, declined to its lowest level in 2005, and has varied since. Biomass has been increasing in the last three years of the assessment model (Figure 42). Initial female spawning stock biomass was estimated to be approximately 43,000 mt and virgin SSB was around 86,000 mt (Figure 43). Annual fishing mortality is reported as the average for fish ages 1-10 (Figure 44). Fishing mortality was below MSY for all years and has been well below $F_{MSY}$ since 2005. Recruitment was variable but the log of the deviations were generally between 0.4 and -0.4 (Figure 45). Current depletion, as estimated as the age 1+ biomass in 2016 compared to the virgin age 1+ biomass was estimated to be 0.23.

## *Diagnostics*

Profiling on $R_0$ showed that the size and length composition data were influential in the model results, although close to the MLE CPUE was also an important contributor (Figure 46) The model indicated more confidence in the lower bound of the population size, and the likelihood profile at values above MLE are relatively flat. Some conflict between length composition data does exist




(

Figure 47), and the Japanese late LL index suggests a smaller R0 than the other two indices

(Figure *48*). Results from the ASPM model showed the same population trend as the full model but estimated a slightly smaller initial spawning stock biomass, 80,000 mt compared to 86,000 mt in the full model (Figure *49*). This suggested that the length composition data did not have substantial conflict with the abundance indices but did scale the population size slightly and the abundance indices are driving the population trend after 1990.

# Conclusions

Fits to both growth models are overall relatively similar. The new growth model appears to be more optimistic with no years below $SSB_{MSY}$ or above $F_{MSY}$ but estimates an overall smaller population size (Figure 50). This is likely due to the smaller $L_{inf}$ for males and females, which would tell the model there is less biomass of very large individuals. The old growth model has years that occasionally go below $B_{MSY}$ or above $F_{MSY}$ but estimates a larger population size (Figure 51). Both models indicate that the stock is not overfished and overfishing is not occurring, and if the trend of the last 3-4 years continues, the stock will not become overfished and overfishing will not occur. Diagnostics for both models indicate that the population trend is being driven by the length composition data and the CPUE indices, which makes both models good candidates for the 2021 Pacific blue marlin base-case assessment model.

# References


Andrews, A.H., Humphreys Jr, R.L. and Sampaga, J.D., 2018. Blue marlin (*Makaira nigricans*) longevity estimates confirmed with bomb radiocarbon dating. Canadian Journal of Fisheries and Aquatic Sciences, 75(1), 17-25.

Brodziak, J., 2013. Combining Information on Length-Weight Relationships for Pacific Blue Marlin. ISC/13/BILLWG-1/01

Carvalho, F., A. E. Punt, Y.-J. Chang, M. N. Maunder and K. R. Piner (2017). Can diagnostic tests help identify model misspecification in integrated stock assessments? Fisheries Research 192: 28-40.

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M. Schirripa, M., Kitakado, T., Yemane, D., Piner, K.R., Maunder, M.N., Taylor, I., Wetzel C.R., Doering, K. Johnson, K.F. and Methot, R.D. (2021). A cookbook for using model diagnostics in integrated stock assessments. Doi: 10.1016/j.fishres.2021.105959

Chang, Y.J., Lee, H.H., Brodziak, J., DiNardo, G., and Sun, C.L., 2013. A Bayesian hierarchical meta-analysis of blue marlin (*Makaira nigricans*) growth in the Pacific Ocean. ISC/13/BILLWG-1/02

Chang, Y.J., Shimose, T., Kanaiwa, M., Chang, X.B., Masubuchi, T., Yamamoto, A., Kanaiwa, M., 2020. Estimation of the two-stanza growth curves with ageing uncertainty for the Pacific blue marlin (*Makaira nigricans*). ISC/20/BILLWG-02/03

Francis, R. I. C. C. (2011). Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68(6): 1124-1138.

Hsu, J. and Chang, Y.J. (2021). CPUE standardization of blue marlin (Makaira nigricans) for the Taiwanese distant-water tuna longline fishery in the Pacific Ocean during 1971 – 2019. ISC/20/BILLWG-03/03.

Ijima, H. (2021a) Update Japanese longline abundance index of Pacific blue marlin (*Makaira nigricans*) estimated by the habitat model. ISC/20/BILLWG-03/01.

Ijima, H. (2021b). Update Japanese catch and length-frequency data of Pacific blue marlin (Makaira nigricans) during 1971-2018. ISC/20/BILLWG-03/05.

Ijima, H. (2021c). Update of Pacific blue marlin (Makaira nigricans) catch and size statistics from the WCPFC and the IATTC. ISC/20/BILLWG-03/08.

ISC BILLWG. (2021). Report of the Billfish Working Group Workshop. 6-7th, 10, 13th November 2020.

Ito, R.Y. and Sculley, M. (2021). U.S. Commercial fisheries for marlins in the North Pacific Ocean. ISC/20/BILLWG-03/07.

Kell, L. T.; Mosqueira, I.; Grosjean, P.; Fromentin, J-M.; Garcia, D.; Hillary, R.; Jardim, E.; Mardle, S.; Pastoors, M.A.; Poos, J.J.; Scott, F.; and Scott, R.D. (2007). FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64 (4): 640-646. Doi: 10.1093/icesjms/fsm012

Lee, H.H and Chang, Y.J., 2013. Age-structured Natural Mortality for Pacific Blue Marlin Based on Meta-analysis and an Ad Hoc Model. ISC/13/BILLWG-1/07

Lee, H.-H., K. R. Piner, R. D. Methot Jr and M. N. Maunder (2014). Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: An example using blue marlin in the Pacific Ocean. Fisheries Research 158: 138-146.

Methot, R.D.; A'mar, T.; Wetzel, C.; and Taylor, I. (2017). Stock Synthesis User Manual Version 3.30.05-3.30.08. November 14, 2017.

Methot, R. D. and I. G. Taylor (2011). Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68(10): 1744-1760.

Methot Jr, R. D. and C. R. Wetzel (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142: 86-99.

R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sculley M. (2021). Standardization of Pacific Blue Marlin Catch Per Unit Effort in the Hawaii Longline Fishery from 1995-2019. ISC/20/BILLWG-03/04.

Sun, C.L., Chang, Y.J., Tszeng, C.C., Yeh, S.Z., and Su, N.J., 2009. Reproductive biology of blue marlin (Makaira nigricans) in the western Pacific Ocean. Fishery Bulletin. 107:420-432.

Taylor, I.G.; Stewart, I.J.; Hicks, A.C.; Garrison, T.M.; Punt, A.E.; Wallace, J.R.; Wetzel, C.R.; Thorson, J.T.; Takeuchi, Y.; Ono, K.; Monnahan, C.C.; Stawitz, C.C.; A'mar, Z.T.; Whitten, A.T.; Johnson, K.F.; Emmet, R.L.; Anderson, S.C.; Lambert, G.I.; Stachura, M.M; Cooper, A.B.; Stephens, A.; and Klaer, N. (2017). r4ss package: R Code for Stock Synthesis. Version 1.28.0. URL https://github.com/r4ss

# Tables and Figures

Table 1. List of fleets with Catch and CPUE indices provided for the 2021 Pacific Blue Marlin Stock Assessment and the source for more information about the standardization of the CPUE series and the size composition data.

| Fleet No | Fleet name | Fishing countries | Gear types | Catch units | Size data | CPUE | Source |
|---|---|---|---|---|---|---|---|
| F1 | JPNEarlyLL | Japan | Offshore and distant water longline | B | Y | S1_JPN_DW&OS LL | ISC/20/BILLWG-03/01, ISC/20/BILLWG-03/05 |
| F2 | JPNLateLL | Japan | Offshore and distant water longline | B | Y | S2_JPN_DW&OS LL | ISC/13/BILLWG-1/05, ISC/20/BILLWG-03/05 |
| F3 | JPNCLL | Japan | Coastal longline | B | N (Mirror to F2) | N | ISC/20/BILLWG-03/05 |
| F4 | JPNDRIFT | Japan | High-sea large-mesh driftnet and coastal driftnet | B | Y | N | ISC/20/BILLWG-03/05 |
| F5 | JPNBait | Japan | Bait fishing | B | N (Mirror to F4) | N | ISC/20/BILLWG-03/05 |
| F6 | JPNOth | Japan | Other gears | B | N (Mirror to F2) | N | ISC/20/BILLWG-03/05 |
| F7 | HWLL | USA (Hawaii) | Longline | B | Y | S3_HW_LL – Excluded | ISC/20/BILLWG-03/04, ISC/20/BILLWG-03/07 |
| F8 | ASLL | USA (American Samoa) | Longline | B | N (Mirror to F7) | N | ISC/20/BILLWG-03/07 |
| F9 | HWOth | USA (Hawaii) | Troll and handline | B | N (Mirror to F7) | N | ISC/20/BILLWG-03/07 |
| F10 | TWNLL | Taiwan | Distant-water longline | B | Y | S4_TW_DWLL, S4_TW_DWLL, and S4_TW_DWLL | ISC/20/BILLWG-01/08, ISC/20/BILLWG-03/03 |

Table 1. Continued

| Fleet No | Fleet name | Fishing countries | Gear types | Catch units | Size data | CPUE | Source |
|---|---|---|---|---|---|---|---|
| F11 | TWNOth | Taiwan | Offshore longline, coastal longline, gillnet, harpoon, and others | B | N (Mirror to F10) | N | ISC/20/BILLWG-01/08 |
| F12 | OthLL | Various flags | Longline | B | Y | N | ISC/20/BILLWG-03/08 |
| F13 | PYFLL | French Polynesia | Longline | B | Y – Excluded Mirror F12 | N | ISC/20/BILLWG-03/08 |
| F14 | EPOPS | Various flags | Purse seine | B | Y | N | ISC/20/BILLWG-03/08 |
| F15 | WCPFCPS | Various flags | Purse seine | B | N (Mirror to F14) | N | ISC/20/BILLWG-03/08 |
| F16 | WCPFCOth | Various flags | Troll, handline, and harpoon and others | B | N (Mirror to F14) | N | ISC/20/BILLWG-03/08 |

Table 2. Key life history, recruitment, and selectivity parameters for the blue marlin old growth model and the new growth model. From Table 2 in the ISC BILLWG Data Preparatory report (2021).

| Parameter | Old Growth | New Growth | Reference |
|---|---|---|---|
| Growth_Age_for_L1 | 1 | 0.5 | Chang et al. (2013), Chang et al (2020) |
| Growth_Age_for_L2 | 26 | 20 | Chang et al. (2013), Chang et al (2020) Andrews (2018) |
| NatM_Fem_GP_1 | $M_0$ =0.42, $M_1$ = 0.37, $M_{4+}$ =0.22 | $M_0$ = 0.41, $M_1$ = 0.35, $M_2$ = 0.33, $M_3$ = 0.32, $M_{4+}$ = 0.3 | Lee and Chang (2013), Brodziak (ISC/21/BILLWG-01/03) |
| L_at_Amin_Fem_GP_1 | 144 | 136.13 | Chang et al. (2013) , Chang et al (2020) |
| L_at_Amax_Fem_GP_1 | 304.178 | 249.1 | Chang et al. (2013), Chang et al (2020) |
| VonBert_K_Fem_GP_1 | 0.107 | 0.31 | Chang et al. (2013), Chang et al (2020) |
| Richards_Fem_GP_1 | NA | 0.000468 | Chang et al (2020) |
| CV_young_Fem_GP_1 | 0.14 | 0.13 | Chang et al. (2013), Chang et al (2020) |
| CV_old_Fem_GP_1 | 0.15 | 0.15 | Chang et al. (2013), Chang et al (2020) |
| NatM_Mal_GP_1 | $M_0$ = 0.42, $M_{1+}$ = 0.37 | $M_0$ = 0.41, $M_{1+}$ = 0.35 | Lee and Chang (2013), Brodziak (ISC/21/BILLWG-01/03) |
| L_at_Amin_Mal_GP_1 | 144 | 136.13 | Chang et al. (2013), Chang et al (2020) |
| L_at_Amax_Mal_GP_1 | 226 | 206.4 | Chang et al. (2013), Chang et al (2020) |
| VonBert_K_Mal_GP_1 | 0.211 | 0.18 | Chang et al. (2013), Chang et al (2020) |
| Richards_Mal_GP_1 | NA | 0.000468 | Chang et al (2020) |
| CV_young_Mal_GP_1 | 0.14 | 0.2 | Chang et al. (2013), Chang et al (2020) |
| CV_old_Mal_GP_1 | 0.1 | 0.1 | Chang et al. (2013), Chang et al (2020) |
| Wtlen_1_Fem | 1.84E-05 | 1.84E-05 | Brodziak 2013 |
| Wtlen_2_Fem | 2.956 | 2.956 | Brodziak 2013 |
| Mat50%_Fem | 179.76 | 179.76 | Sun et al. (2009) |
| Mat_slope_Fem | -0.2039 | -0.2039 | Sun et al. (2009) |
| Fecundity | Proportional to spawning biomass | Proportional to spawning biomass | Sun et al. (2009) |
| Wtlen_1_Mal | 1.37E-05 | 1.37E-05 | Brodziak 2013 |
| Wtlen_2_Mal | 2.975 | 2.975 | Brodziak 2013 |
| Spawning season | 2 | 2 | Sun et al. (2009) |
| R0 | 0.6 | 0.4 | Rescaled |
| Steepness | 0.9 | 0.9 | Estimated |

Table 3. Selectivity patterns used for length and size composition data in the 2021 Pacific blue marlin stock assessment base-case models.

| Fleet | Selectivity |
|---|---|
| F1 JPN LL Early | 4-parameter cubic spline |
| F2 JPN LL Late | 3-parameter cubic spline |
| F3 JPN CLL | Mirror F2 |
| F4 JPN DRIFT | Double normal |
| F5 JPN Bait | Mirror F2 |
| F6 JPN Oth | Mirror F2 |
| F7 HW LL | 3-parameter cubic spline |
| F8 AS LL | Mirror F7 |
| F9 HW Oth | Mirror F7 |
| F10 TWN LL | Double normal |
| F11 TWN Oth | Mirror F10 |
| F12 Oth LL | Double Normal |
| F13 PYF LL | Mirror F12 |
| F14 EPO PS | Asymptotic logistic |
| F15 WCPFC PS | Mirror F14 |
| F16 WCPFC Oth | Mirror F14 |

Table 4. Mean CV, calculated RSME, and additional variance added in the base-case models for the three CPUE Indices.

| Fleet | Model | Mean CV | RSME | Added Variance |
|---|---|---|---|---|
| S1 JPN LL Early | Old Growth | 0.20 | 0.19 | 0.00 |
| S2 JPN LL Late | Old Growth | 0.20 | 0.42 | 0.22 |
| S4 TWN LL | Old Growth | 0.33 | 0.42 | 0.08 |
| S1 JPN LL Early | New growth | 0.20 | 0.25 | 0.05 |
| S2 JPN LL Late | New growth | 0.20 | 0.38 | 0.18 |
| S4 TWN LL | New growth | 0.33 | 0.48 | 0.14 |

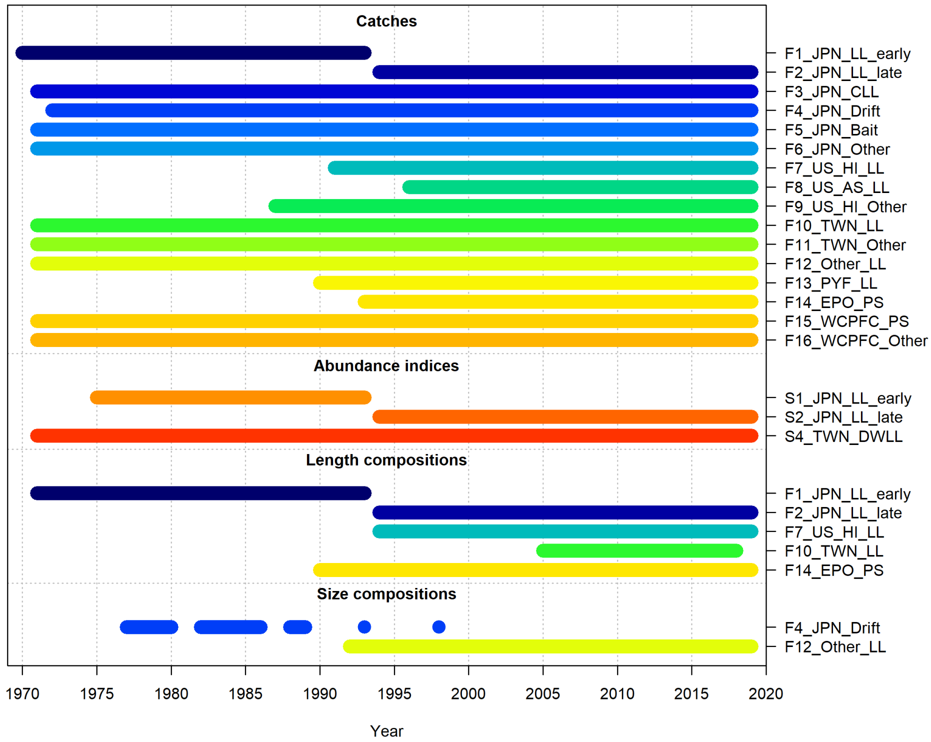



Figure 1. Catch, CPUE index, and size composition data included in the 2021 BUM stock assessment.

