Progress towards a Base-case Model in Stock Synthesis 3.30 for the 2026 Pacific Blue Marlin (*Makaira nigricans*) Stock Assessment

Erin C. Bohaboy*, Michelle Sculley*

*NOAA Fisheries

Pacific Islands Fisheries Science Center

Email: erin.bohaboy@noaa.gov

Working paper submitted to the International Scientific Committee for Tuna and Tuna-like Species (ISC), Billfish Working Group Meeting, April 28 – May 4, 2026, Honolulu, H. Document not to be cited without authors' permission.

Abstract

This working paper describes two preliminary base-case models in Stock Synthesis 3.30 for Pacific blue marlin (*Makaira nigricans*) for consideration as the 2026 base-case model. Both models include years 1971–2024 with 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). The two preliminary base-case models have identical data inputs, with the exception of the catch per unit effort (CPUE) indices included: the "Japan" model excludes CPUE indices from Taiwan, whereas the "Taiwan" model excludes CPUE indices from Japan. Initial diagnostics reveal problematic performance of both models. Preliminary results from both models suggest current (2024) spawning stock biomass is below SSB_{MSY}, whereas only the Japan model suggests the Pacific blue marlin stock is being recently (2022–2024 average) fished above F_{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 agreed to by the billfish working group (BILLWG) at the data preparatory meeting in January 2026, where a new growth curve was presented (Chang *et al.,* 2025). The WG agreed to use ~~new growth curve and~~the updated natural mortality based upon the new growth curve (Brodziak 2026). Due to conflict between some of the CPUE indices, two models were run using only the CPUE indices from Japan and US, the "Japan" model, and only CPUE indices from Taiwan and the US, the "Taiwan" model. Aside from the inclusion of CPUE indices, all other data included were the same. 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*; BUM) is assessed as a Pacific-wide stock. Generally, BUM are found in the warm tropical waters of the Pacific Ocean and are known to make long distance movements for feeding and spawning over relatively short periods of time, potentially moving east and west across the Pacific Ocean and crossing the equator (Su et al., 2011; Carlisle et al., 2016). The Billfish Working Group (WG) drafted a conceptual model which details the seasonal and ontogenetic movements of BUM (ISC BILLWG, 2026). Size composition data generally indicate larger individuals are encountered by fisheries in the eastern and more poleward areas of the Pacific Ocean, whereas smaller individuals are more encountered by fishing fleets in the western and equatorial areas. So, although assessed as a single stock, the WG agreed to account for this spatial structure whenever possible by dividing index and size data among east and west areas (boundaried by 170°W longitude) and within or widespread beyond the main equatorial fishing grounds (20°S–20°N latitude).

The WG agreed to run the base model from 1971 to 2024 when catch, CPUE, and size-frequency data are all available. The stock was assumed to be in a fished condition at the start of the model. There are limited data to estimate catches going back to 1950 when the stock is generally believed to be at essentially unfished levels. Given the rapid growth of BUM and the diverse seasonal nature of Members' fisheries, the model was run using a quarterly timestep, see Base-case model description for more details on the model specifications.

Definition of fisheries

Twenty different fleets are available for inclusion in the base-case model: 12 catch time series, 8 CPUE indices, 8 fleets with length composition data, and 5 fleets with weight composition data. The fleet names, numbers, descriptions, and data available are detailed in Table 1 and shown graphically in Figure 1. The acronyms in the fleet names and used throughout this working paper are: ISC Members: JPN, Japan; TWN, Taiwan; US, United States; IATTC, Inter-American Tropical Tuna Commission; WCPFC, Western and Central Pacific Fisheries Commission; VAR, any combination of flags; Fishing gears: LL, longline; LLdeep, deepset longline; DRIFT, driftnet; BAIT, bait fishery; Other, other fishing gear (e.g. troll, handline, net, harpoon, and others); PS, purse seine gear; Fleet characteristics: LTLL, large tonnage longline (referring to the size of vessels engaged in fishing); STLL, small tonnage longline (referring to the size of vessels engaged in fishing).

Catch

Japan, Taiwan, and the US (Hawaii and American Samoa) contributed catch time series, Also, catch from countries reporting to the WCPFC and IATTC were obtained from each RFMO (Figure 2). Catch was input to the model in biomass (mt) for all fleets except F11_US_LL which was input in thousands of individuals. The CV for catch was fixed at 0.05 for all fleets. Catch for fleets with only annual data were divided equally into each quarter.

Four Japanese catch time series were used in the assessment. F1_JPN_EarlyLL included coastal, offshore, and distant water longline, and other gears for years 1971–1993. F1_JPN_EarlyLL was assumed to be active prior to the start of the model in 1971, so initial quarterly catches for the start of the assessment model were assumed equal to the average 1971–1972 quarterly catches for that fleet. F2_JPN_LateLL_East and F3_JPN_LateLL_West included all longline gears east and west, respectively, of 170°W longitude for 1994–2024. The catch series for the Japanese "Early" period (F1) and "Late" period (F2 and F3) were divided in 1993 because of significant changes in the logbook reporting system, which required the CPUE data to be standardized in different time periods and the corresponding harvest fleets to be divided similarly. F4_JPN_DRIFT_BAIT was catch from highseas large-mesh driftnet, coastal driftnet, and bait fishing for 1972–2024. Catch for this fleet was available annually, so was divided equally among the 4 quarters each year.

Three Taiwanese catch time series, available annually 1971–2024, were divided equally among quarters within each year for each fleet and used in the assessment. F5_TWN_LTLL and F6_TWN_STLL included catch from the highseas longline fishery by large tonnage and small tonnage longline vessels, respectively. F6_TWN_STLL was assumed to be active prior to the start of the model in 1971, so initial quarterly catches for the start of the assessment model were assumed equal to the average 1971–1972 quarterly catches for that fleet. F7_TWN_Other included the aggregated catch for all other Taiwanese vessels and gear types, operating primarily nearer to Taiwan, including coastal longline, gillnet, setnet, harpoon, and other gears.

Two catch time series were available from the US. F11_US_LL included quarterly catches (in numbers) from longline gear in American Samoa and Hawai'i from 1993–2024. F12_US_Other was the aggregation of quarterly catch for all gear types, primarily troll and handline, recorded in Hawai'i from 1971–2024. This catch timeseries also included early estimates of longline catch in Hawai'i between 1971–1992, which was before the main development of the F12 longline fleet and only available in metric tons, so are maintained separately from the F11 timeseries starting in 1993.

Catch estimates were compiled from the IATTC and WCPFC. F8_IATTC_LL included

quarterly estimated catch of all longline vessels in the eastern Pacific Ocean as recorded by the IATTC for all flags other than Japan, Taiwan, and the US for 1983–2024. F9_WCPFC_LL included annual catch estimates, divided quarterly, of all longline vessels in the eastern Western Ocean as recorded by the WCPFC for all flags other than Japan, Taiwan, and the US for 1971–2024. F10_VAR_PS_Other was the summation of catch from all flags and gear types Pacific-wide that were not included in any of the other catch fleets from 1971–2024. These included catch reported by both the IATTC and WCPFC for purse seine, troll, handline, harpoon, and miscellaneous or unspecified gears.

Relative Abundance Indices

The eight CPUE indices available for inclusion in the model were assigned to the modal quarter of the highest catch of the corresponding harvest fleet based upon the recommendations of the country providing the index. This corresponded to the first quarter for all indices except S2_JPN_LateLL_East and S8_US_LLdeep, which were assigned to the second quarter. 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 CPUE indices for the Japanese “Early” period (S1) and “Late” period (S2 and S3) were divided in 1993 because of significant changes in the logbook reporting system, which required the CPUE data to be standardized in different time periods. The WG also agreed to divide the Japanese Late time period into an “East” (S2) and a “West” (S3) index, boundaried by 170°W longitude, to account for expected differences in the size selectivities of the longline fleets in these two areas. The Japan Late LL indices were also limited to the central fishing grounds (20°S–20°N latitude) in order to exclude inter-annual variation in fleet spatial coverage as fishing boats may move more pole-ward in some years.

There were 4 CPUE indices from Taiwan which were standardized over three separate time periods due to changes in the logbook reporting for the longline fleet, with breaks in 1978/1979 and 1999/2000. Similar to the late longline indices from Japan, the Taiwan late period longline indices were divided into “East” (S6) and “West” (S7) indices at 170°W longitude to account for expected differences in the size-based selectivity, and were also limited to the central fishing grounds (20°S–20°N latitude) to exclude the effects of spatiotemporal variability in fishing behavior or stock range.

There was a single CPUE index provided by the US (S8_US_LLdeep) which was based on logbook data for deep longline sets only; although shallow sets targeting swordfish do catch some BUM, most catch is on deep longline sets targeting tunas. Hence, choosing only a single set type for the index could help eliminate artifacts caused by high zero-catch sets from

the shallowset, and interannual variability in fishing fleet targeting behavior.