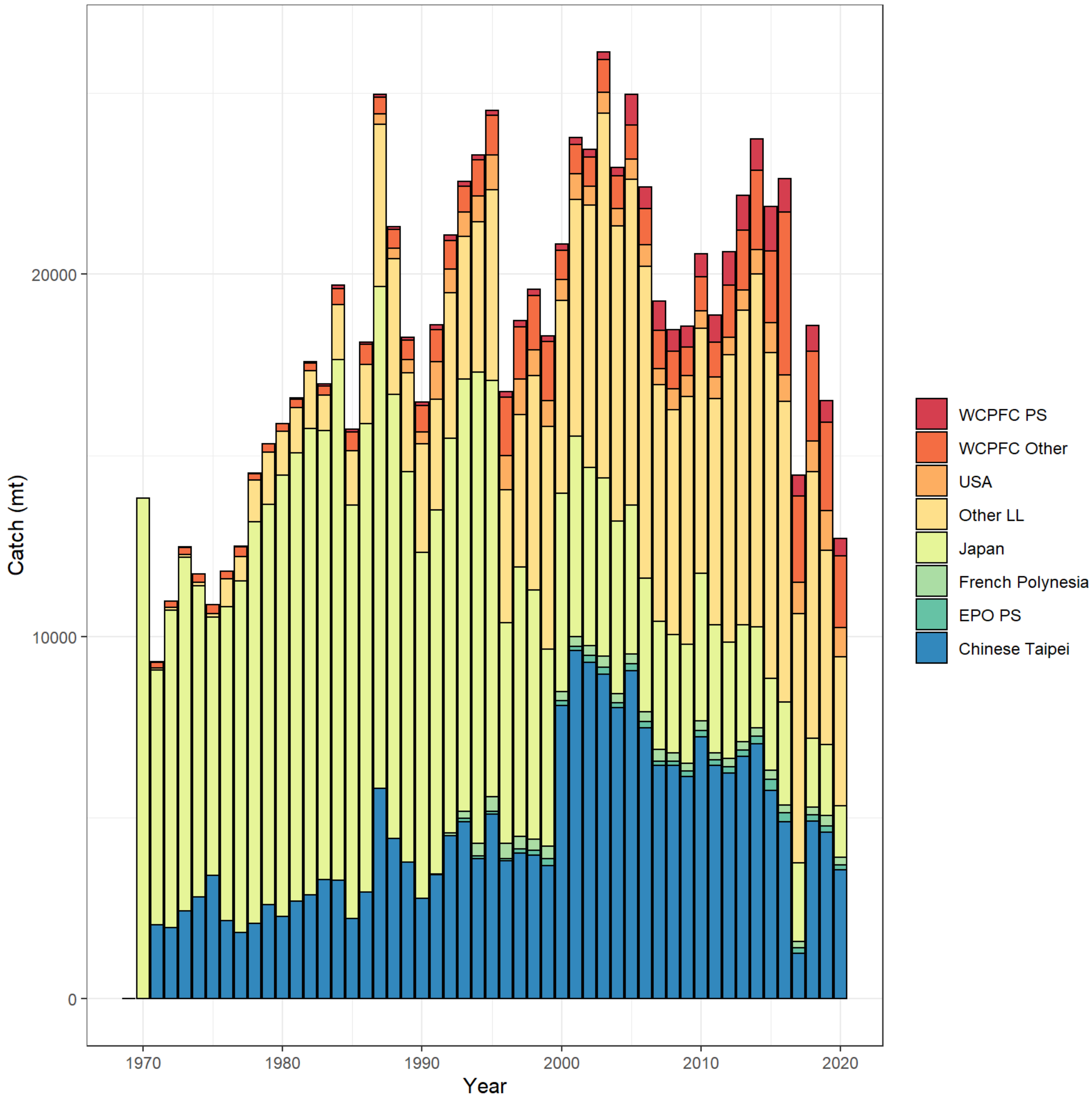



Figure 2. Annual catch of Pacific blue marlin by country or RFMO and gear used in the 2021 base-case assessment model.

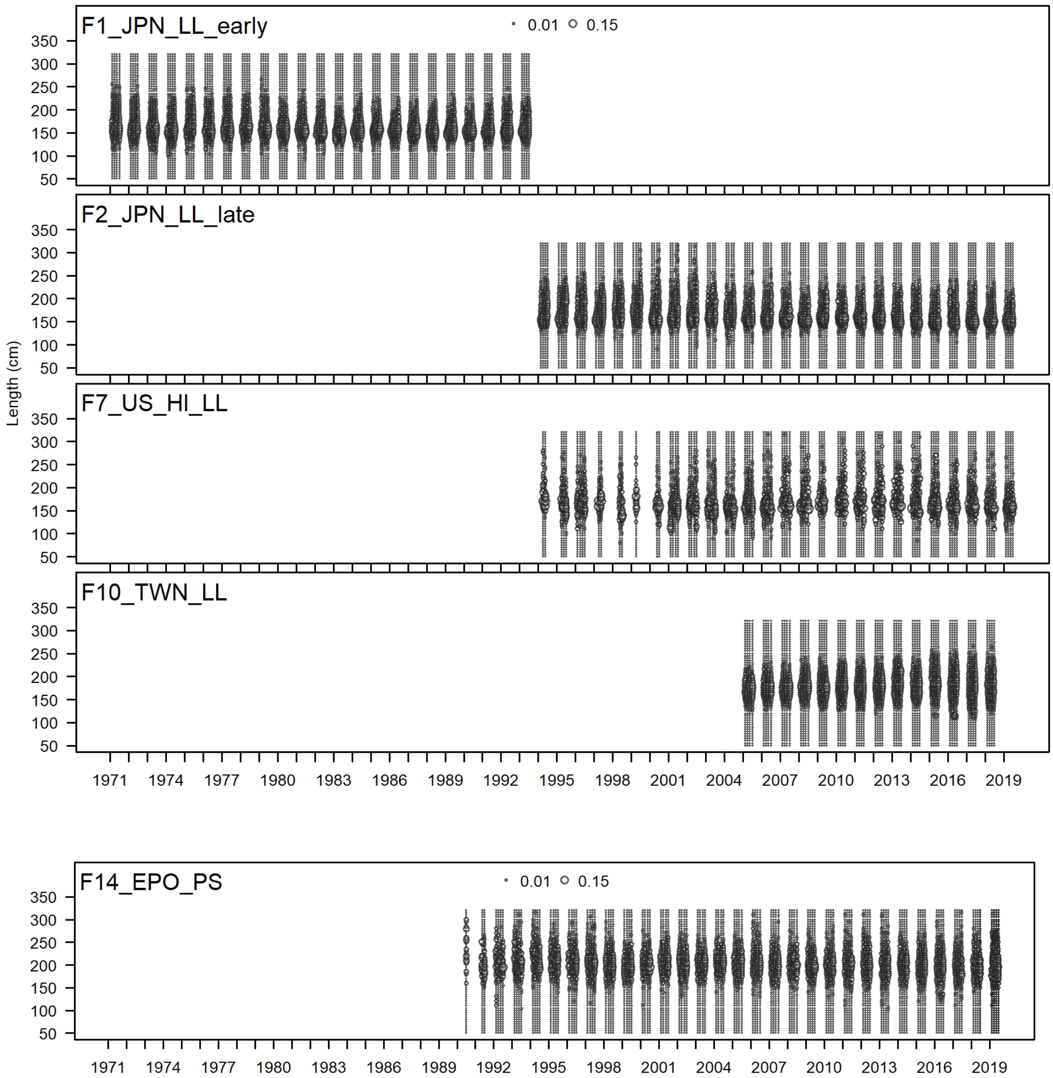



Figure 3. Length Composition data available in 5cm size bins for the 2021 Pacific blue marlin stock assessment.

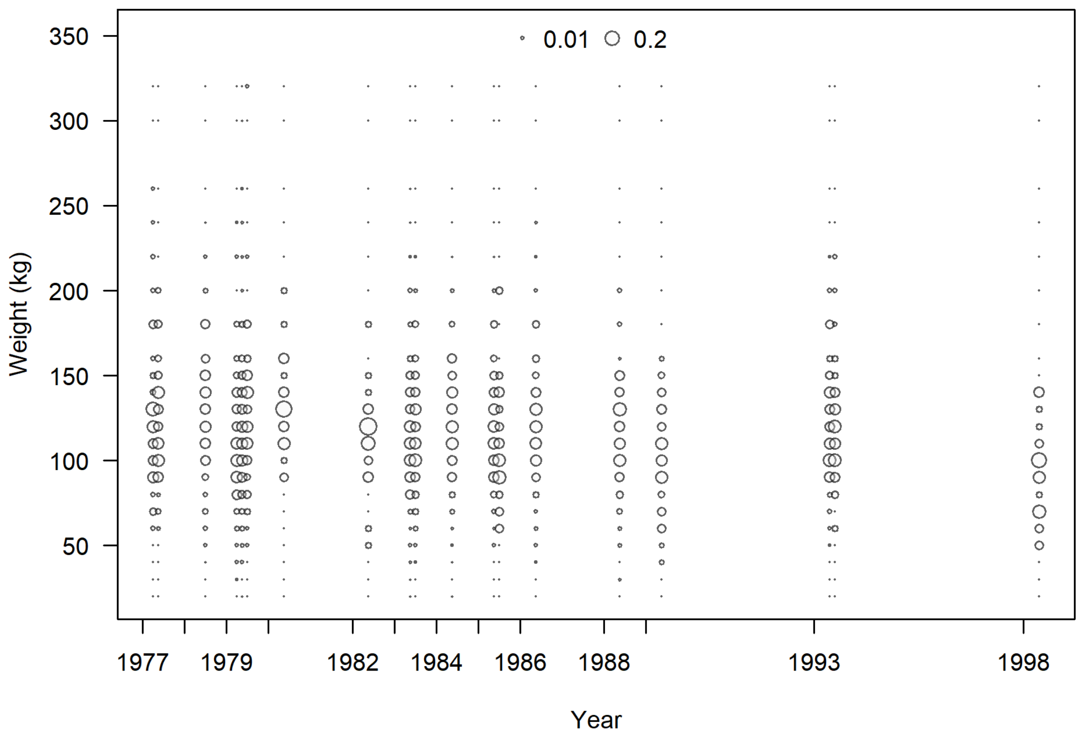



Figure 4. Weight composition data available from F4 JPNDRIFT in 5kg bins for the 2021 Pacific blue marlin assessment.

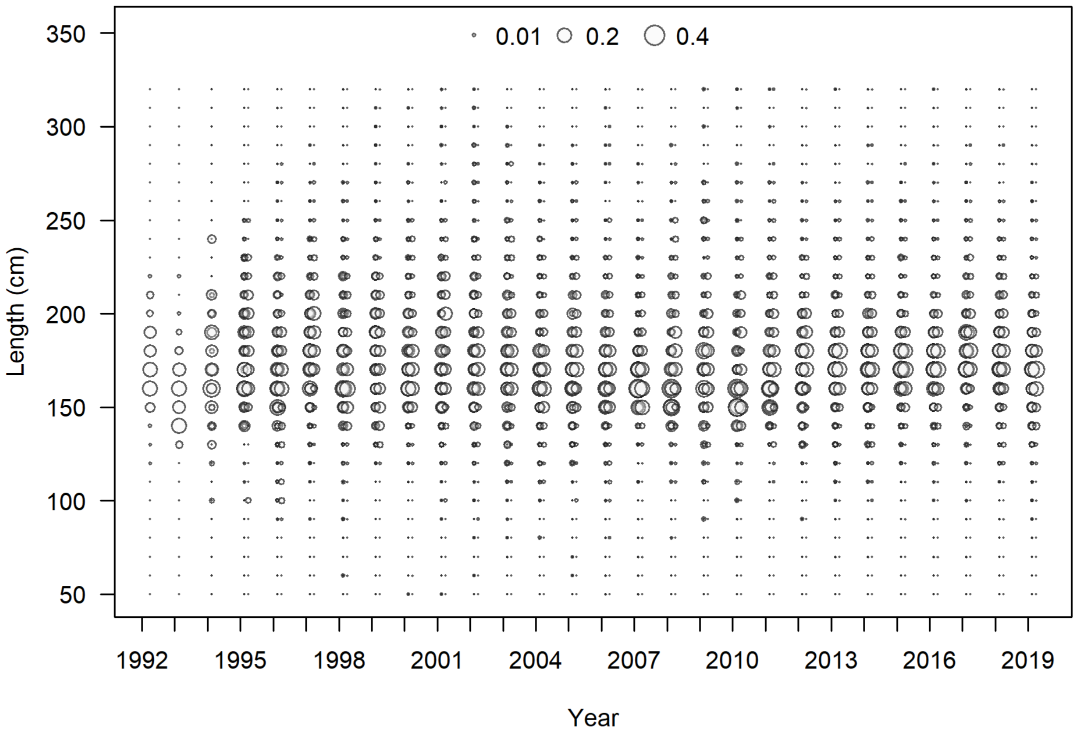



Figure 5. Length composition data available from F12 Other LL in 10cm size bins for the 2021 Pacific blue marlin assessment.

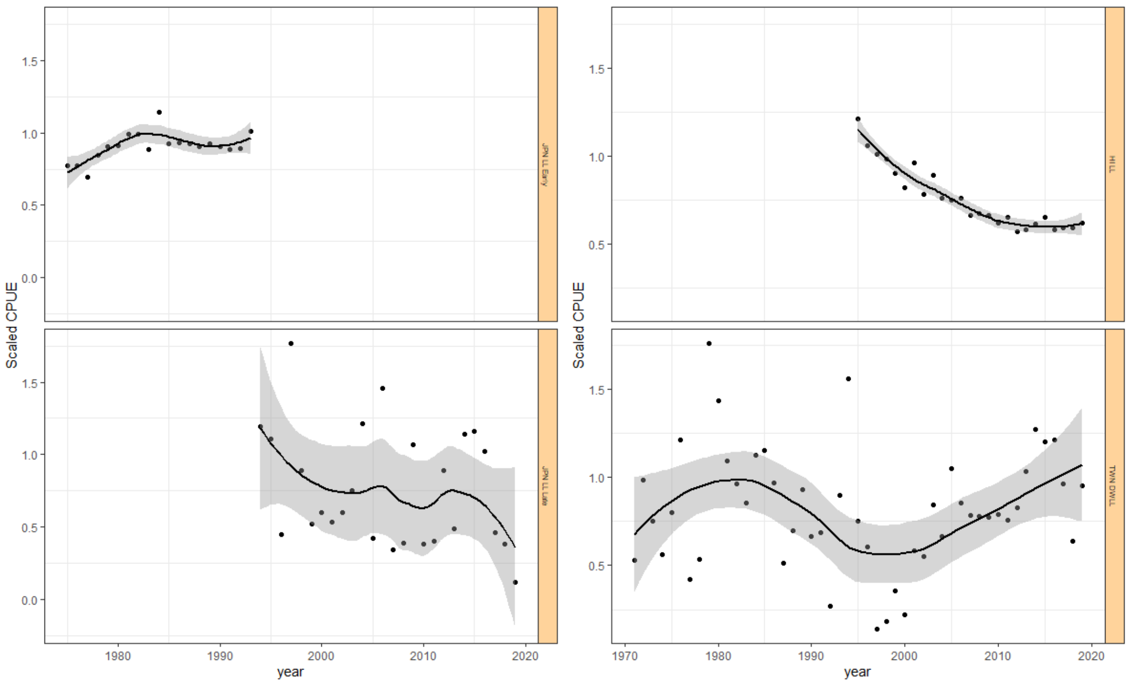



Figure 6. Time series of CPUE indices; continuous black line is a loess smoother showing the average trend by area (i.e. fitted to year for each area with series as a factor). Top left is Japan LL Early, bottom left is Japan LL Late, top right is US HI LL, and bottom right is Taiwan LL.

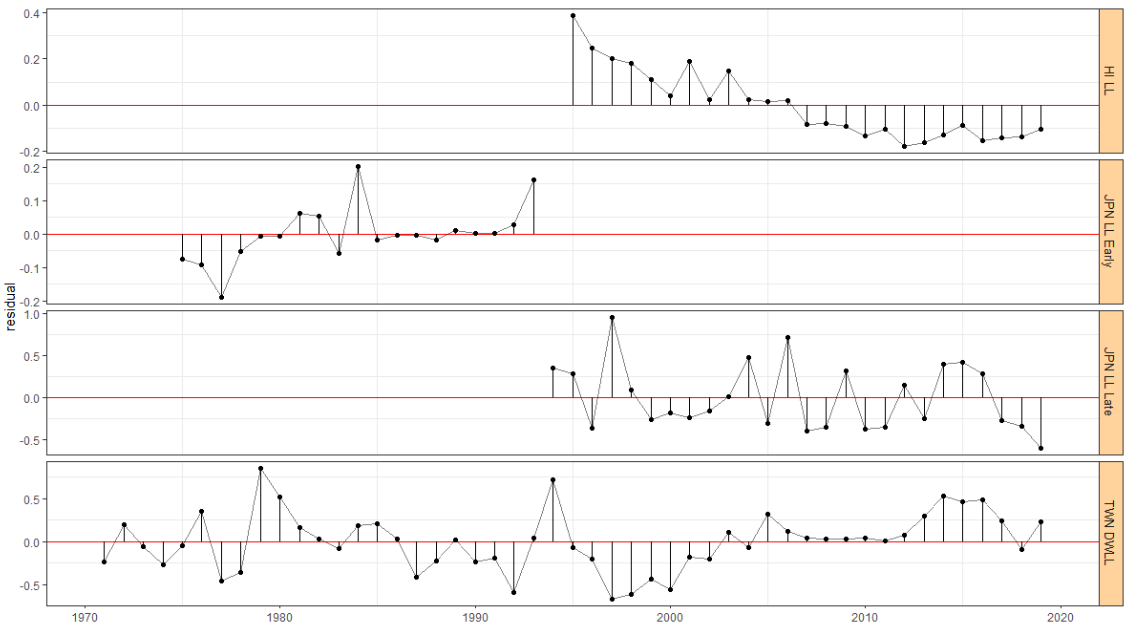



Figure 7. Time series of residuals from the Loess fit.

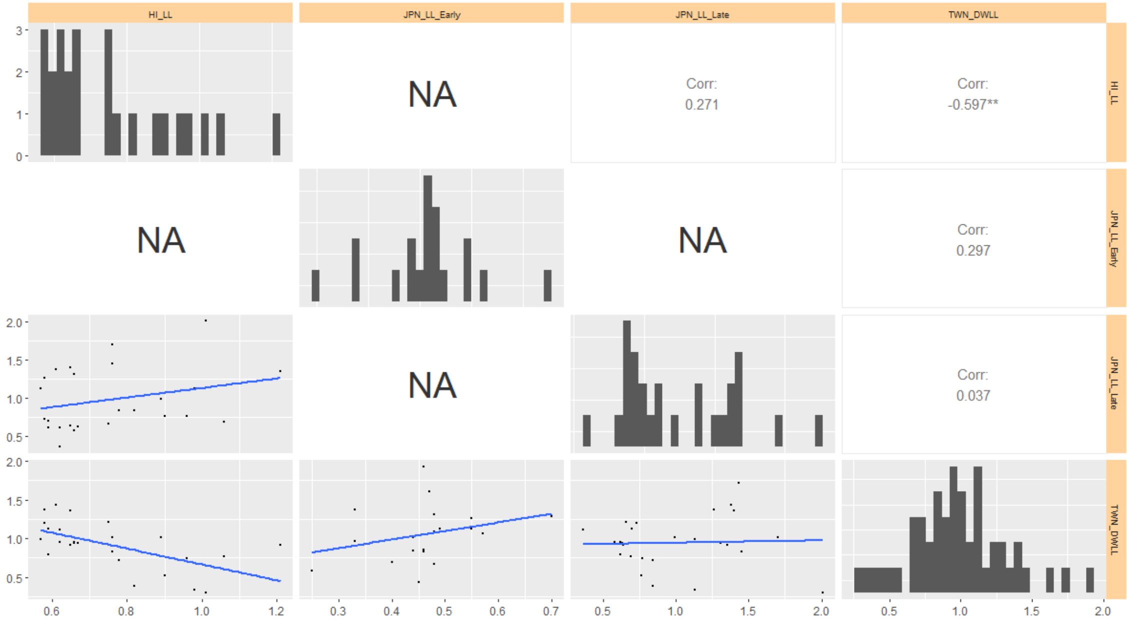



Figure 8. Pairwise scatterplots with blue regression lines (lower left), correlation coefficients (top right), and the range of observations to illustrate correlations among all CPUE indices (central diagonal).

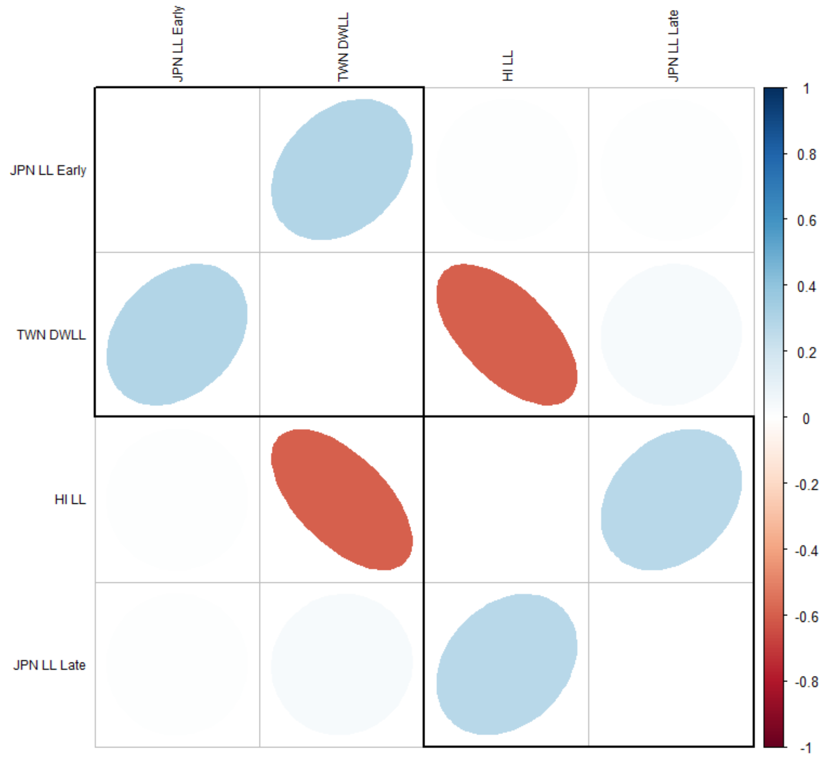



Figure 9. Plot of the correlation matrix for CPUE indices. Blue indicates a positive correlation and red negative. The order of the indices and the rectangular boxes are chosen based on a hierarchical cluster analysis using a set of dissimilarities for the indices being clustered.

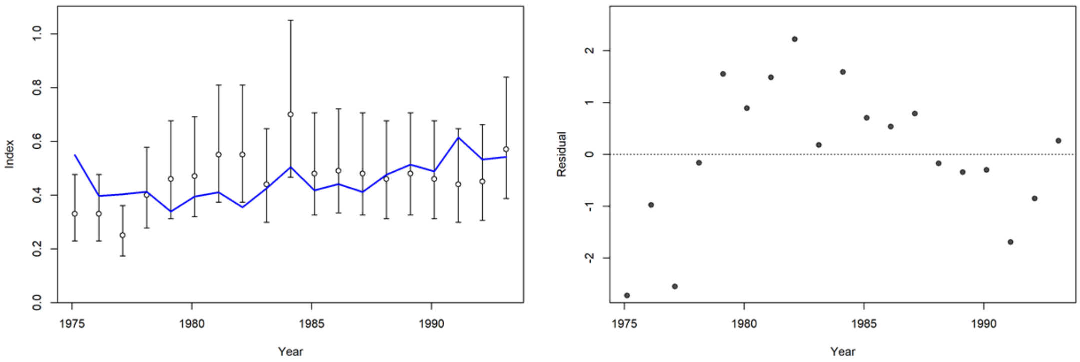



Figure 10. Fit to the S1 Japanese Early LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

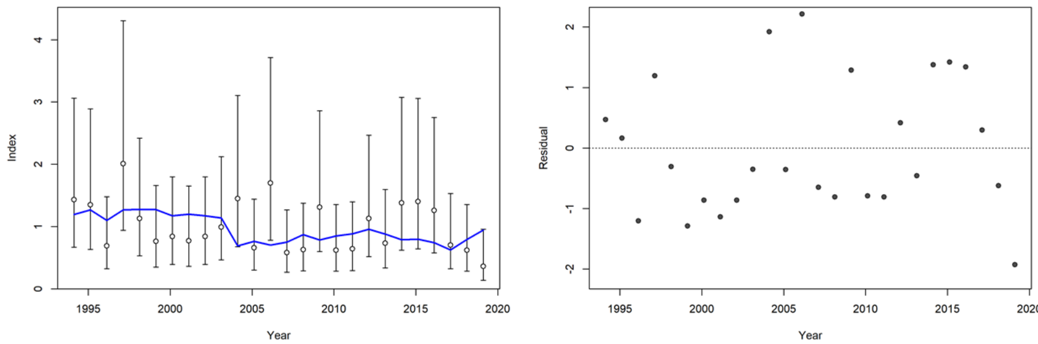



Figure 11. Fit to the S2 Japanese Late LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

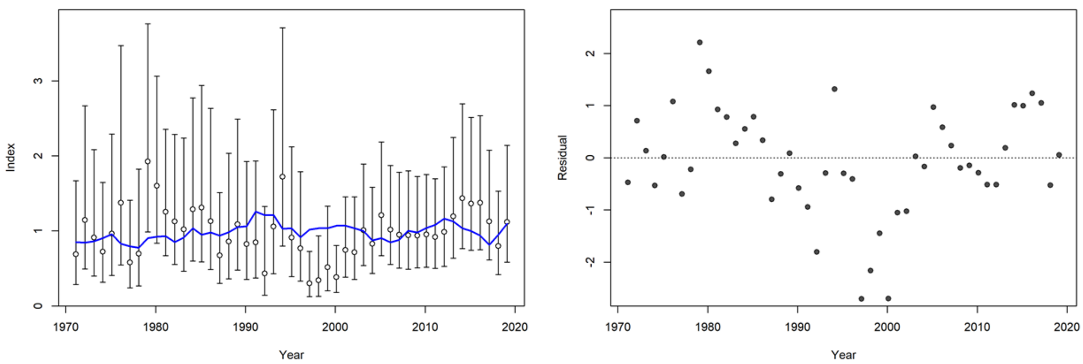



Figure 12. Fit to the S4 Taiwanese LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit. Note that catchability changes in 1979 and 1999.

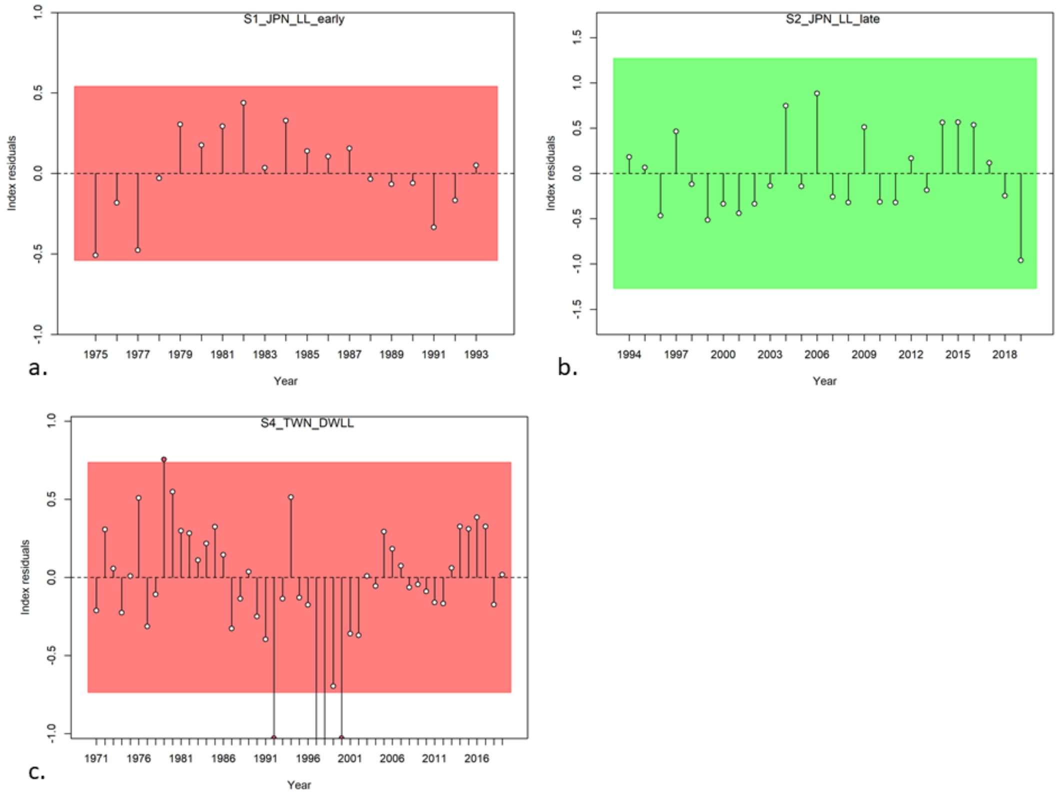



Figure 13. Results from a runs test for each CPUE index. Red indicates the index failed the test (residuals are not random), green indicates the index passed the test; subplot a.) S1 JPN Early LL, subplot b.) S2 JPN Late LL, subplot c.) S4 TWN LL.

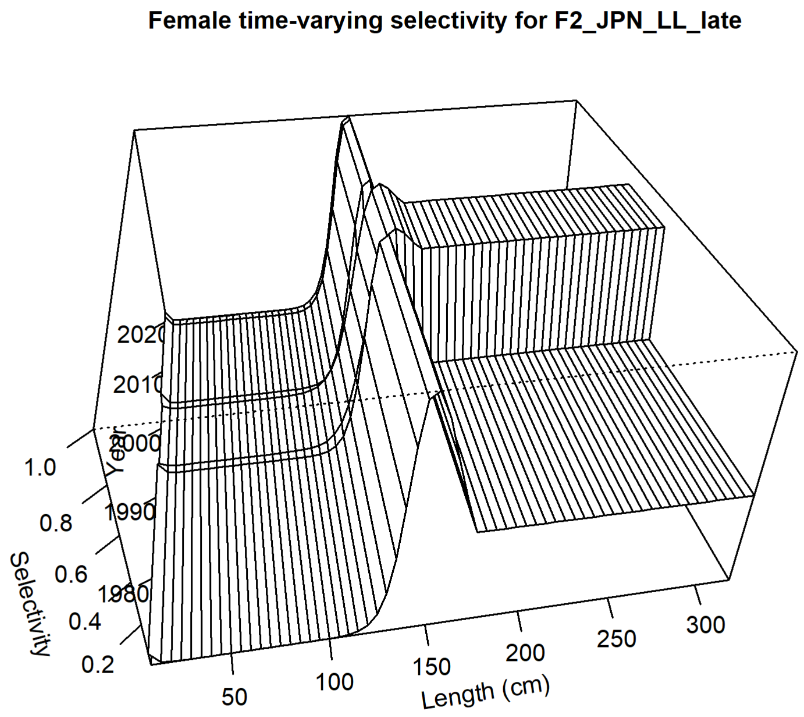



Figure 14. Time-varying selectivity estimated for F2 Japan LL Late.

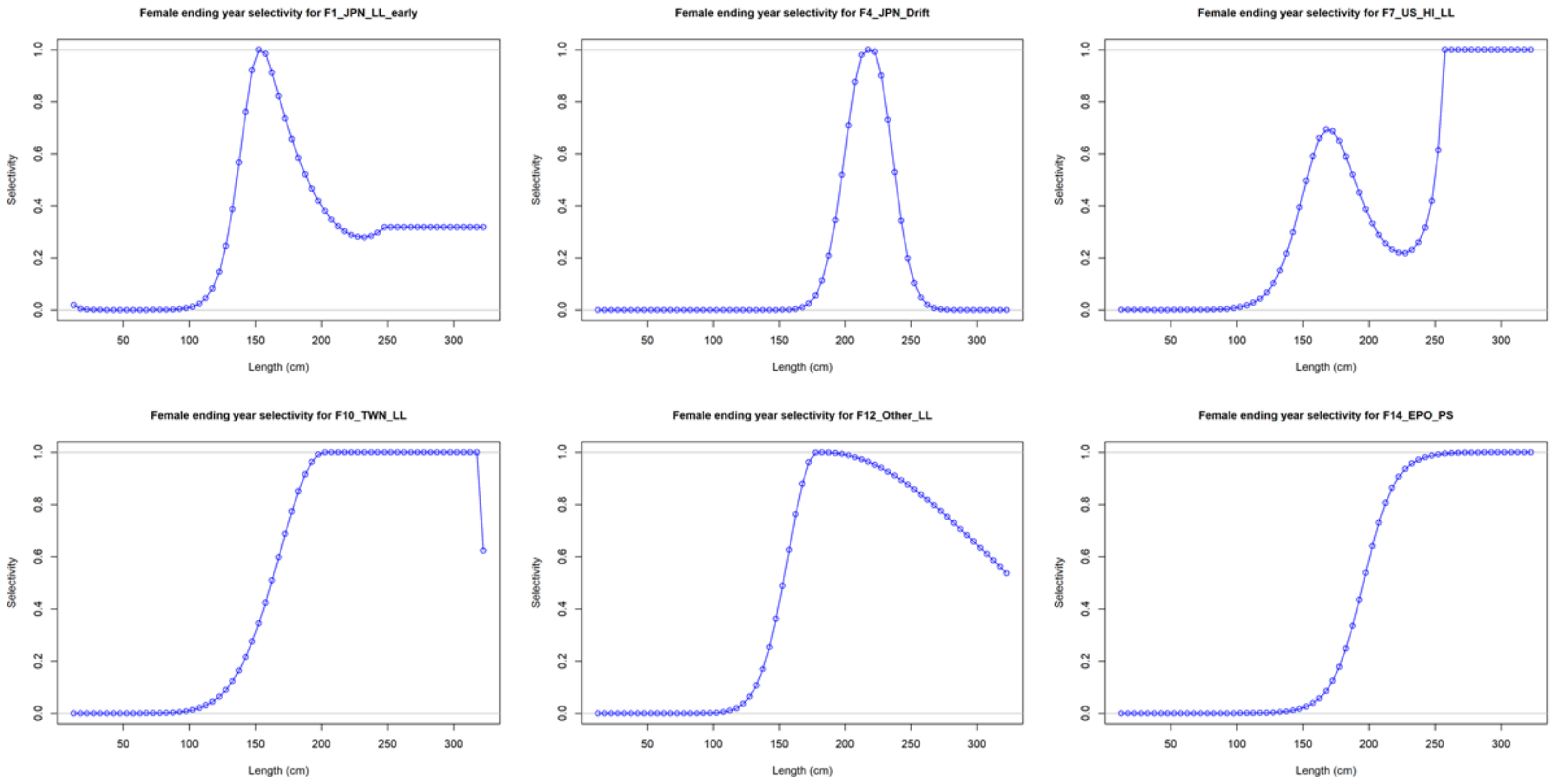



Figure 15. Selectivity estimates for each of the 6 fleets without time-varying parameters. Clockwise from the top left: F1 Japan LL Early, F4 Japan Driftnet, F7 US HI LL, F14, EPO Pure Seine, F12 Other LL, F10 Taiwanese LL.