An *a priori* analysis of the CPUE indices was conducted to evaluate the potential for conflict within the indices. The analysis was performed based upon the methods from the 'diags' component of the FLCore package (Version 2.6.6, Kell *et al.* 2007) in R (version 3.4.0, R Core Team, 2026). This analysis provides 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 CVs for each CPUE index were assumed to be equal to the SE on the log scale. The minimum CV was scaled to an average of at least 0.2. If the input SE was greater than 0.2, no additional variance was added. RSME values for each index are listed in Table *2*.

Size Composition Data

Thirteen size composition time series were provided by year and quarter for consideration in the 2026 Stock Synthesis 3 (SS3) assessment model (Table 1); eight as length composition data (Figure *3*) and five as weight composition data (Figure *4* and Figure *5*). All length composition data were input as counts per 5 cm bin between 50 and 320 cm eye to fork length (EFL). In the data and as used in SS3, all length and weight bins are lower edge inclusive, e.g., the 70 cm length bin is all individuals 70 cm and larger, but less than 75 cm. The very few individuals recorded under 50 cm were added to the 50 cm bin in the input data.

Length composition data were available for all years of the Japanese longline fleets F1, F2, and F3 for 1971–2024, noting that lengths were divided between the east and west areas for fishing fleets F2_JPN_LateLL_East and F3_JPN_LateLL_West. Length composition data for survey fleets S2_JPN_LateLL_East and S3_JPN_LateLL_West were subset from F2 and F3, respectively, to include only fish caught between 20°S–20°N latitude to correspond to the reduced spatial extent of these survey fleets. Length data for the Japanese late longline fishery and survey fleets were initially included as a mix of specified female, male, and unknown (combined) lengths. However, initial analyses of these length data revealed that only 10–20% of fish identified to sex in these fleets were female, and it was unlikely that the wide spatial coverage of these fleets would cause such vastly different selectivity of the sexes, therefore, all length observations for Japanese fishing and survey fleets were input to the SS3 models as combined sex.

Length composition from IATTC purse seine vessels from 1990–2024 were used for F10_VAR_PS_Other. Length composition data were initially included starting in 1994 for F11_US_LL (for all longline set types), however, preliminary analyses revealed an abrupt shift in apparent size composition from small to larger fish in 2004 was likely due to a shift in reporting, not behavior of the fishing fleet. Therefore, only size data starting in 2004 were included for F11_US_LL. Length data for S8_US_LLdeep were from only the deepset longline fishery. Length data for both F11_US_LL and S8_US_LLdeep were included as a mix of specified female, male, and combined lengths, allowing for differential female and male selectivity.

Size composition data were provided by weight (kg) for F4_JPN_DRIFT_BAIT and four of the Taiwan longline fleets. Weight compositions are referred to as "generalized size compositions" in SS3 and bins can be specified differently among fleets or irregularly binned to be more narrow at smaller weights when fish are expected to grow quickly, and wider at larger weights. For F4_JPN_DRIFT_BAIT, weight data were available sporadically between 1977 and 1998, and were binned in 10 kg increments from 30 to 150 kg, and 20 kg increments from 160 to 320 kg, with individuals less than 30 kg comprising the smallest bin. Weight composition data were available from 2014–2024 for F5_TWN_LTLL and 2017–2024 for F6_TWN_STLL, binned in 10 kg increments from 20 to 110 kg, and 20 kg increments from 120 to 260 kg, with individuals less than 20 kg comprising the smallest bin. Weight composition data for survey fleets S6_TWN_LTLL_LateEast and S7_TWN_LTLL_LateWest were subset from F5_TWN_LTLL to include only fish caught between 20°S–20°N latitude and further divided between east and west areas (boundaried by 170°W longitude) to correspond to the spatial extent of these indices.

All size composition data were available in quarterly time steps. Data were minimally filtered and initial sample sizes (effective sample size; effN) were calculated consistently across all fleets for SS3 input. First, any quarter with fewer than 15 individuals measured were excluded from the data. For fleets with sex-specific length data, length information was entered in two matrices: an combined sex matrix and a female-male matrix. Any quarter for female or male length composition with less than 30 individuals measured was re-assigned to combined sex. The effN input to SS3 were calculated as the number of individuals measured for each quarter and sex (combined or female-male), divided by 10, and capped at 50 to reduce the influence of quarters with very large sample sizes.

Within the SS3 model, size composition data were further tuned among fleets and downweighted relative to other likelihood components following a modified Francis (2011) multinomial distribution T.A1.8 approach. First, the model was fit on the input effN values described in the previous paragraph, with all CPUE indices included. The resulting fitted

model variance was compared to the input data variance using the r4ss function tune_comps() to yield the recommended Francis (2011) multinomial distribution T.A1.8 effN multiplier for each size composition fleet. These "tuning" suggestions are scalars that can be used to adjust the effN for each size composition fleet, and balance the relative informativeness among the size composition fleets. Values greater than 1 decrease the observed variance of the input data, giving it more "weight" within the SS3 likelihood function. For the 2026 BUM assessment, these suggested tuning scalars ranged from 1.3 to 6 (Table 3). In order to reduce the contrast between fleets, these values were natural log transformed and then scaled to 0.2 to reduce the overall influence of size compositions within the model likelihood. In order to prevent any fleet's size composition data from losing influence in the likelihood altogether, the final effN adjustment values were assigned a minimum of 0.1.

Base-case model description

The assessment was conducted with Stock Synthesis (SS3) version V3.30.24.1-safe released September 18, 2025 using ADMB version 13.2. (Methot and Wetzel, 2013). All model specifications were discussed at the BILLWG Data Preparatory Meeting (ISC BILLWG, 2026). The model was set up as a single area model with two sexes and four seasons (quarters). Spawning was assumed to occur in April, while annual recruitment was partitioned equally across 3 settlement events in May, July, and September. Fish at each settlement event were assigned age 0 at settlement and experienced growth and natural mortality according to their real ages. Sex-specific growth parameters were from the Chang *et al*. (2025) 2-stanza growth model, re-parameterized for the SS3 Schnute-Richards formulation using Growth_Age_for_L1= 0.5 years and Growth_Age_for_L2 = 27 years (Table *4*; Figure *6*). Weight and female maturity were logistic functions of length (Table 2; Figure *7* Figure *8*) and the fraction female at recruitment was fixed at 0.5. The first mature age was 2 years and fecundity was proportional to female spawning biomass. Natural mortality was sex- and age-specific based on the maximum age estimator using Lorenzen scaling of the updated Chang et al. (2025) two-stage growth model as detailed in Brodziak, 2026 (Table 5).

Stock-recruitment followed a Beverton-Holt spawner-recruit relationship with steepness (h) fixed at 0.87 for the base models (Brodziak and Mangel 2011). Main recruitment deviations were estimated from 1975–2024 with sum to zero and sigmaR (σ_r) fixed equal to 0.4, and early recruitment deviations were estimated from 1961–1974 to enable the population to have non-equilibrium length and age composition. Recruitment deviations were bias-adjusted based upon the estimates from Methot and Taylor (2011), ramping linearly between zero in 1954 to 0.77 in 1998, with full bias adjustment continuing until the terminal year of the model.

The model start year was 1971 and the fishery was assumed to be in an initial fished state by the F1_JPN_EarlyLL and F6_TWN_STLL fleets. Initial conditions were fixed using the average of the first two-years of catch for each fleet. When initial conditions were freely estimated, the depletion in 1971 was over 50%, and with initial conditions fixed to the early average catch, it was around 80% depleted. Runs starting in 1950 with the assumption that the stock was unfished at the start of the model produced biomass trajectories and depletion consistent with fixing initial F. Additional runs varying the initial equilibrium catch showed initial depletion to be fairly constant. Catchability (q) was freely estimated for all survey fleets.

The SS3 population model used a 5 cm length bin from 30 to 320 cm, with 320 cm being the accumulator bin for larger fish. Fishery selectivity was length-based for all fleets, meaning removals from fleets with weight composition data are converted to length within the model. The seven fleets that lacked size composition data were assigned to mirror the selectivity of the fleet with size data and the most similar fishing practices (Table 6). The length-based selectivities of F7_TWN_Other, F8_IATTC_LL, F9_WCPFC_LL, S4_TWN_LTLL_Early, and S5_TWN_LTLL_Mid were all assumed to be the same as F5_TWN_LTLL. Length-based selectivity of F12_US_Other mirrored F11_US_LL and S1_JPN_EarlyLL mirrored F1_JPN_EarlyLL.

For most fleets with size data, selectivity was specified as a dome-shaped function of length, using the 4-parameter “dblnorm” pattern (24) in SS. The parameters were unconstrainted, therefore the model had the flexibility to mimic logistic (asymptotic) selectivity for these fleets. F11_US_LL and S8_US_LLdeep, which had sex-specific length composition data input to the model, each had 4 additional parameters to allow for the estimation of differential male and female selectivity. The only fleet with specified logistic selectivity (2-parameter pattern 1 in SS) was F10_VAR_PS_Other, which showed the overall largest size fish caught among all fleets during initial data examination and can reasonably be assumed to be able to catch the largest fish.