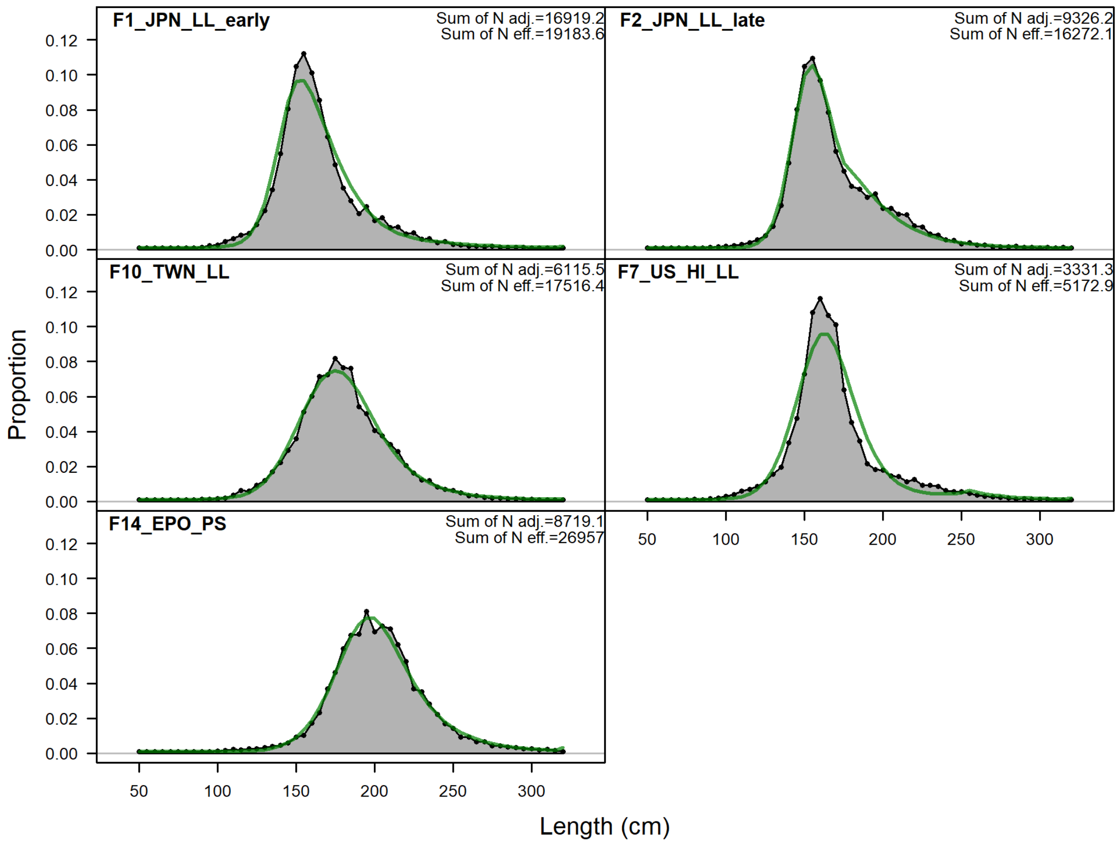



Figure 16. Overall fits to the length composition data for Japan Early and Late LL, Taiwanese LL, Hawaii LL, and EPO Purse Seine. Grey shading indicates input data, the green line indicates expected length distribution based upon estimated selectivity.

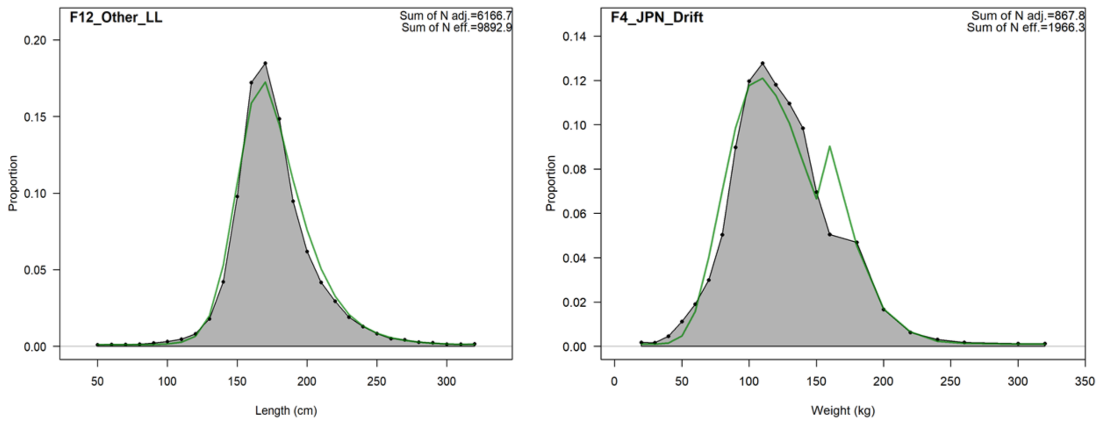



Figure 17. Overall fits to the length composition data for Other LL and Japan Driftnet. Grey shading indicates input data, the green line indicates expected length distribution based upon estimated selectivity.

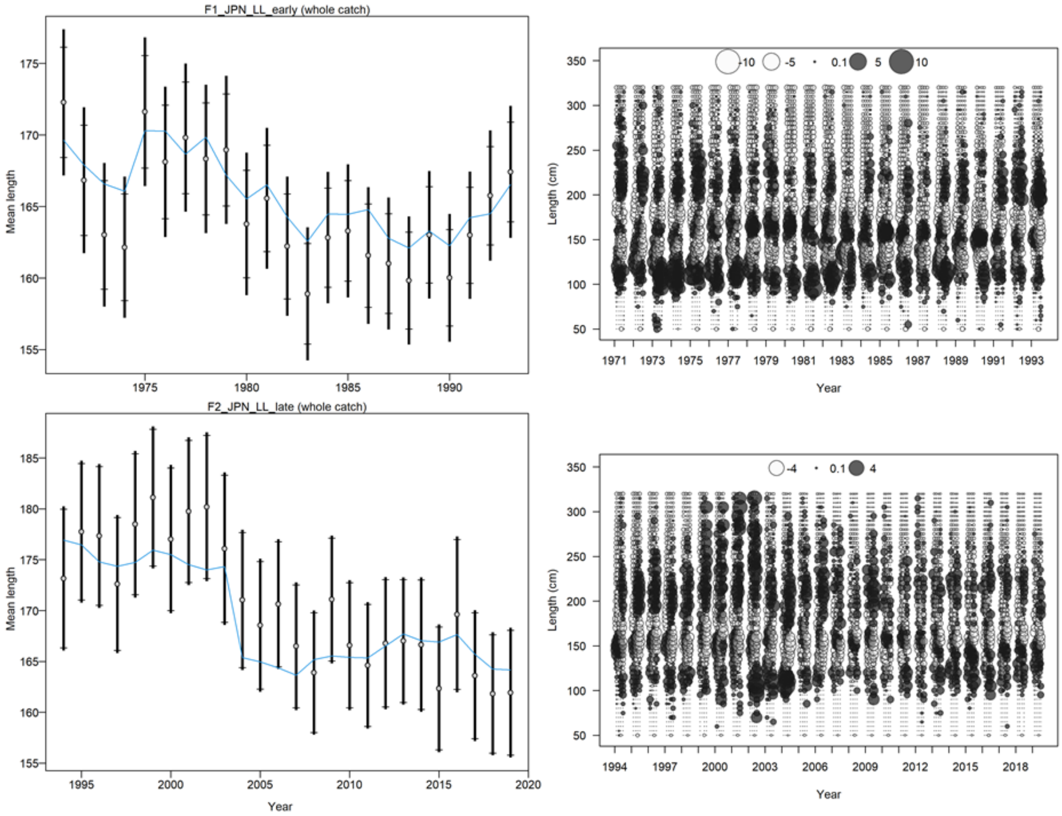



Figure 18. Fits to the annual mean length (left panels) and quarterly residuals (right panels) for Japan LL early (top) and late (bottom) length composition data. The blue line indicates the estimated mean length, open dots indicate input mean length with black bars indicating the distribution of the length data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

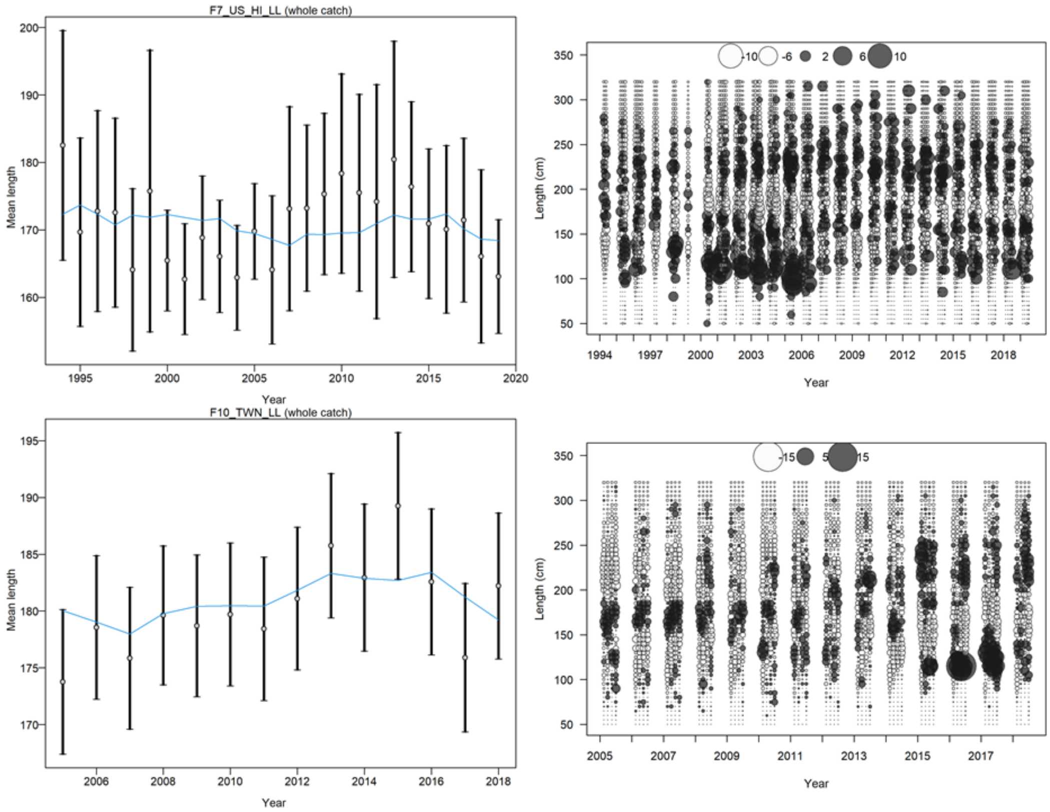



Figure 19. Fits to the annual mean length (left panels) and quarterly residuals (right panels) for US Hawaii LL (top) and Taiwanese LL (bottom) length composition data. The blue line indicates the estimated mean length, open dots indicate input mean length with black bars indicating the distribution of the length data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

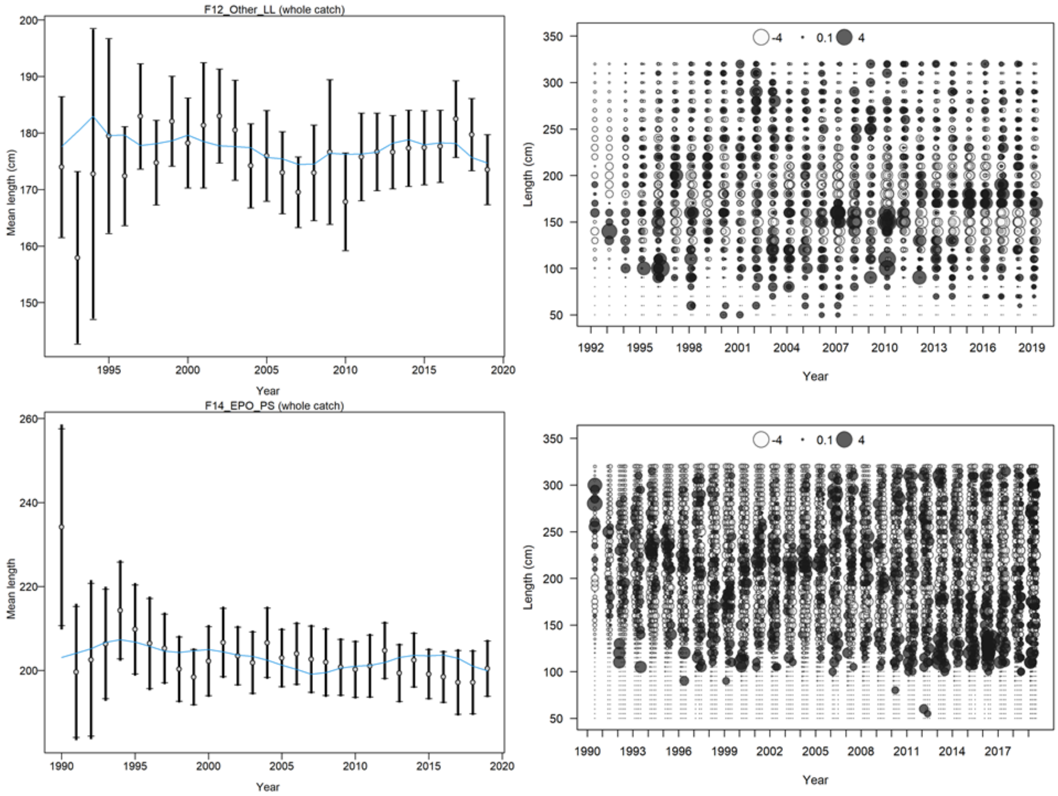



Figure 20. Fits to the annual mean length (left panels) and quarterly residuals (right panels) for Other LL (top) and EPO purse seine (bottom) length composition data. The blue line indicates the estimated mean length, open dots indicate input mean length with black bars indicating the distribution of the length data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

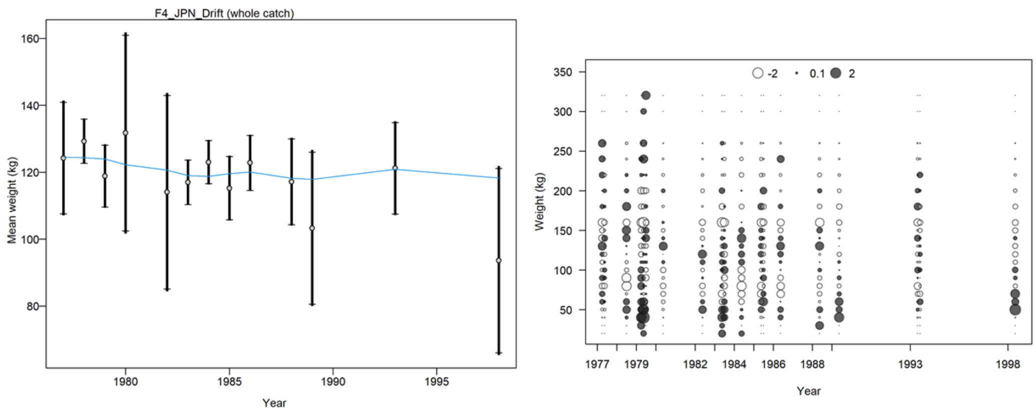



Figure 21. Fits to the annual mean weight (left panels) and quarterly residuals (right panels) for Japan

driftnet weight composition data. The blue line indicates estimated mean weight, open dots indicate input mean weight with black bars indicating the distribution of the weight data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

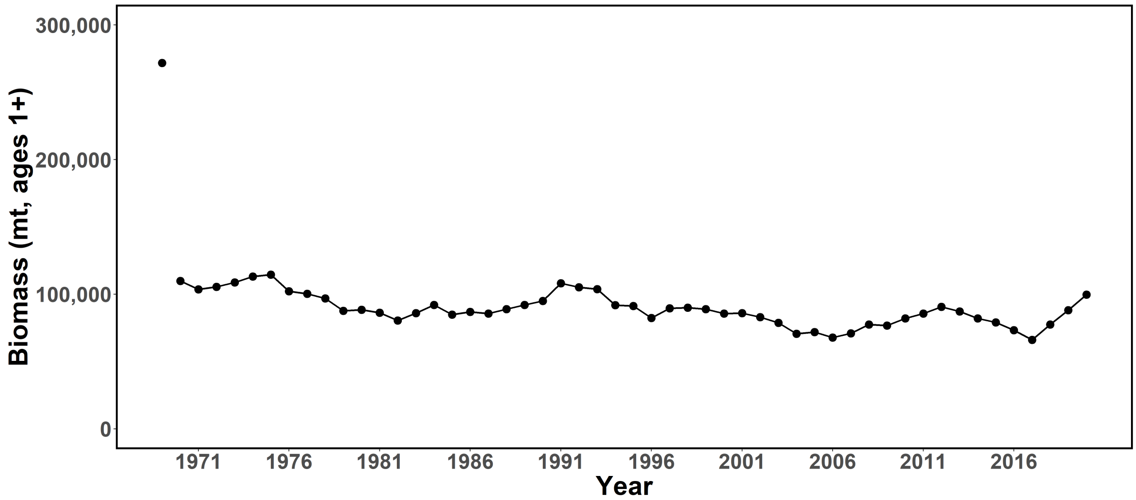



Figure 22. Estimated biomass (mt) of Pacific blue marlin ages 1+ from the old growth base-case model. Virgin population size is indicated by the first data point.

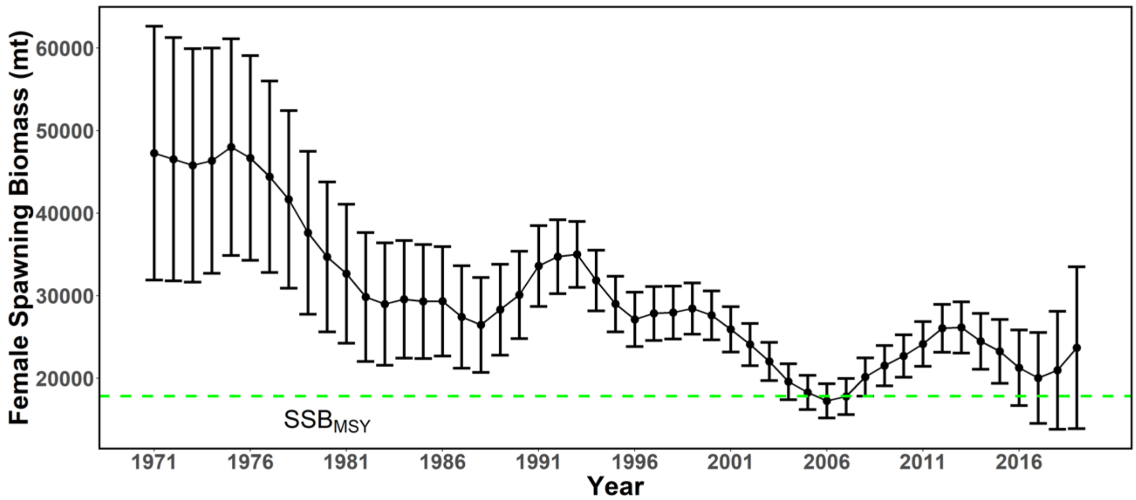



Figure 23. Estimated Female Spawning Stock Biomass (SSB) from the old growth assessment model with 95% confidence intervals. SSBMSY is indicated by the dashed green line.

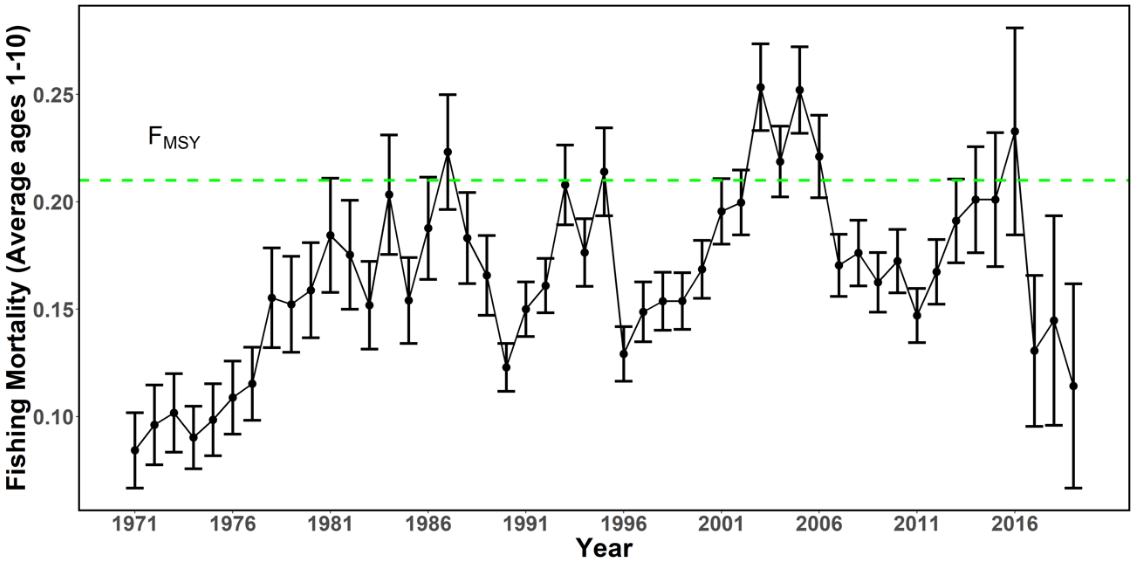



Figure 24. Estimated annual fishing mortality (ages 1-10) from the old growth assessment model with 95% confidence intervals. $F_{MSY}$ is indicated by the dashed green line.

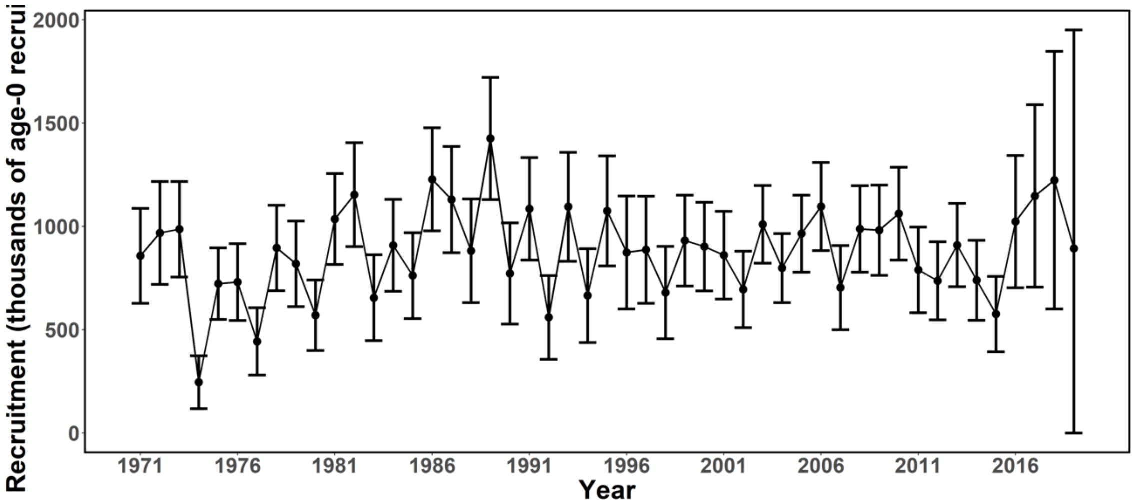



Figure 25. Estimated annual recruitment (thousands of age-0 fish) with 95% confidence intervals from the old growth base-case model.

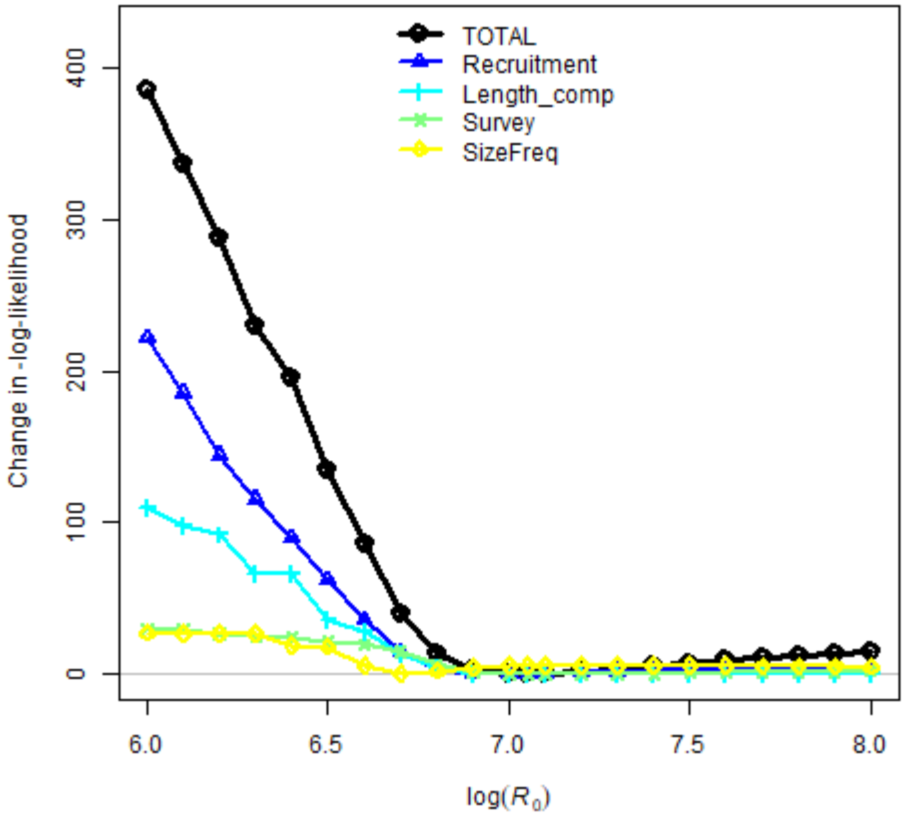



Figure 26. Likelihood profile over R0 for the old growth base-case model: total likelihood (black circles), recruitment (blue triangles), length composition data (light blue vertical bars), survey/CPUE indices (green x's), and generalized size frequency index (yellow diamonds).

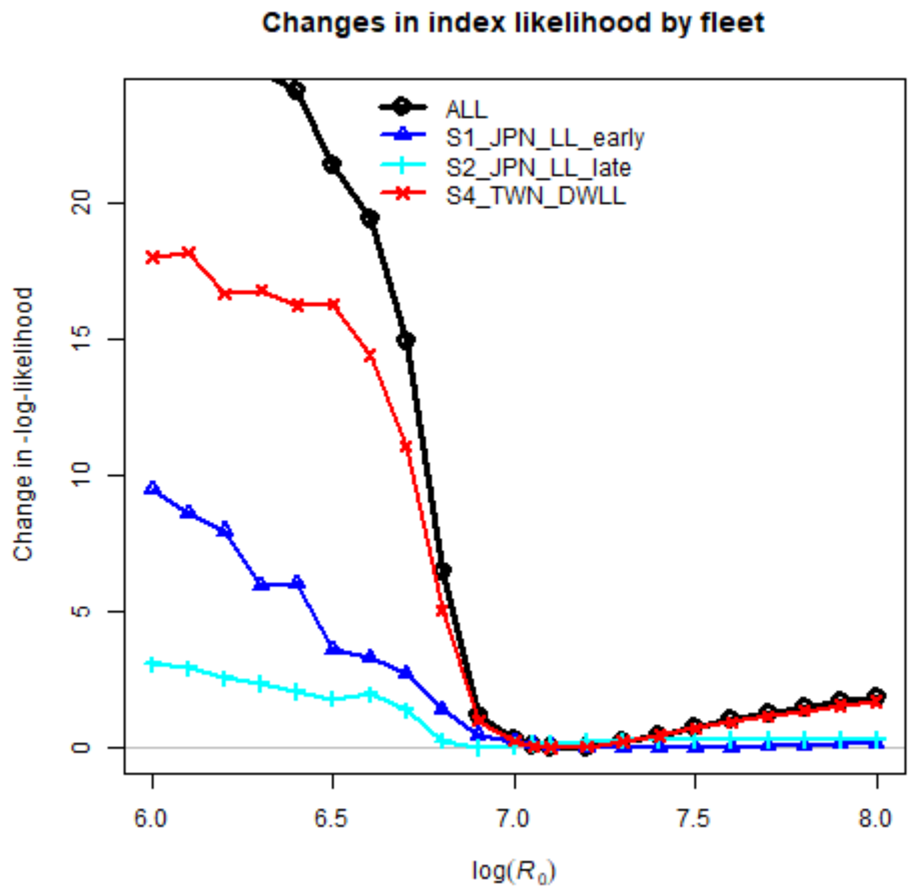



Figure 27. Likelihood profile over R0 by CPUE index for the old growth base-case model.

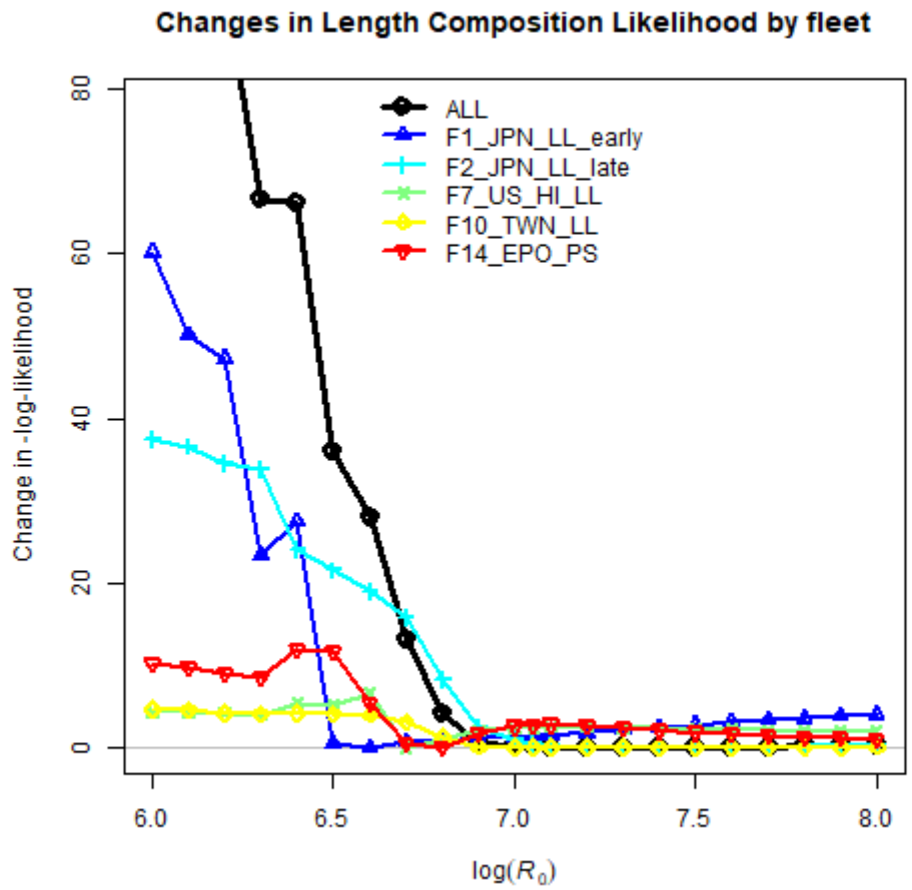



Figure 28. Likelihood profile over R0 for each length composition time series for the old growth base-case model.

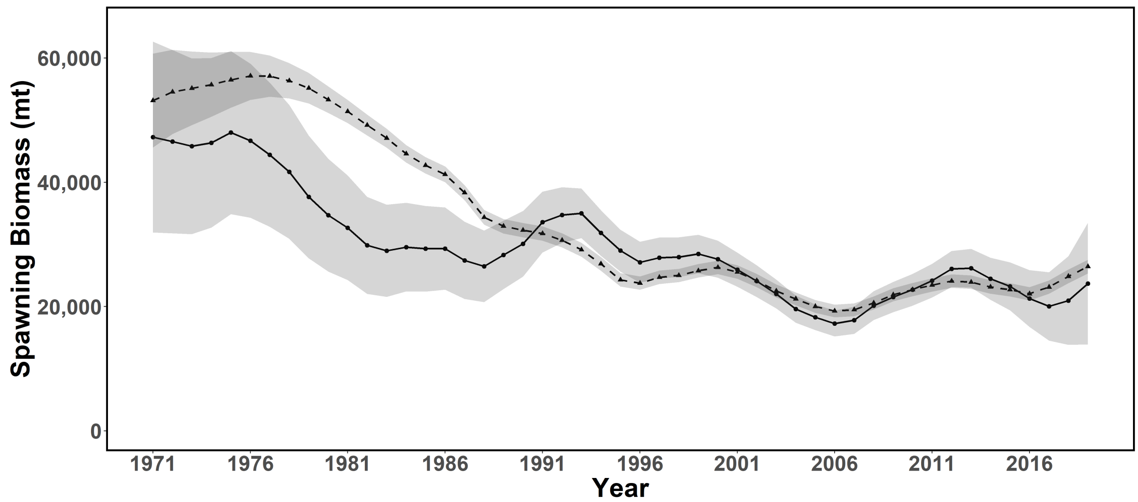



Figure 29. Female spawning stock biomass trend for the ASPM model run (dashed line, triangles) and the old growth base-case model (solid line, circles). Grey shading indicates 95% confidence intervals for each model.