An important consideration for the 2026 BUM assessment model is that the size composition data for the survey fleets are not distinct observations, instead they are a subset (e.g., by year or area) of the same length composition observations already used for the fishing fleets. Therefore, including survey fleets’ size composition data within the likelihood of the model would over-represent the true information and appropriate influence of these data, as they would be essentially counted twice. So, models were parameterized using a 2-step process. During the first step, all fleets’ size composition data were included in the likelihood (objective function to be minimized by ADMB) and selectivity parameters were estimated. In the second step, selectivities for the five survey fleets were fixed to these values and the

contribution of these fleets' size composition data was removed from the likelihood.

Model estimated time series of female spawning biomass (SSB mt) and recruitment (R in 1000s of fish) were tabulated on an annual basis. The annual exploitation rate (F) was calculated as Catch / Biomass. Stock status indicators were calculated relative to the SSB and F at maximum sustainable yield (MSY).

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 and length composition 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 9, to compare trends by stock. Generally, the only two indices with moderate negative correlations are the US LL index and the Taiwan Mid LTLL index. The Japan LL late west index appears to show the abundance in the west Pacific have been in a steady decline since the 1990s, although this trend is not observed in any other fleet. Many indices do show a decline in abundance through the 2000s, but show a flat or increasing trend thereafter. Both the early Japan and Taiwan LL indices show a rapid increase in the 1970-1980s, followed by a decline into 1990.

To look at deviations from the overall trends, the residuals from the LOESS smoothers to each observation within an index are compared in Figure 10. 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. The Japan LL Late west, and Both Taiwan LTLL late indices show a sharp increase in CPUE in the final year (2024). This is consistent with a large increase in the catch of small BUM throughout the Pacific, and the WG believes this to be a strong recruitment pulse from the 2020s COVID period.

Figure 11 illustrates the correlation between indices; the lower-left triangle displays the pairwise scatter plots of one index plotted against another with a linear regression line, the upper right triangle displays the correlation coefficients, and the diagonal displays the smoothed frequency of observation values. 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 TWN LL early fleet (corr = 0.0.68) and the Japanese late LL West appears to be positively correlated with the TWN LTLL Mid (corr =0.81). The US LL has a high negative and statistically significant correlation with Taiwan LTLL Mid (corr = -0.49), which could indicate potential conflict between these two indices. The US LL index is also positively correlated with the JPN LL Late East index (corr = 0.77). If we compare indices that index the same area of the Pacific, the JPN LL Late East index is positively correlated with the TWN LTLL Late East index (corr = 0.26), and the JPN LL Late west index appears to be positively correlated with the TWN LL Late West index, although that correlation may be driven by the 2024 value (corr = 0.68).

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 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 12 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. Two clusters were identified with the generally the US and JPN indices in one cluster and the TWN indices in the second.

Overall, the results of the CPUE analysis suggest that there may be the potential for conflict between the candidate CPUE indices from JPN and TWN . Due to this analysis and discussions with the modeling team, two models are provided for consideration, one including the US and JPN CPUE indices, and one including the US and TWN indices.

Results

Japan Model

Model fit

The Japan base-case model ran in about 2 hours 30 minutes, estimated 112 parameters, and had a total adjusted loglikelihood of 169.433. 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 4×10^{-5}. None of the parameter estimates hit a bound, no parameters had correlations above 0.95 and two selectivity parameters had correlations below 0.01. A plot of the estimated recruitment deviations and their uncertainty can be found in Figure 14 All of the early recruitment deviations (1960-1974) and the main recruitment deviations had CVs > 50%. Twelve of 34 selectivity parameters had CVs >50%. These parameters were from the dome-shaped selectivity functions and seven were the width of the plateau, the rest were poorly estimated parameters for the sex-specific selectivity for the US LL fleet.

Fits to the abundance indices were moderate, the model fits the US Hawaii index and JPN LL early reasonably well, but not the JPN LL late indices Figure 15 and Figure 17Figure 17. U.S. CPUE indices (points) and fits of the Japan (blue) and Taiwan (red) models. Error bars are 95% confidence intervals of the adjusted variance of the index.. Both of these indices failed the runs test and show substantial patterning in their residuals (Figure 18).

Estimated selectivity for each fleet are in Figure 24 through Figure 28a. Fits to the length composition data were also poor(Figure 29- Figure 41), although the annual trends were generally flat for all fleets. Results from the runs test indicate that only one fleet passed the runs test, F2_JPN_LLLate_East (Figure 19 and Figure 20).

Model estimates of age 1+ biomass show an initial decrease in biomass from 1971 to 1980, then biomass flattened to 1994, declined to its lowest level in 2021, and has slightly increased since. Biomass is on a slight increasing trend in the last three years of the assessment model (エラー! 参照元が見つかりません。). Initial female spawning stock biomass was estimated to be approximately 118, 000 mt and virgin SSB was around 170,000 mt (Figure 48). Annual fishing mortality is reported as the exploitation rate in biomass (Figure 49). Fishing mortality was below MSY until 2000 and above MSY for the rest of the time series. Recruitment was variable but the log of the deviations were generally between 0.4 and -0.4 (Figure 50). Early recruitment deviations were negative to better fit the JPN early size data. Current depletion, as estimated as the age 1+ biomass in 2016 compared to the virgin age 1+ biomass was estimated to be 0.20.

Diagnostics

Profiling on R_0 showed that the recruitment estimates were highly influential in the model results, especially below the MLE estimate while length composition data provided the majority of the information for the upper bound of the MLE estimate (Figure 42Figure 45). The CPUE indices contributed to the lower bound of the MLE estimate close to the MLE value. Generally, there was conflict between the CPUE and length and size composition data, with the CPUE indices pushing for a lower ln(R0) than the size and length data, This conflict is driven mainly by F2 the JPN LL Late East index. Results from the ASPM model are reasonably consistent with the full model, with similar population trajectories and scale, suggesting the CPUE indices do contain enough information to inform the productivity function and drive the dynamics of the model (Figure 51).

Taiwan Model

Model fits

The Taiwan base-case model ran in about 2 hours 30 minutes, estimated 113 parameters, and had a total adjusted loglikelihood of 189.171. 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 8×10^{-6}. None of the parameter estimates hit a bound, no parameters had correlations above 0.95, and two selectivity parameters had correlations below 0.01. A plot of recruitment deviations can be found in Figure 14. All of the early recruitment deviations (1960-1974) and 47 of 50 94% of the main recruitment deviations had CVs > 50%. Twelve of 34 selectivity parameters had CVs >50%. Seven parameters were from the dome-shaped selectivity functions and were the width of the plateau. The rest were poorly estimated parameters for the sex-specific selectivity for the US LL fleet..

Fits to the abundance indices were relatively good for the indices from TWN (S4, S5, S6, and S7) as all passed the runs test (Figure 16, Figure 17, and Figure 21Figure 23). A very low CPUE in 2000 for the TWN LTLL Late East fleet was very influential to the fit of that index and a large outlier, and should be considered for removal from future model runs. Fits to the US LL index were poor and failed the urns test, which is to be expected as the correlation analysis suggested that the US LL index was negatively correlated with several of the TWN indices.

Estimated selectivity for each fleet are in Figure 24Figure 28a. Fits to the length composition data were also consistent with the Japan model, and annual trends tended to be flat (Figure 29Figure 41). Only three of the length composition time-series passed the runs test,

F3_JPN_LLLate_East, S2_JPN_LLLate_East, and S3_JPN_LLLate_West, with both of the CPUE fleets having fixed selectivity in the estimation model (Figure 22 Figure 23)

Model estimates of age 1+ biomass show a long term decline in biomass since 1971 through 2000, where it remained fairly stable until 2020, and has drastically risen in the last four years of the model (Figure 47). Initial female spawning stock biomass was estimated to be approximately 117,000 mt and virgin SSB was around 169,000 mt (Figure *48*). Annual fishing mortality is reported as the exploitation rate in biomass (Figure *49*). Fishing mortality was below MSY from 1971-1994, and 1996-1998. It was just above MSY in 1995 and from 1999-2019. Fishing mortality has been below MSY since 2020.. Recruitment was variable but the log of the deviations were generally between 0.4 and -0.4, a very low recruitment estimated in 1997-1999 was likely driven by a very low CPUE value in 2000 from the TWN LTLL Late west index (エラー! 参照元が見つかりません。). Current depletion, as estimated as the age 1+ biomass in 2024 compared to the virgin age 1+ biomass was estimated to be 0.32.

Diagnostics

Profiling on R_0 showed that the recruitment provided the majority of the information on the lower bound of R_0 while the length composition data and CPUE indices provided information on the upper bound of R_0 (Figure 43). There is substantial conflict between the length composition data and the CPUE data, with the CPUE data preferring a minimum MLE of ln (R0) around 6.4, and the length composition data preferring a minimum around 6.8 (Figure 43Figure 45). This conflict is primarily driven by the TWN LTLL Late west and the TWN LTLL Mid indices. Results from the ASPM model showed the same population trend as the full model however, the scale of the population was significantly higher in the APSM model (Figure 52). This, and the very large uncertainty around the ASPM estimates suggests that the CPUE data do not provide information on the size of the population, but does have some information on the trajectory.