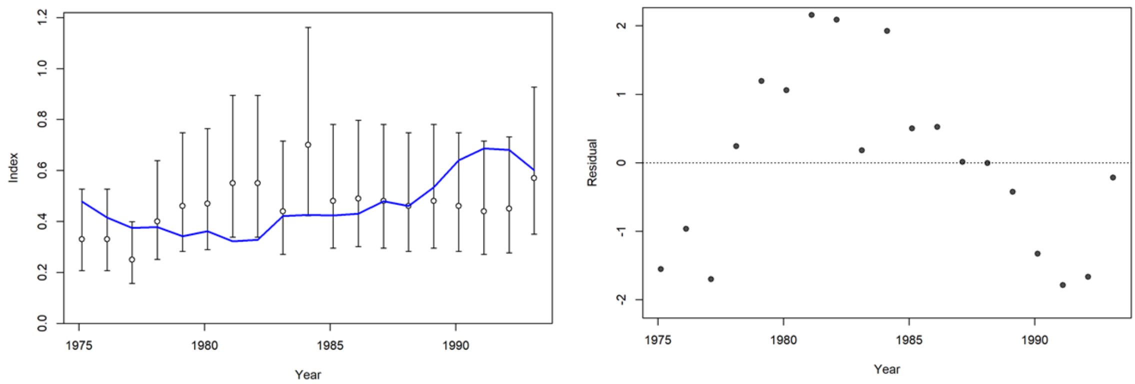



Figure 30. Fit to the S1 Japanese Early LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

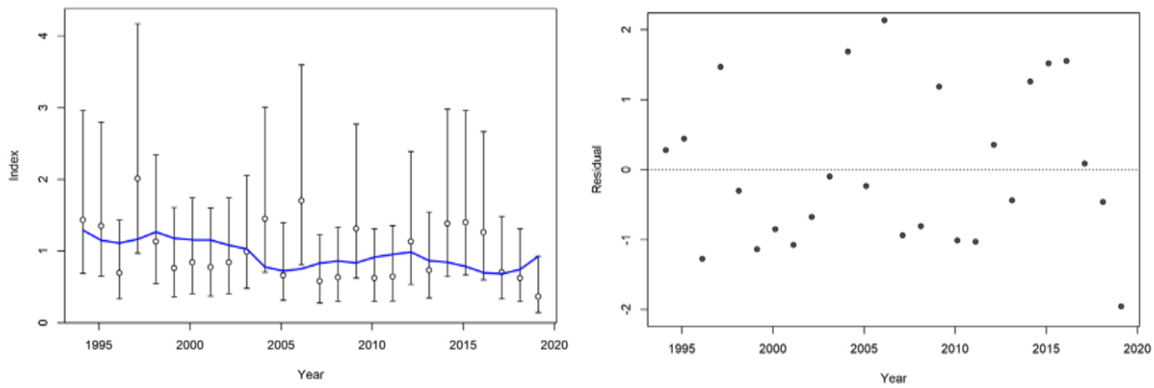



Figure 31. Fit to the S2 Japanese Late LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

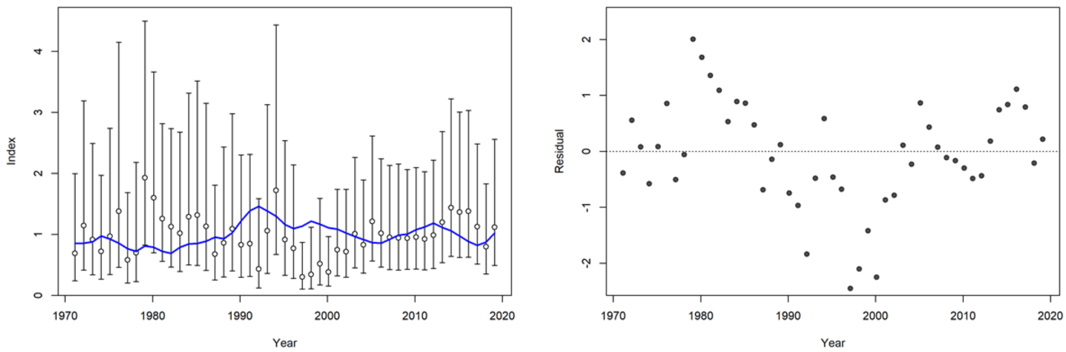



Figure 32. Fit to the S4 Taiwanese LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit. Note that catchability changes in 1979 and 1999.

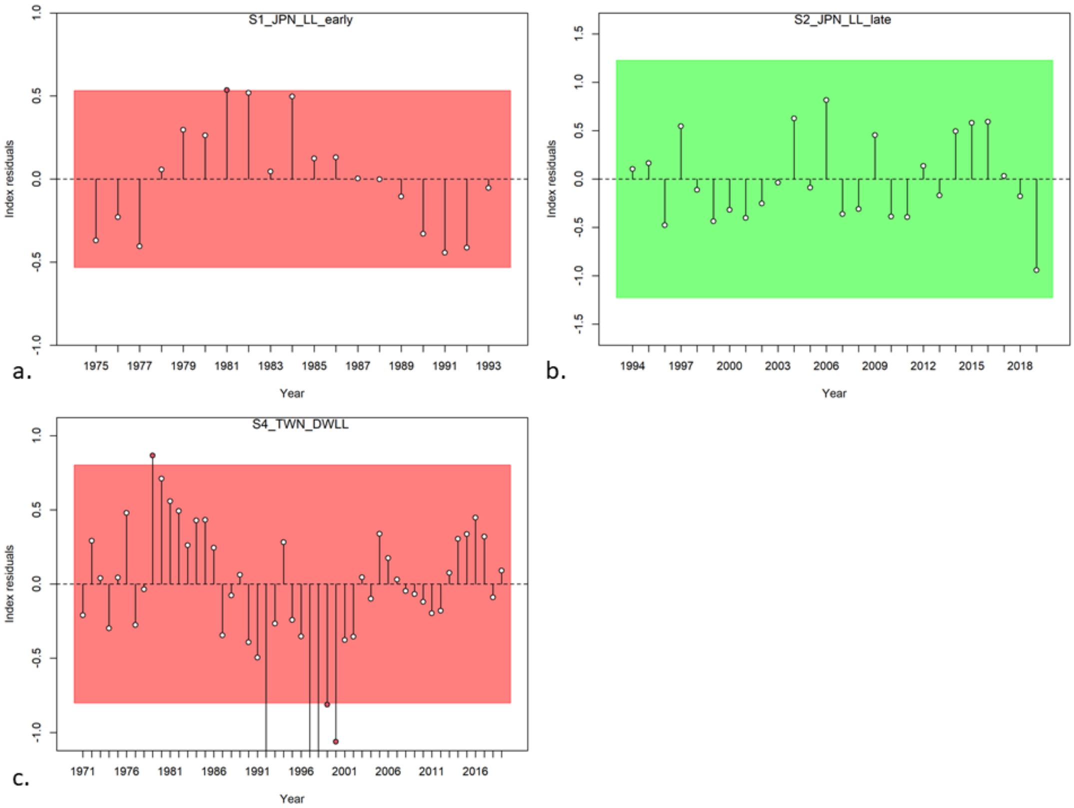



Figure 33. Results from a runs test for each CPUE index. Red indicates the index failed the test (residuals are not random), green indicates the index passed the test; subplot a.) S1 JPN Early LL, subplot b.) S2 JPN Late LL, subplot c.) S4 TWN LL.

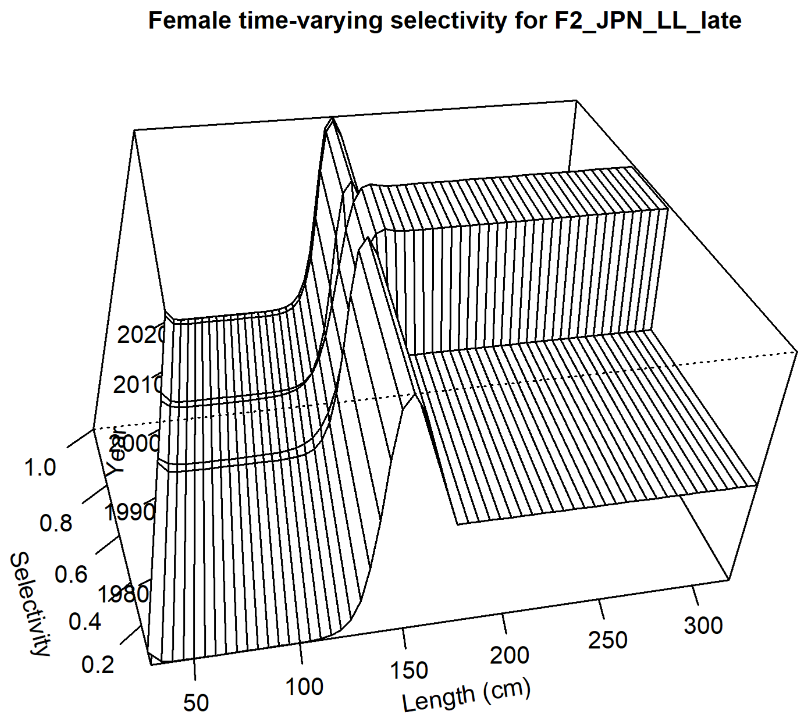



Figure 34. Time-varying selectivity estimated for F2 Japan LL Late.

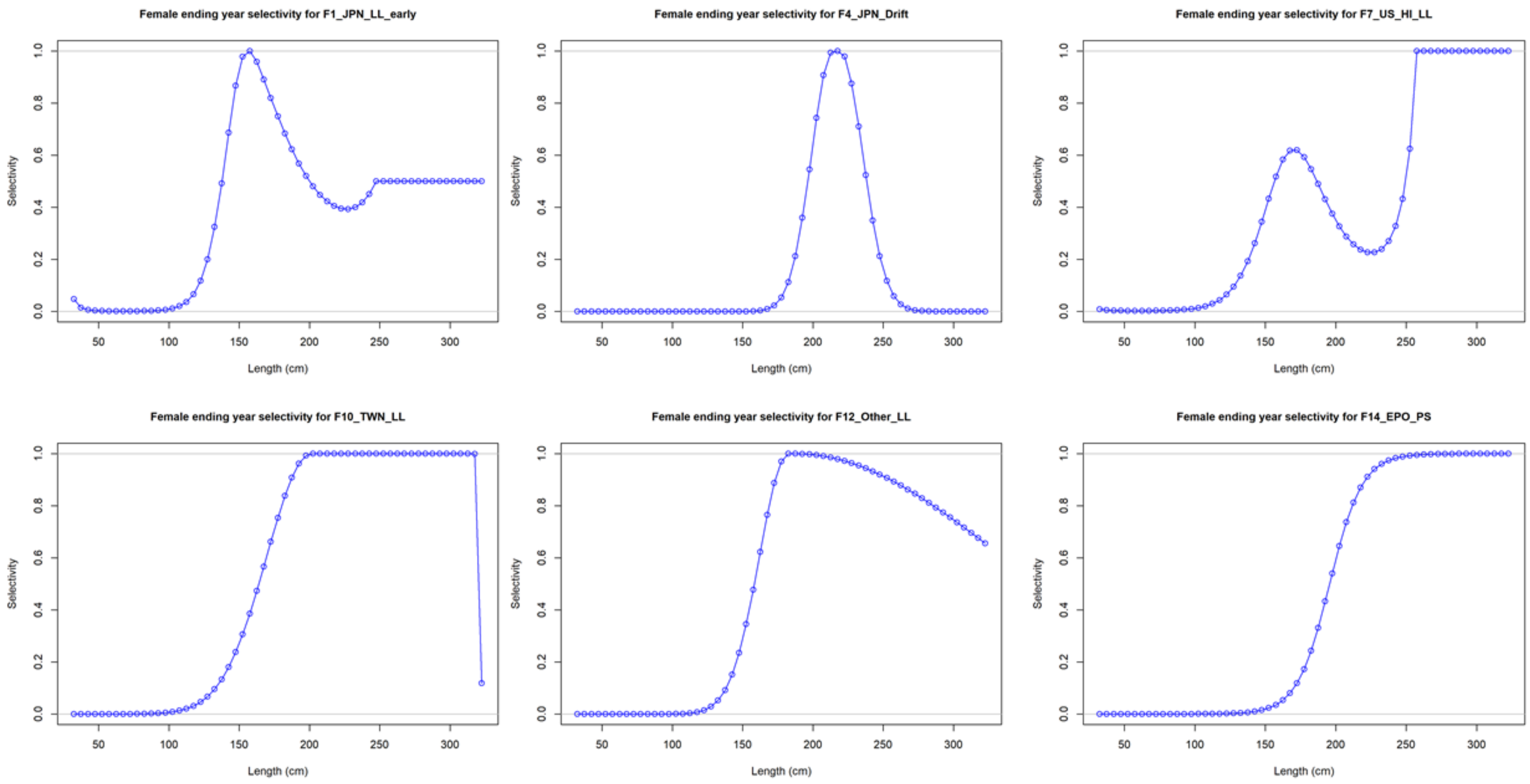



Figure 35. Selectivity estimates for each of the 6 fleets without time-varying parameters. Clockwise from the top left: F1 Japan LL Early, F4 Japan Driftnet, F7 US HI LL, F14, EPO Pure Seine, F12 Other LL, F10 Taiwanese LL.

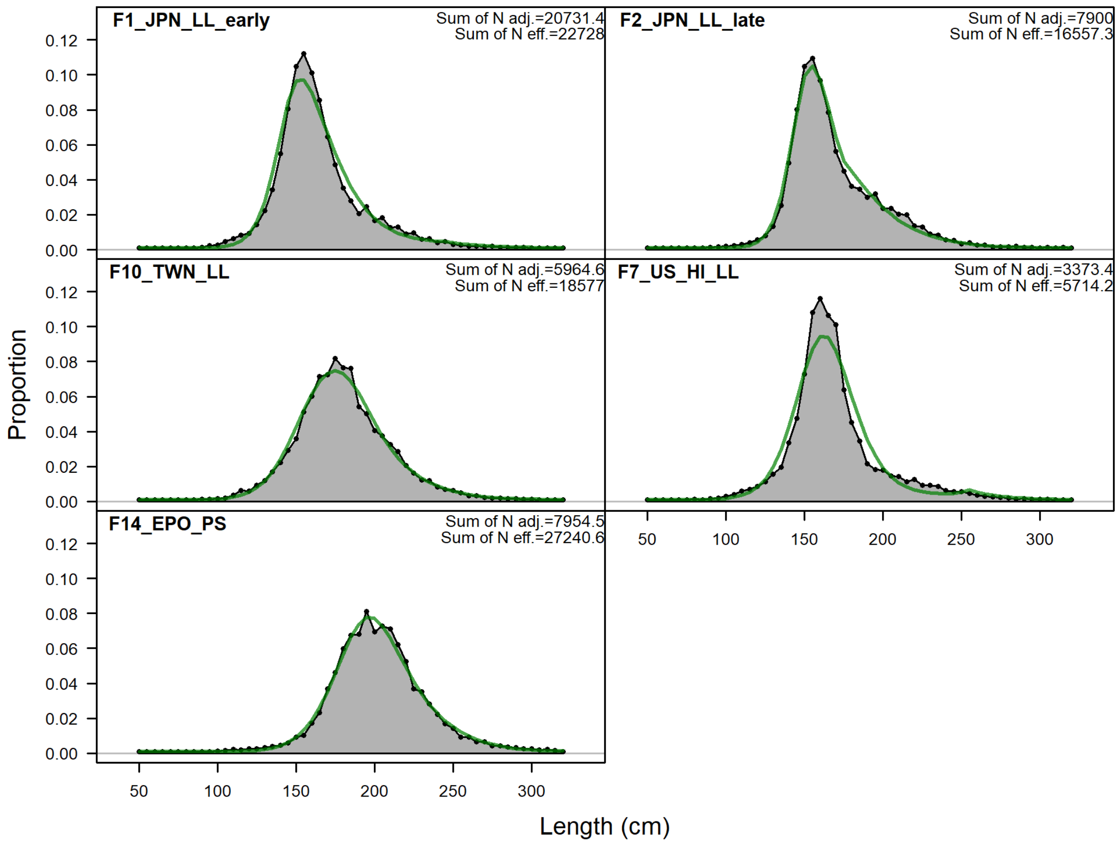



Figure 36. Overall fits to the length composition data for Japan Early and Late LL, Taiwanese LL, Hawaii LL, and EPO Purse Seine. Grey shading indicates input data, green line indicates expected length distribution based upon estimated selectivity.

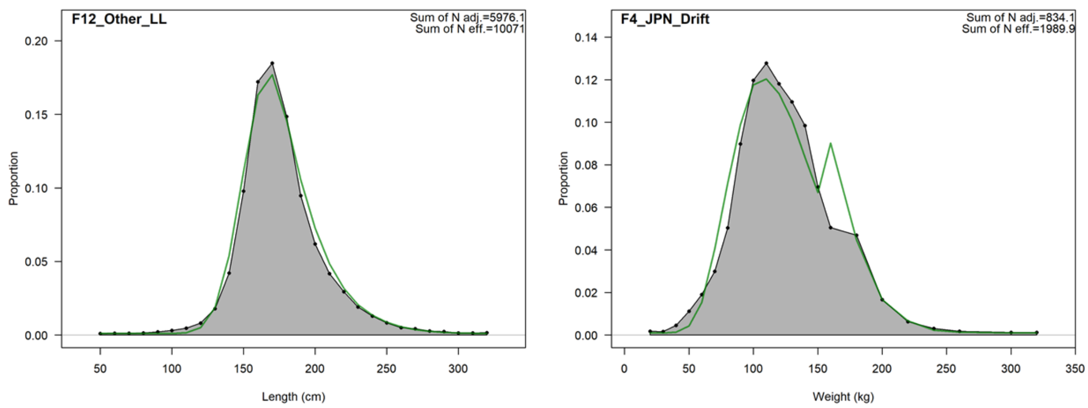



Figure 37. Overall fits to the length composition data for Other LL and Japan Driftnet. Grey shading indicates input data, green line indicates expected length distribution based upon estimated selectivity.