Conclusions

Both models show some problems with fit to the available data. The Japan model fails to fit the CPUE indices adequately as two of the four indices fail the runs test. Fits to the size composition data for the US LL sex specific data are poorly estimated and may need to be combined into the unknown sex data. The likelihood profile on $\ln(R_0)$ suggests that the JPN LL Late West index is in conflict with the rest of the CPUE indices and size and length composition data. This is consistent with the correlation analysis findings that suggest this index is not tracking the same abundance trends as the rest of the CPUE fleets.

The Taiwan model fits to the CPUE indices appear to be better than the Japan model, but fits to the US LL index is poor and fails the runs test. There also is more conflict in the model between the CPUE indices for the US LL index and the TWN LTLL Late west index. Generally the Taiwan model fits the size and length composition data better than the Japan model, although difference between the two models are minimal for these fleets.

Overall, both models show conflict and problems with fitting to the data that need to be further explored before a base-case model can be recommended. At a minimum, a model run and diagnostics for a model dropping only the JPN LL Late east index should be explored to better understand the contribution of that index to the model, and collapsing the US length composition data into a single combined time series should be undertaken.

References

Andrews, A.H., Humphreys Jr, R.L. and Sampaga, J.D., 2018. Blue marlin (*Makaira*
Brodziak, J. and Mangel, M. 2011. Probable values of Stock-Recruitment Steepness for North Pacific Striped Marlin. ISC/11/BILLWG-02/11.

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

Brodziak, J. 2026. Natural Mortality Rates for the Pacific Blue Marlin Stock. ISC/26/BILLWG-01/05.

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., Shimose, T. Chang, X.B., Furuyama, A., Hsu, J. Kanaiwa, M., Chiang, W.C., Brodziak, J.K.T., Kai, M., Ijima, H. Sculley, M.L., and Kinney, M. 2025. Estimation of two-stanza growth curves with age determination uncertainty for Pacific blue marlin (*Makaira nigricans*) in the western and central North Pacific Ocean. ICES Journal of Marine Science, 82(8): 1-15.

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.

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

ISC BILLWG. (2026). Report of the Billfish Working Group Workshop. 13-19, January 2026.

Jusup, M. and Kai. M. 2026a. Update of the Japanese longline abundance index for Pacific blue marlin (Makaira nigricans) using the habitat model (1994-2024). ISC/26/BILLWG-01/06.

Jusup, M., and Kai, M. 2026b. Update of the Japanese catch and length-frequency data for Pacific blue marlin (Makaira nigricans) from 1971 to 2024. ISC/26/BILLWG-01/09.

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., 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. 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 2026. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sculley M. (2026a). Standardization of Pacific Blue Marlin Catch Per Unit Effort in the Hawaii Longline Fishery from 1995-2024. ISC/26/BILLWG-01/08.

Sculley M. (2026b). Summary of US Catch and Size of Pacific Blue Marlin. ISC/26/BILLWG-01/11.

Shimose T, Fujita M, Yokawa K, Saito H, Tachihara K. 2009. Reproductive biology of blue marlin *Makaira nigricans* around Yonaguni Island, southwestern Japan. Fish Sci, 75:109119.

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.

Su, N.J, Sun, C.L., Punt, A.E., Yeh, S.Z, and DiNardo, G. 2011. Modelling the impacts of environmental variation on the distribution of blue marlin, *Makaira niricans*, in the Pacific Ocean. ICES Journal of Marine Science. 68(8): 1072-1080.

Yeh, Z.W., and Chang, Y.J. 2026a CPUE standardization of blue marlin (*Makaira nigricans*) for the Taiwanese distant-water tuna longline fishery in the Pacific Ocean during 1971 - 2024. ISC/26/BILLWG-01/07.

Yeh, Z.W., and Chang, Y.J. 2026b. Blue marlin (*Makaira nigricans)* catch and size data of Taiwanese fisheries in the Pacific Ocean. ISC/26/BILLWG-01/10.

Tables and Figures

Table 1. List of fleets with Catch and CPUE indices provided for the 2026 Pacific Blue Marlin Stock Assessment and the source for more information about the standardization of the CPUE series and the size composition data. Size data are available as combined sex (U), female (F), and male (M).

Fleet name	Fleet number	Member	Description	Units	Size data	Source
F1_JPN_EarlyLL	1	Japan	Coastal, offshore, and distant water longline, other gears, 1971–1993	B	Length (U), 1971–1993	Ijima 2021, Jusup and Kai 2026b
F2_JPN_LateLL_East	2	Japan	Offshore and distant water longline east of 170°W, 1994–2024	B	Length (U), 1994–2024	Jusup and Kai 2026b
F3_JPN_LateLL_West	3	Japan	Coastal, offshore, and distant water longline, west of 170°W, 1994–2024	B	Length (U), 1994–2024	Jusup and Kai 2026b
F4_JPN_DRIFT_BAIT	4	Japan	High-seas large-mesh driftnet, coastal driftnet, and bait fishing 1972–2024	B	Weight (U), some years from 1977–1998	Jusup and Kai 2026b
F5_TWN_LTLL	5	Taiwan	Large tonnage longline, 1971–2024	B	Weight (U), 2014–2024	Yeh and Chang 2026b
F6_TWN_STLL	6	Taiwan	Small tonnage longline, 1971–2024	B	Weight (U), 2017–2024	Yeh and Chang 2026b
F7_TWN_Other	7	Taiwan	Coastal longline, gillnet, setnet, harpoon, and others, 1971–2024	B	NA: Selectivity mirrors F5_TWN_LTLL	Yeh and Chang 2026b
F8_IATTC_LL	8	Various flags	Longline, 1983–2024	B	NA: Selectivity mirrors F5_TWN_LTLL	IATTC
F9_WCPFC_LL	9	Various flags	Longline, 1971–2024	B	NA: Selectivity mirrors F5_TWN_LTLL	WCPFC
F10_VAR_PS_Other	10	Various flags	All gears and flags not otherwise included: purse seine, troll, handline, harpoon, others, 1971–2024	B	Length (U), 1990–2024 from IATTC purse seine (only fleet w/ *a priori* logistic selectivity)	IATTC, WCPFC

F11_US_LL	11	USA	American Samoa and Hawaii, longline shallow and deep combined, 1993–2024.	Num	Length (U/F/M), 2004–2024. Sex-specific selectivity	Sculley 2026b
F12_US_Other	12	USA (Hawaii)	Troll, handline, 1971–2024, and longline, 1971–1992	B	NA: Selectivity mirrors F11_US_LL	Sculley 2026b
S1_JPN_EarlyLL	13	Japan	Offshore and distant water longline, 1975–1993	Num	NA: Selectivity mirrors F1_JPN_EarlyLL	Ijima 2021
S2_JPN_LateLL_East	14	Japan	Offshore and distant water longline east of 170°W, 20°S–20°N, 1994–2024	Num	Length (U), 1994–2024	Jusup and Kai 2026a, Jusup and Kai 2026b
S3_JPN_LateLL_West	15	Japan	Offshore and distant water longline west of 170°W, 20°S–20°N, 1994–2024	Num	Length (U), 1994–2024	Jusup and Kai 2026a, Jusup and Kai 2026b
S4_TWN_LTLL_Early	16	Taiwan	Large tonnage longline, 1971–1978	Num	NA: Selectivity mirrors F5_TWN_LTLL	Yeh and Chang 2026a
S5_TWN_LTLL_Mid	17	Taiwan	Large tonnage longline, 1979–1999	Num	NA: Selectivity mirrors F5_TWN_LTLL	Yeh and Chang 2026a
S6_TWN_LTLL_Late_East	18	Taiwan	Large tonnage longline east of 170°W, 20°S–20°N, 2000–2024	Num	Weight (U), 2014–2024	Chang, Pers. Comm.
S7_TWN_LTLL_Late_West	19	Taiwan	Large tonnage longline west of 170°W, 20°S–20°N, 2000–2024	Num	Weight (U), 2014–2024	Chang, Pers. Comm.
S8_US_LLdeep	20	USA (Hawaii)	Longline, deep-set only, 1995–2024	Num	Length (U/F/M), 2004–2024. Sex-specific selectivity	Sculley 2026a, Sculley 2026b

Table 2. Mean initial CV, final CV, and additional variance added *a priori* in the base-case models for the eight CPUE Indices.

Fleet	Initial CV (average across years)	Final CV (average across years)	Added Variance each year
S1_JPN_EarlyLL	0.198	0.200	0.002
S2_JPN_LateLL_East	0.088	0.200	0.112
S3_JPN_LateLL_West	0.139	0.200	0.061
S4_TWN_LTLL_Early	0.295	0.295	0
S5_TWN_LTLL_Mid	0.174	0.200	0.026
S6_TWN_LTLL_Late_East	0.136	0.200	0.064
S7_TWN_LTLL_Late_West	0.136	0.200	0.064
S8_US_LLdeep	0.203	0.203	0

Table 3. Effective sample size (effN) adjustment values applied to size composition fleets within the models.