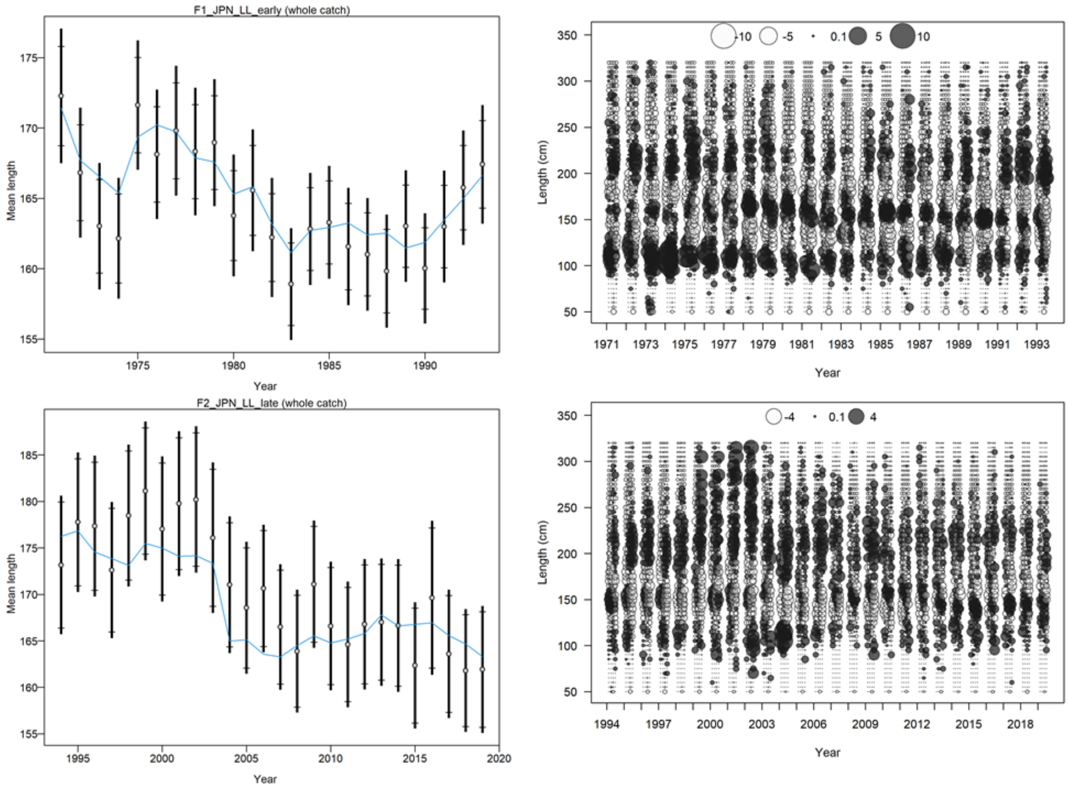



Figure 38. Fits to the annual mean length (left panels) and quarterly residuals (right panels) for Japan LL early (top) and late (bottom) length composition data. The blue line indicates the estimated mean length, open dots indicate input mean length with black bars indicating the distribution of the length data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

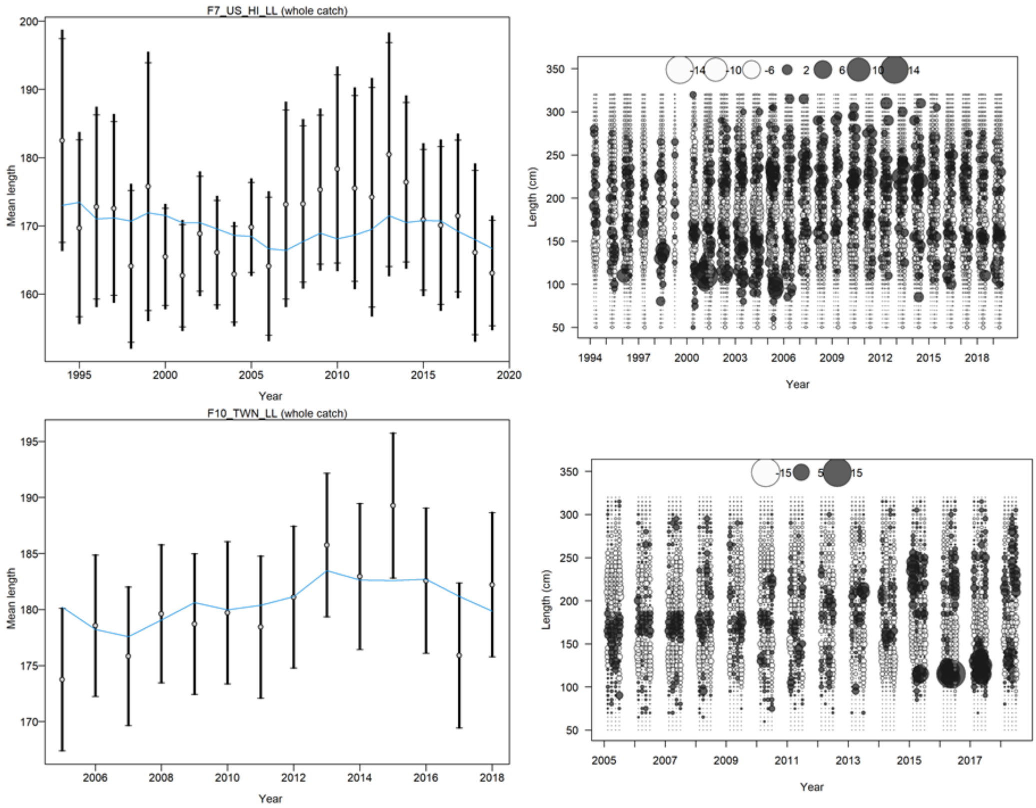



Figure 39. Fits to the annual mean length (left panels) and quarterly residuals (right panels) for US Hawaii LL (top) and Taiwanese LL (bottom) length composition data. The blue line indicates the estimated mean length, open dots indicate input mean length with black bars indicating the distribution of the length data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

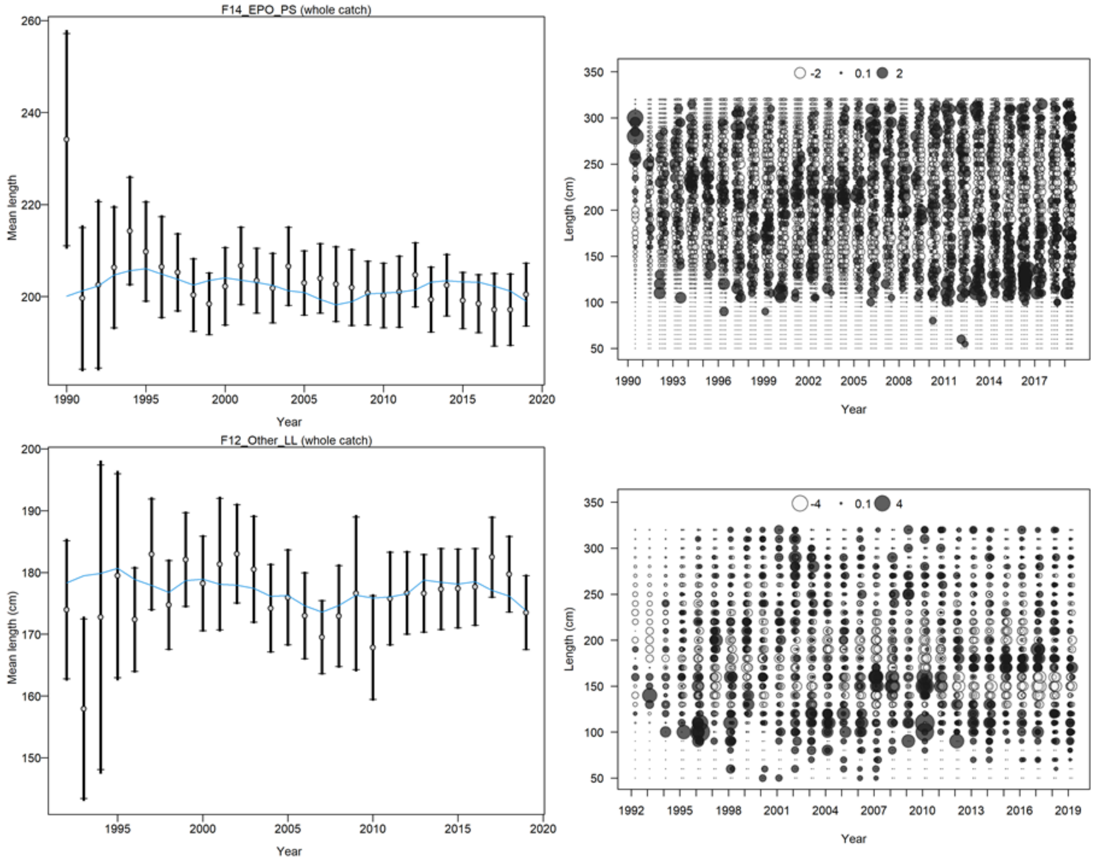



Figure 40. Fits to the annual mean length (left panels) and quarterly residuals (right panels) for Other LL (top) and EPO purse seine (bottom) length composition data. The blue line indicates the estimated mean length, open dots indicate input mean length with black bars indicating the distribution of the length data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

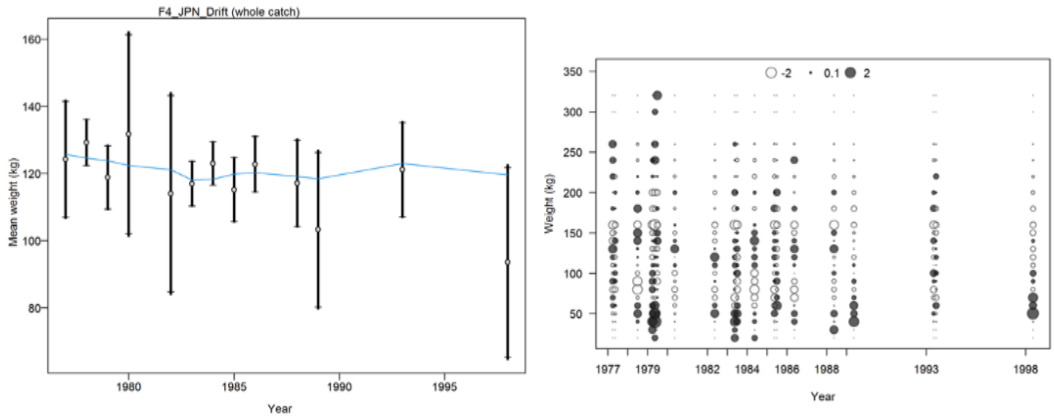



Figure 41. Fits to the annual mean weight (left panels) and quarterly residuals (right panels) for Japan driftnet weight composition data. The blue line indicates estimated mean weight, open dots indicate input mean weight with black bars indicating the distribution of the weight data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

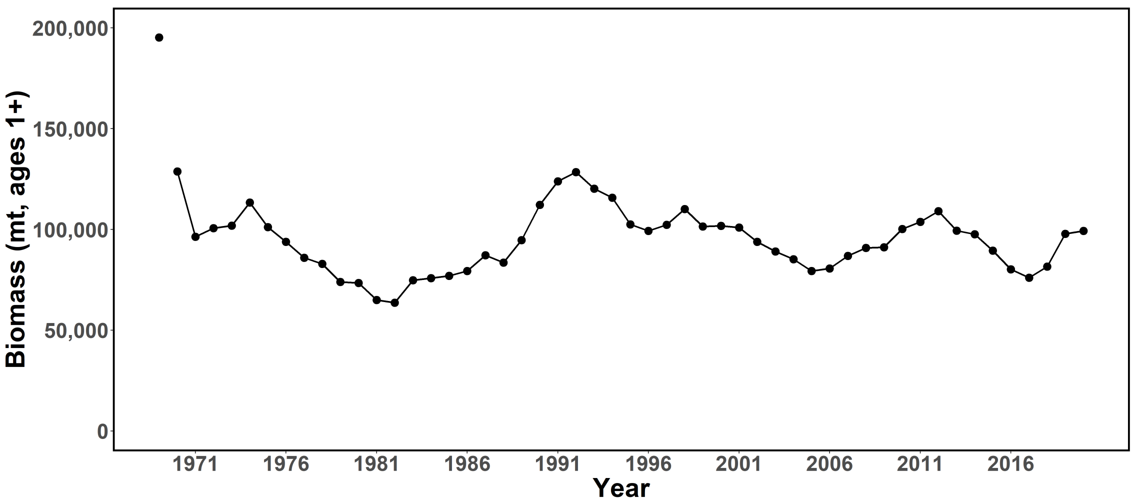



Figure 42. Estimated biomass (mt) of Pacific blue marlin ages 1+ from the new growth base-case model. Virgin population size is indicated by the first data point.

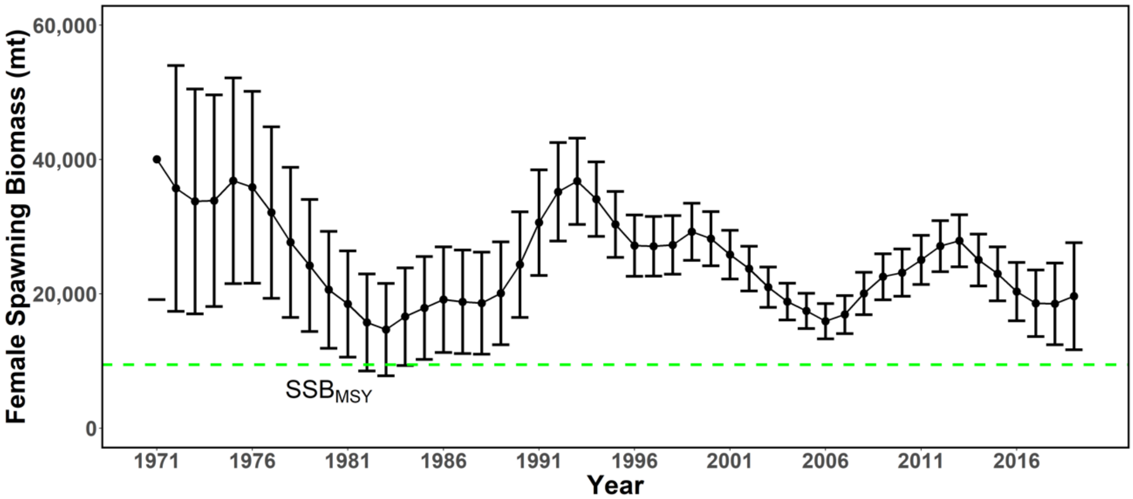



Figure 43. Estimated Female Spawning Stock Biomass (SSB) from the new growth assessment model with 95% confidence intervals. SSBMSY is indicated by the dashed green line.

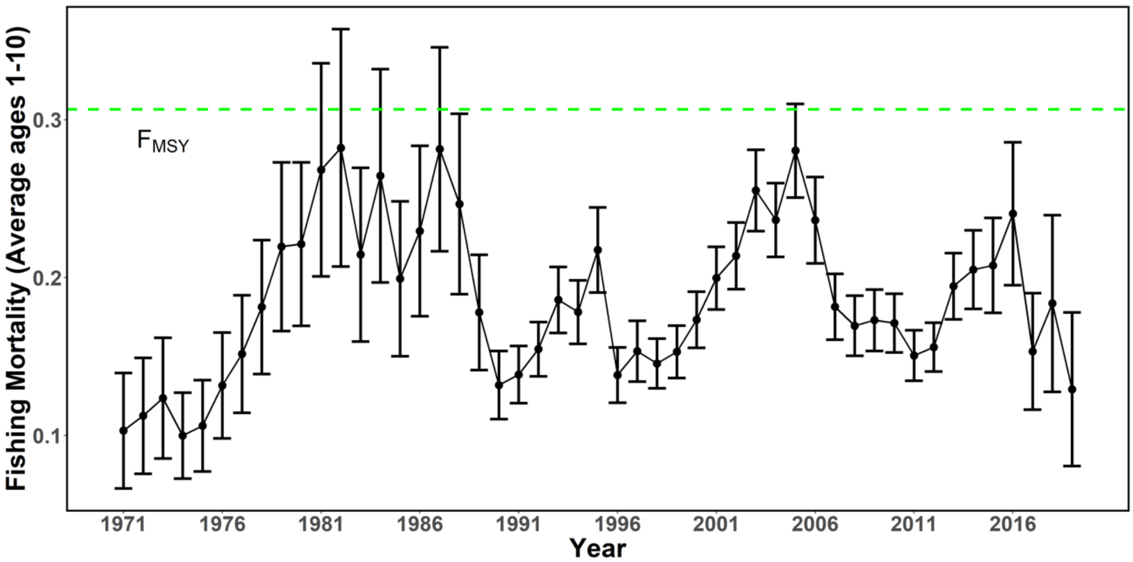



Figure 44. Estimated annual fishing mortality (ages 1-10) from the new growth assessment model with 95% confidence intervals. $F_{MSY}$ is indicated by the dashed green line.

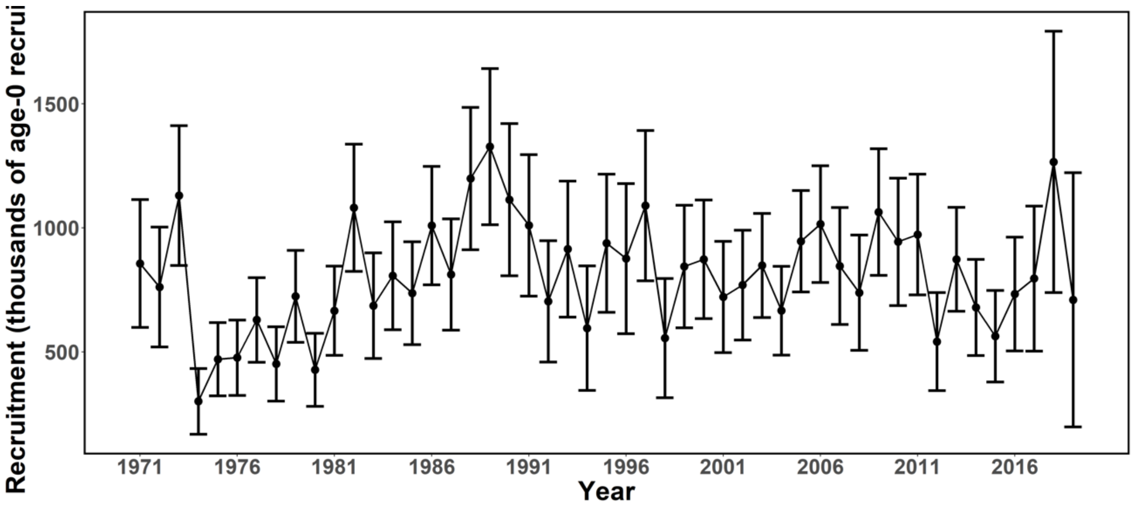



Figure 45. Estimated annual recruitment (thousands of age-0 fish) with 95% confidence intervals from the new growth base-case model.