Fleet name	Size data type	Suggested Francis T.A1.8 value	Ln(value)	Scaled to maximum 0.2	Final effN adjustment values
F1_JPN_EarlyLL	Length	2.912	1.069	0.119	0.119
F2_JPN_LateLL_East	Length	2.682	0.987	0.11	0.11
F3_JPN_LateLL_West	Length	1.325	0.281	0.031	0.1
F4_JPN_DRIFT_BAIT	Weight	2.002	0.694	0.077	0.1
F5_TWN_LTLL	Weight	2.137	0.759	0.085	0.1
F6_TWN_STLL	Weight	2.516	0.923	0.103	0.103
F10_VAR_PS_Oth	Length	4.181	1.431	0.16	0.16
F11_US_LL	Length	4.948	1.599	0.179	0.179
S2_JPN_LateLL_East	Length	2.382	0.868	0.097	0.1
S3_JPN_LateLL_West	Length	3.591	1.278	0.143	0.143
S6_TWN_LTLL_Late_East	Weight	1.791	0.583	0.065	0.1
S7_TWN_LTLL_Late_West	Weight	5.994	1.791	0.2	0.2

S8_US_LLdeep	Length	5.727	1.745	0.195	0.195

Table 4. Key life history parameters used in the 2026 Pacific Blue Marlin Stock Assessment. From Table 2 in the ISC BILLWG Data Preparatory report (2026).

Parameter	Value	Reference
Growth_Age_for_L1	0.5	
Growth_Age_for_L2	27	
L_at_Amin_Fem	134.69	Chang et al. 2025
L_at_Amax_Fem	267.54	Chang et al. 2025
VonBert_K_Fem	0.22	Chang et al. 2025
Richards_Fem	0.77	Chang et al. 2025
CV_young_Fem	0.08	Chang et al. 2025
CV_old_Fem	0.09	Chang et al. 2025
Wtlen_1_Fem	0.00001844	Brodziak 2013
Wtlen_2_Fem	2.956	Brodziak 2013
Mat50%_Fem	179.76	Sun et al. (2009), Shimose et al. (2009)
Mat_slope_Fem	-0.2039	Sun et al. (2009), Shimose et al. (2009)
L_at_Amin_Mal	133.28	Chang et al. 2025
L_at_Amax_Mal	199.52	Chang et al. 2025
VonBert_K_Mal	0.12	Chang et al. 2025
Richards_Mal	0.000468	Chang et al. 2025
CV_young_Mal	0.07	Chang et al. 2025
CV_old_Mal	0.08	Chang et al. 2025
Wtlen_1_Mal	0.0000137	Brodziak 2013
Wtlen_2_Mal	2.975	Brodziak 2013

Table 5. Natural mortality.

	Natural Mortality Rate at Age		Standard Deviation	
Age	Female	Male	Female	Male
0	0.40	0.45	0.13	0.15
1	0.32	0.37	0.11	0.12
2	0.29	0.35	0.10	0.12
3	0.26	0.33	0.09	0.11

4	0.25	0.32	0.08	0.11
5	0.23	0.31	0.08	0.10
6	0.22	0.30	0.07	0.10
7	0.22	0.30	0.07	0.10
8	0.21	0.29	0.07	0.10
9	0.21	0.29	0.07	0.10
10	0.20	0.29	0.07	0.10
11	0.20	0.29	0.07	0.10
12	0.20	0.29	0.07	0.10
13	0.20	0.28	0.07	0.09
14	0.19	0.28	0.06	0.09
15	0.19	0.28	0.06	0.09
16	0.19	0.28	0.06	0.09
17	0.19	0.28	0.06	0.09
18	0.19	0.28	0.06	0.09
19	0.19		0.06	
20	0.19		0.06	
21	0.19		0.06	
22	0.19		0.06	
23	0.19		0.06	
24	0.19		0.06	
25	0.19		0.06	
26	0.19		0.06	
27	0.19		0.06	

Table 6. List of selectivities for each fleet in the BUM assessment model, Estimated indicates if they are estimated in the final model run.

FLEET	SELECTIVITY	ESTIMATED?
F1_JPN_EARLYLL	Double normal	Yes
F2_JPN_LATELL_EAST	Double normal	Yes
F3_JPN_LATELL_WEST	Double normal	Yes
F4_JPN_DRIFT_BAIT	Double normal	Yes
F5_TWN_LTLL	Double normal	Yes
F6_TWN_STLL	Double normal	Yes
F7_TWN_OTHER	Mirror F5	-
F8_IATTC_LL	Mirror F5	-
F9_WCPFC_LL	Mirror F5	-
F10_VAR_PS_OTHER	Asymptotic logistic	Yes
F11_US_LL	Double normal	Yes
F12_US_OTHER	Mirror F11	-
S1_JPN_EARLYLL	Mirror F1	-
S2_JPN_LATELL_EAST	Double normal	Estimated, then fixed
S3_JPN_LATELL_WEST	Double normal	Estimated, then fixed
S4_TWN_LTLL_EARLY	Mirror F5	-
S5_TWN_LTLL_MID	Mirror F5	-
S6_TWN_LTLL_LATE_EAST	Double normal	Estimated, then fixed
S7_TWN_LTLL_LATE_WEST	Double normal	Estimated, then fixed
S8_US_LLDEEP	Double normal, sex specific	Estimated, then fixed

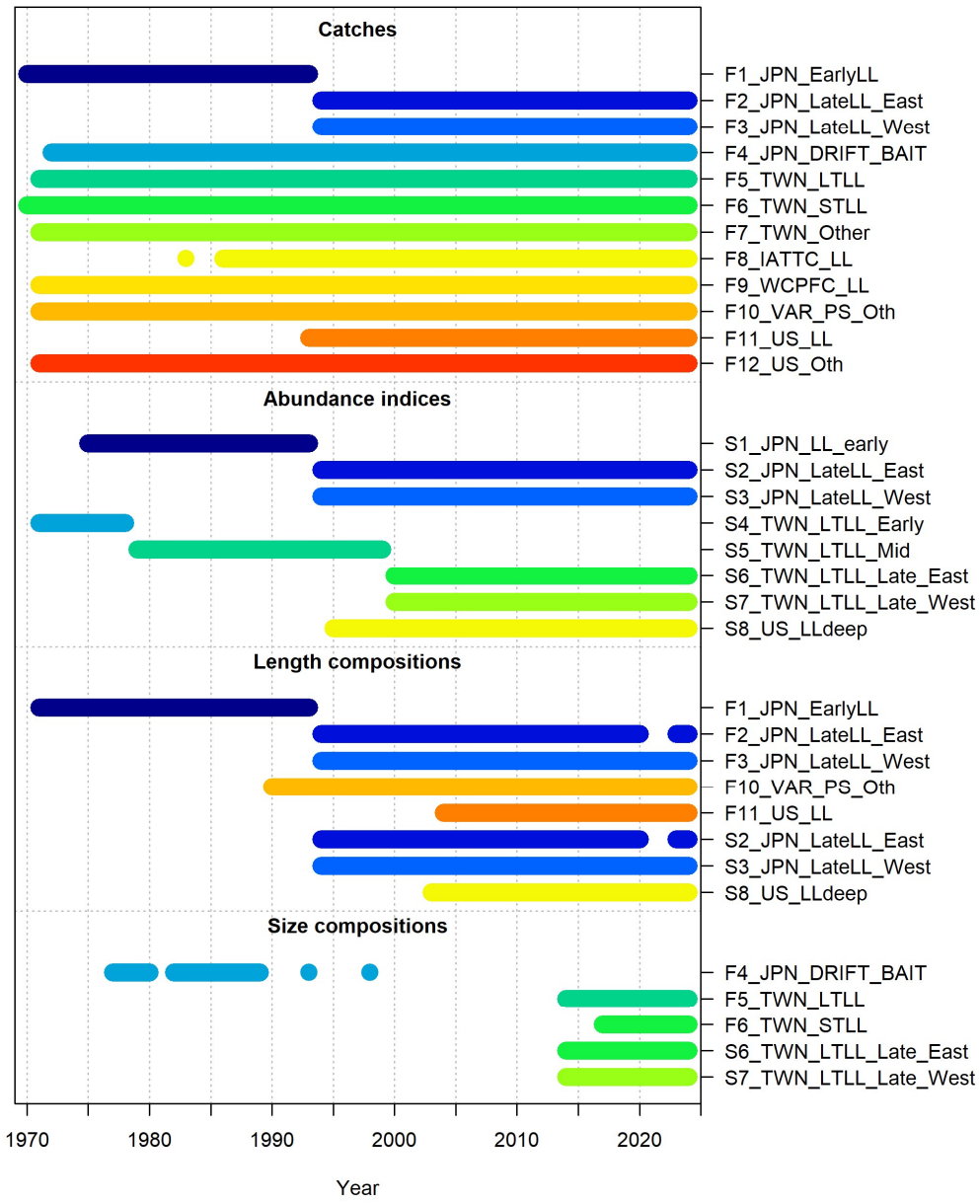


Figure 1. Catch, CPUE index, and size composition data included in the 2026 BUM stock assessment.