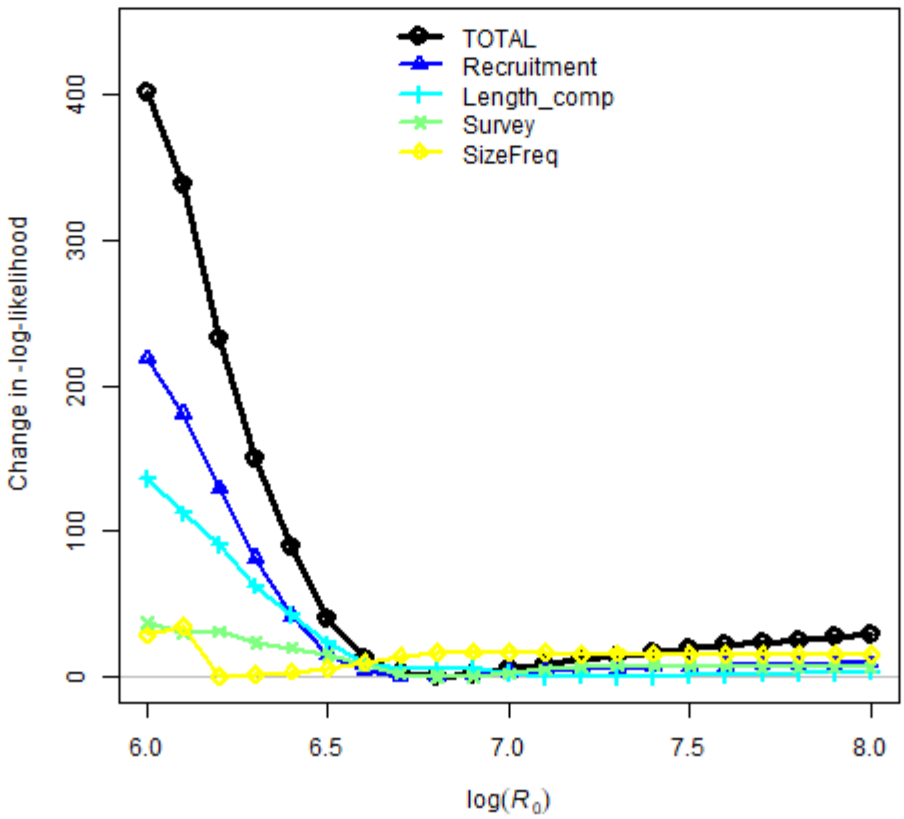



Figure 46. Likelihood profile over R0 for the new growth base-case model: total likelihood (black circles), length composition (blue triangles), survey/CPUE indices (light blue vertical bars), generalized size frequency (green x's), and recruitment index (yellow diamonds).

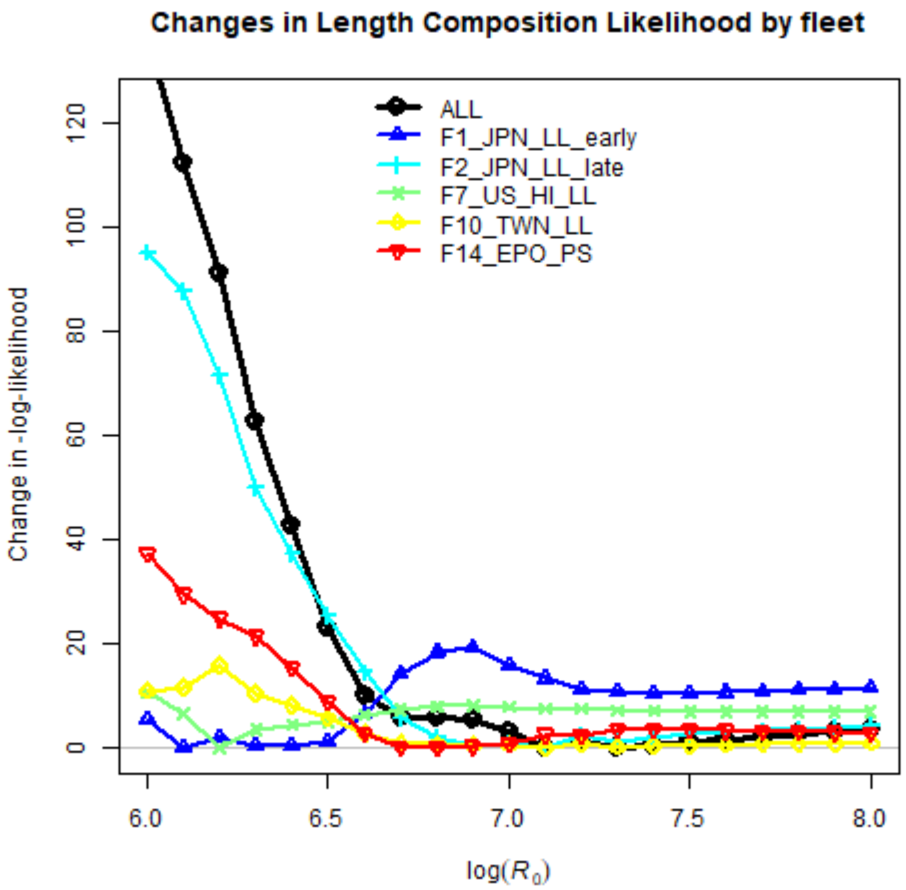



Figure 47. Likelihood profile over R0¬ for each length composition time series for the new growth base-case model.

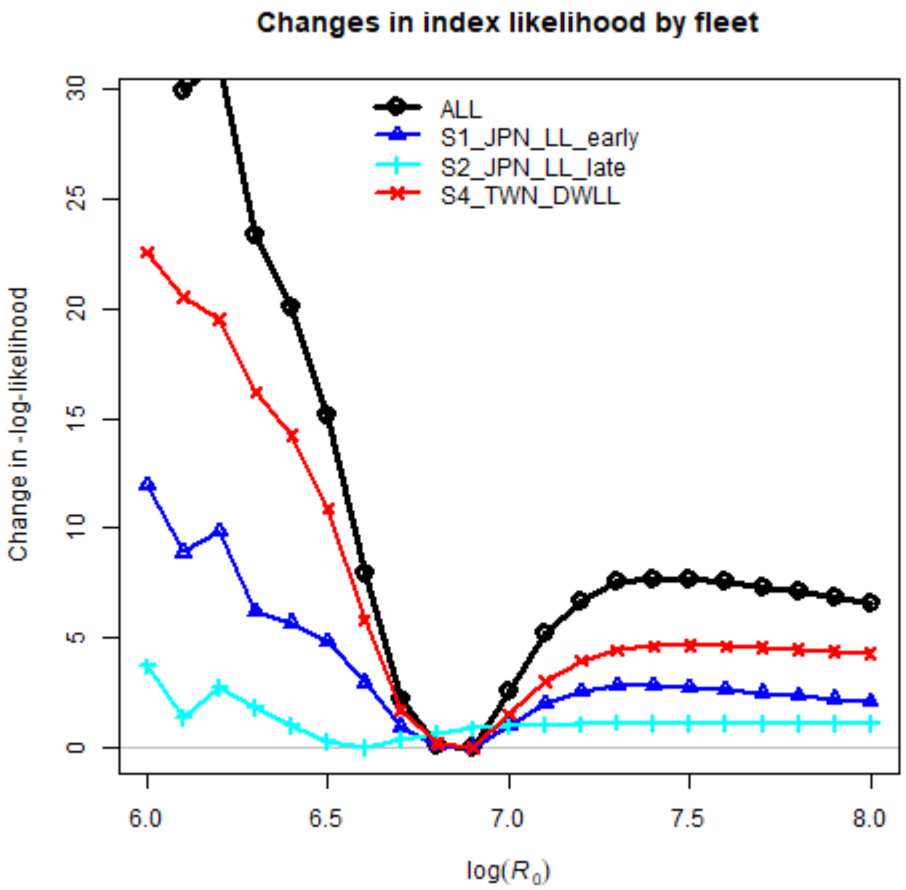



Figure 48. Likelihood profile over R0 by CPUE index for the new growth base-case model.

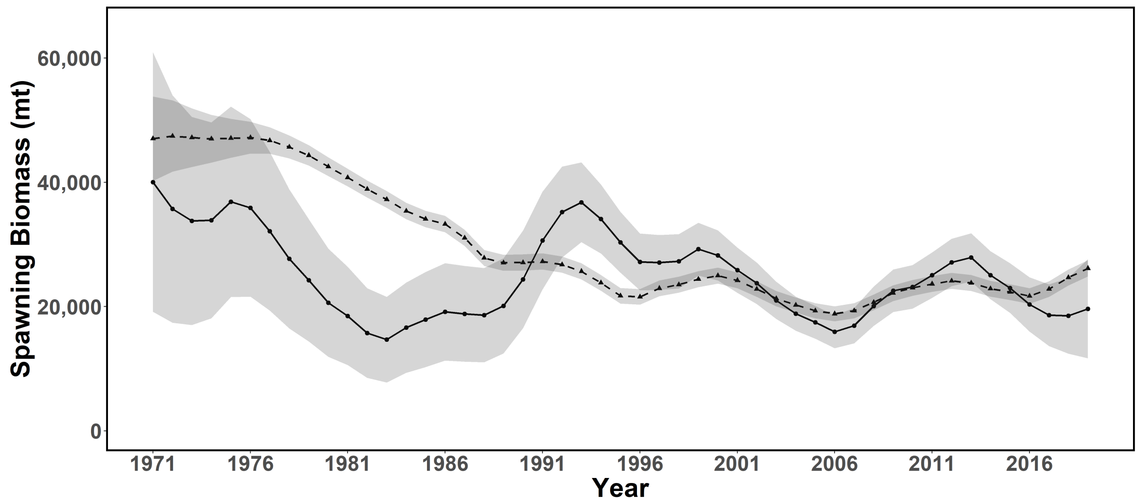



Figure 49. Female spawning stock biomass trend for the ASPM model run (dashed line, triangles) and the new growth base-case model (solid line, circles). Grey shading indicates 95% confidence intervals for each model.

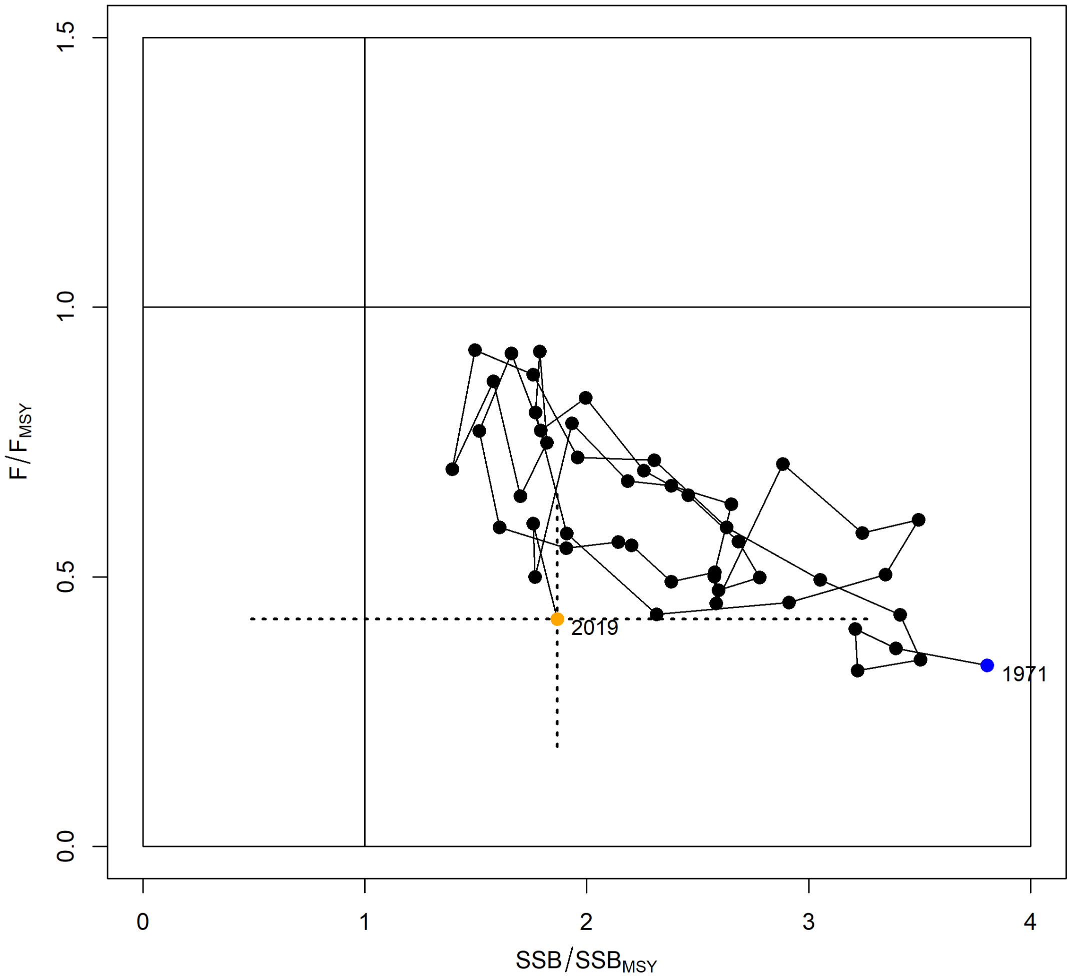



Figure 50. Kobe plot for the 2021 Pacific blue marlin new growth base-case model. The blue dot indicates the start of the assessment, 1971, the orange dot indicates the 2019 status, and the dashed lines indicate 95% confidence intervals.

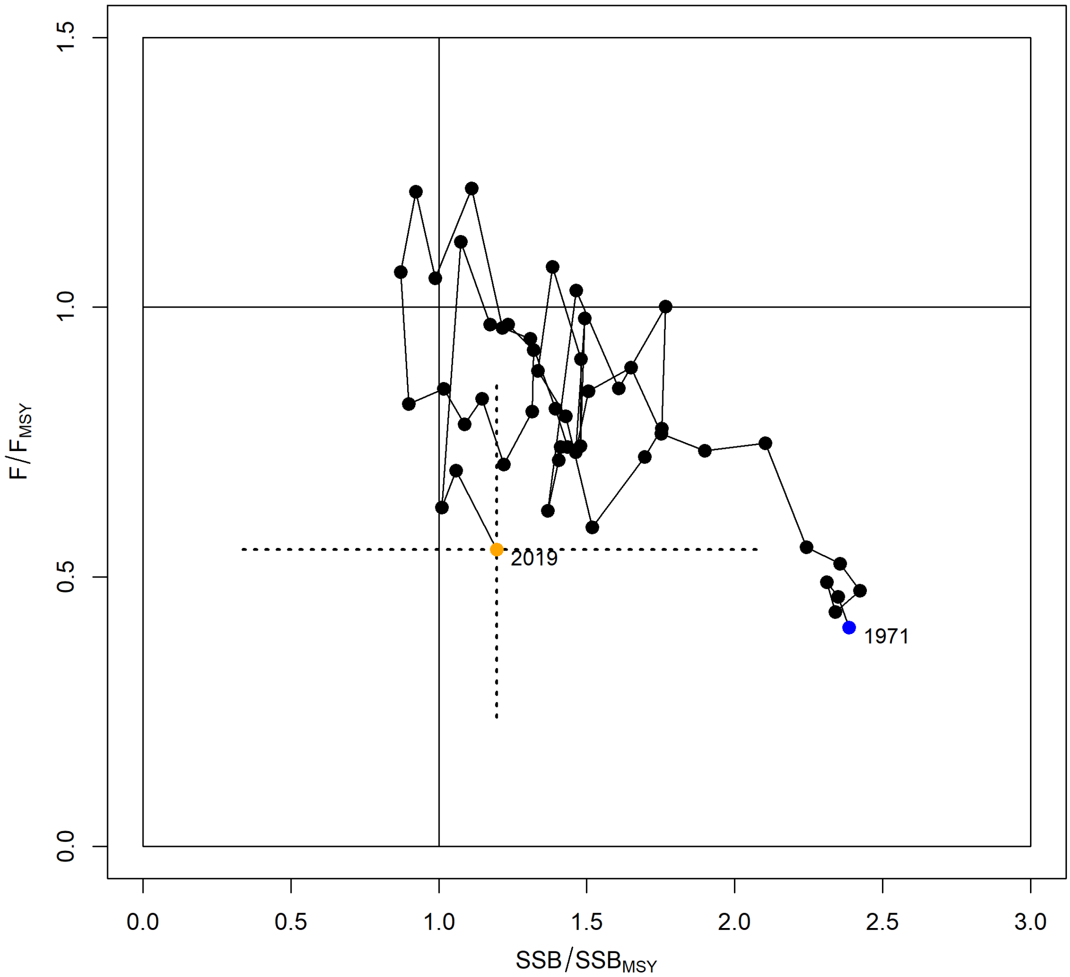



Figure 51. Kobe plot for the 2021 Pacific blue marlin old growth base-case model. The blue dot indicates the start of the assessment, 1971, the orange dot indicates the 2019 status, and the dashed lines indicate 95% confidence intervals.