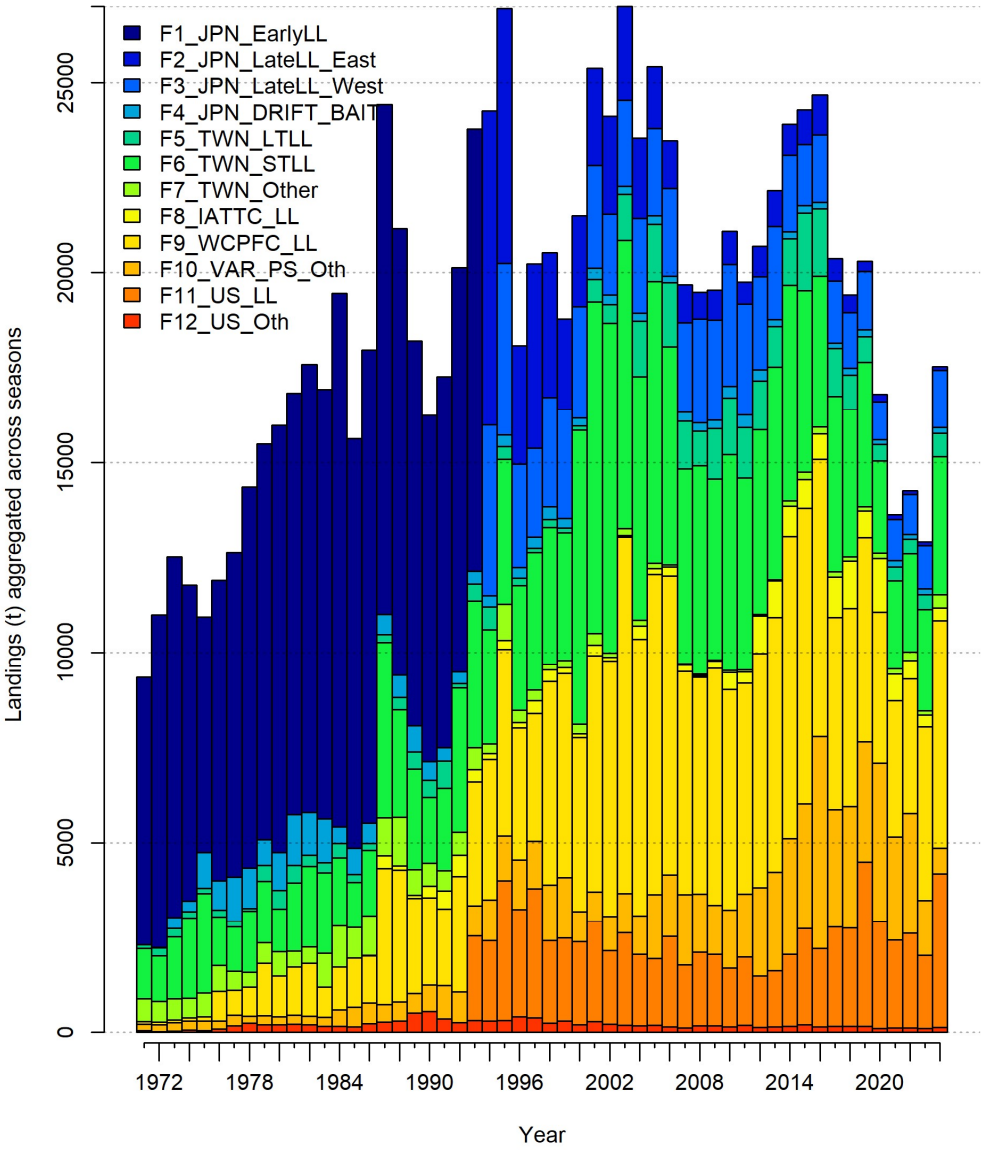


Figure 2. Annual catch of Pacific blue marlin in metric tons by fleet in the 2026 base-case assessment model. Catch was reported quarterly but has been summed annually for clarity. Note that catch for Fleet 11, US Longline (F11_US_LL) was provided in terms of number, but is shown in this figure as biomass calculated from the base case model to allow comparable scales accross fleets.

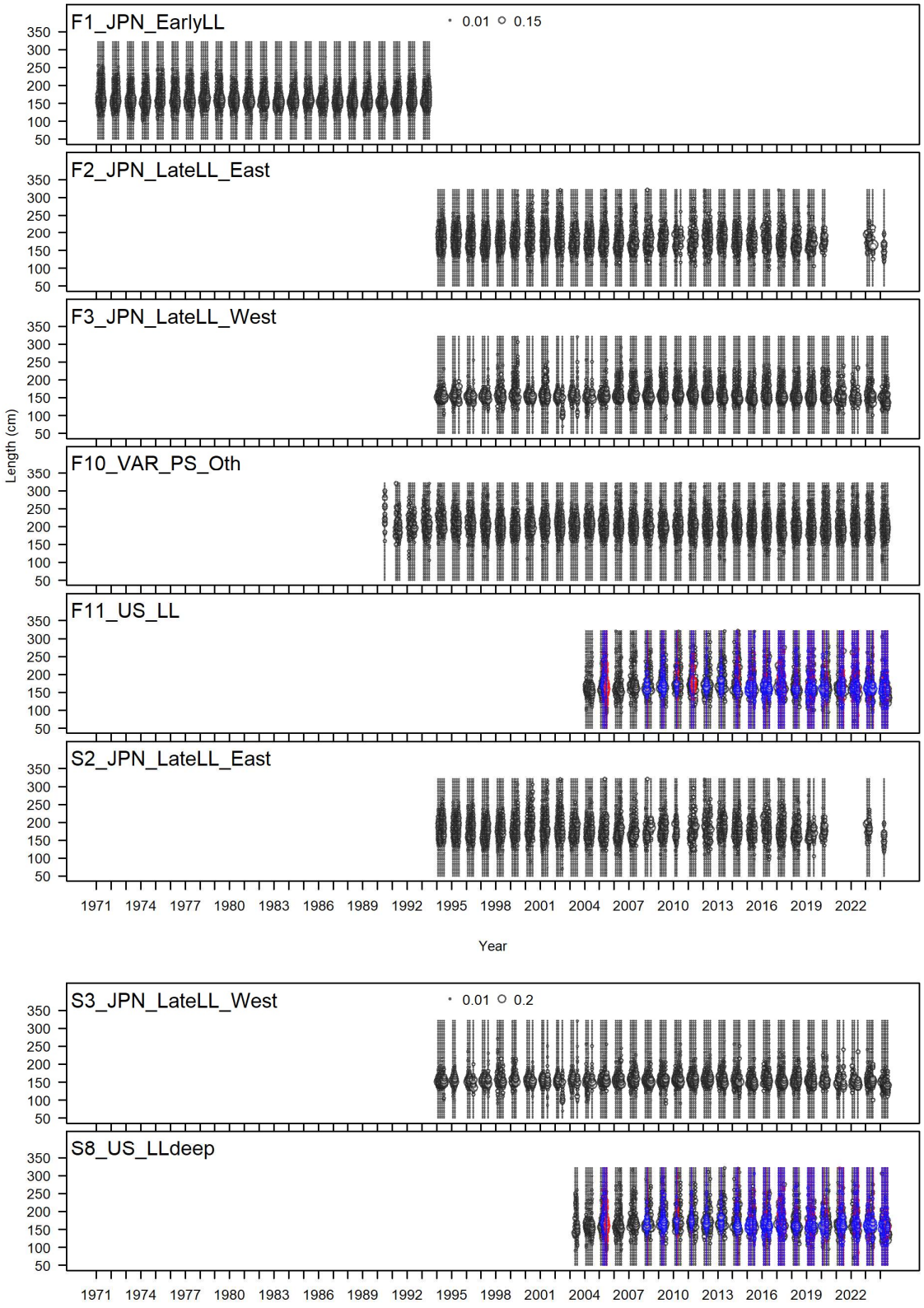


Figure 3. Length Composition data available in 5cm size bins for the 2026 Pacific blue marlin stock assessment. F11_US_LL and S8_US_LLdeep had sex-specific length composition. Data for females is red, males is blue, and unspecified is uncolored.

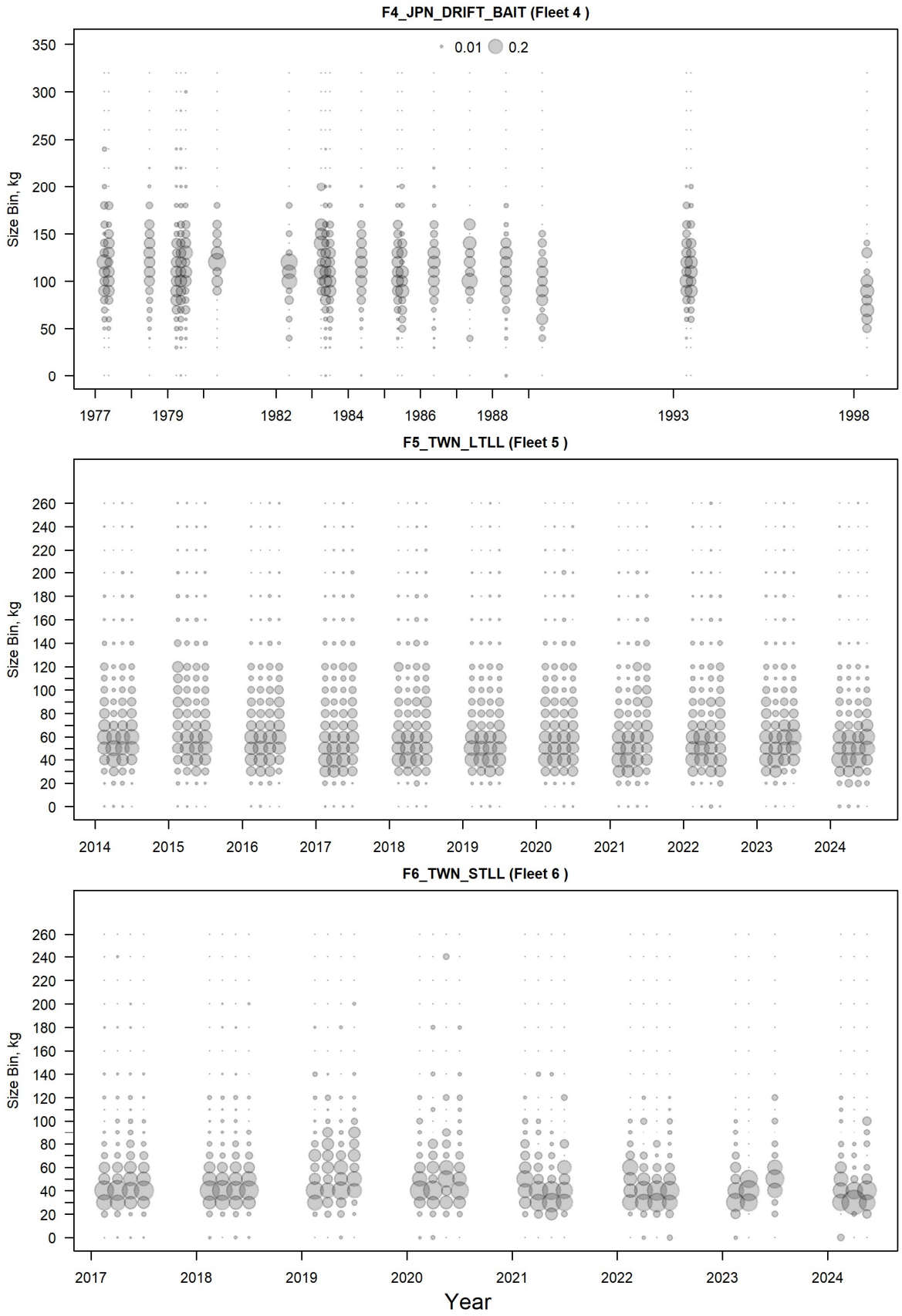


Figure 4. Weight composition data available F4, F5, and F6.

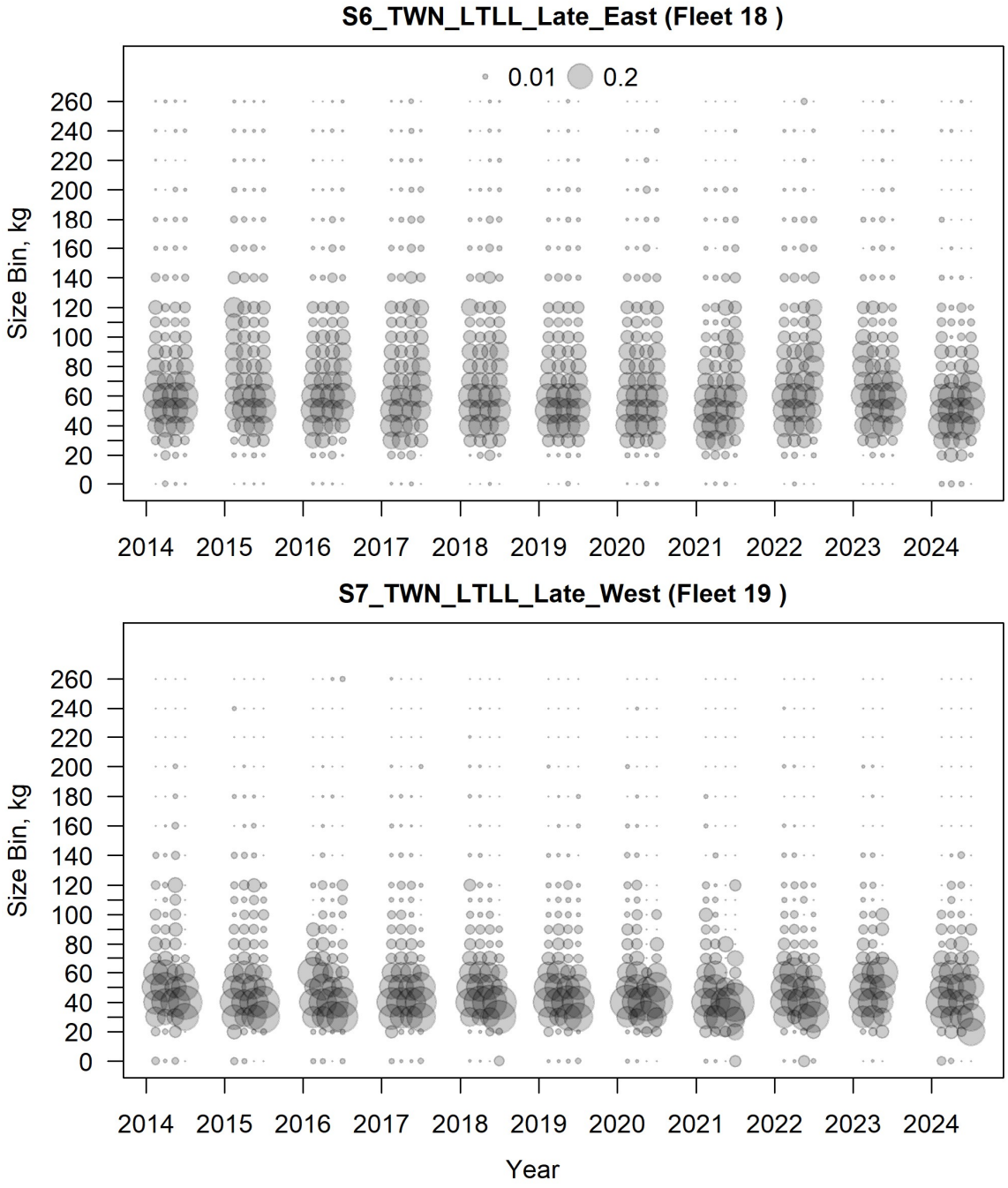


Figure 5. Weight composition data available.

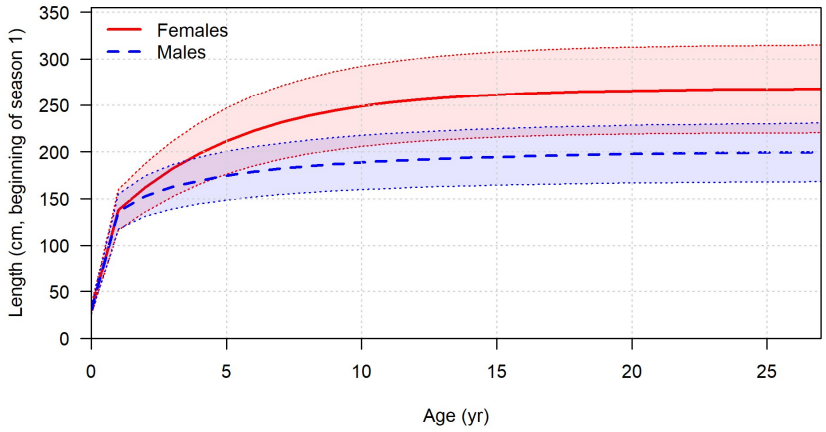


Figure 6. Length as a function of age of female (red) and male (blue) BUM in the 2026 assessment following the two-stanza growth model from Chang et al. (2025). Shaded areas are the 95% distribution of length.

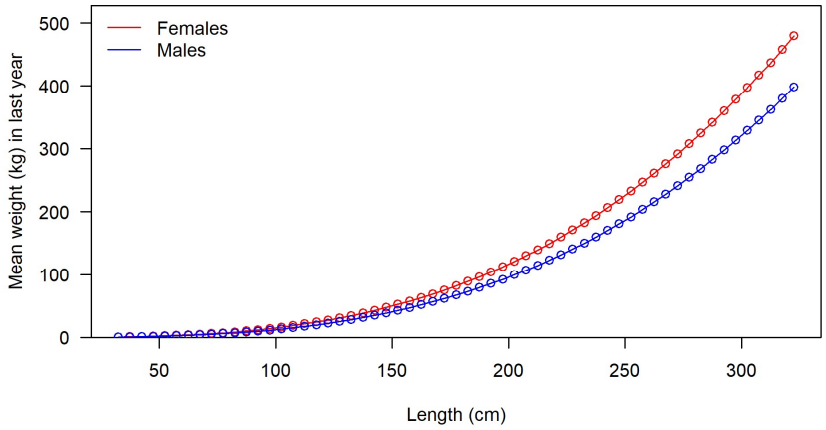


Figure 7. Weight as a function of length for female (red) and male (blue) BUM in the 2026 assessment based on [citation].

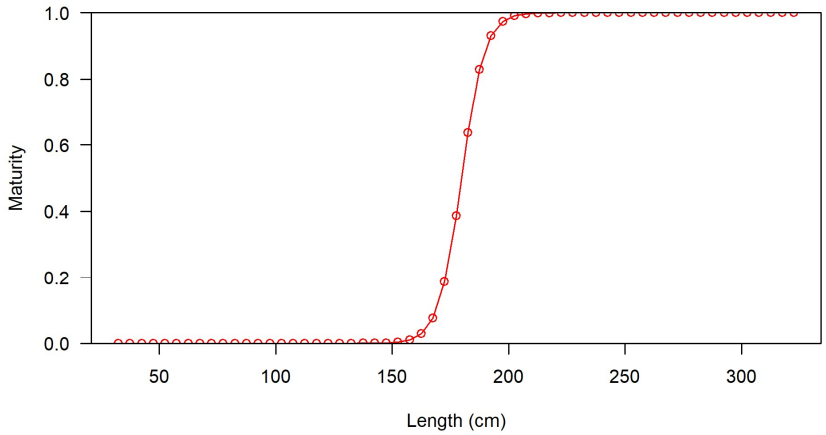


Figure 8. Female proportional maturity as a function of length in the 2026 assessment based on [citation].

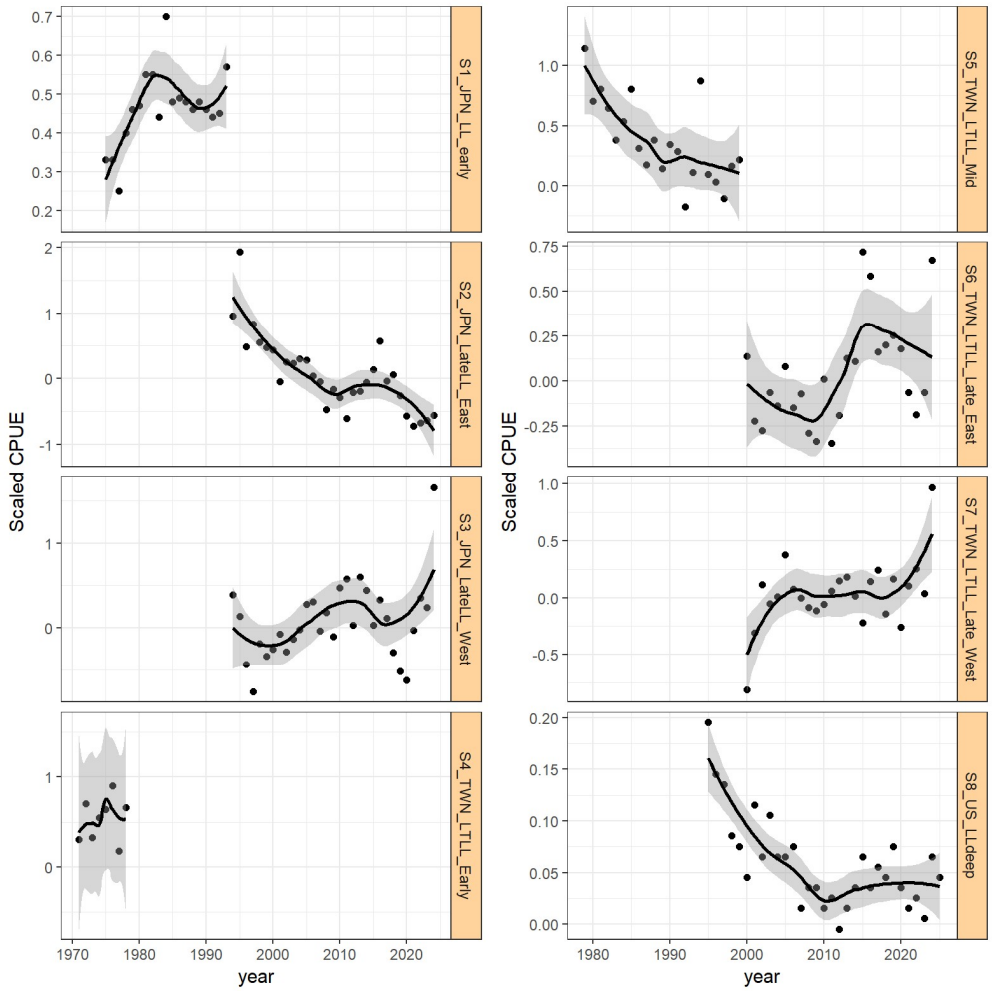


Figure 9. Time series of CPUE indices; continuous black line is a loess smoother showing the average trend by area (i.e. fitted to year for each area with series as a factor).

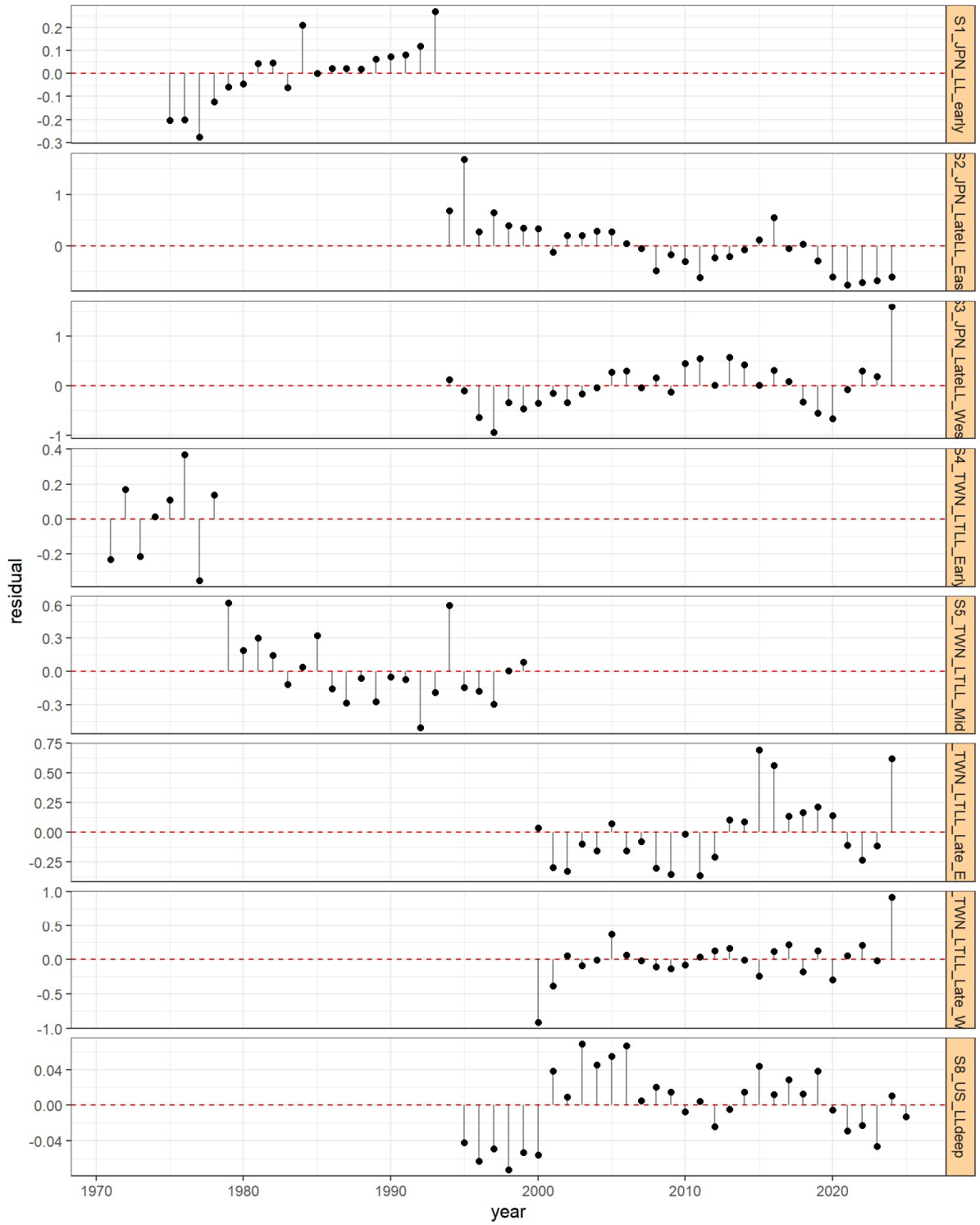


Figure 10. Time series of residuals from the Loess fit.

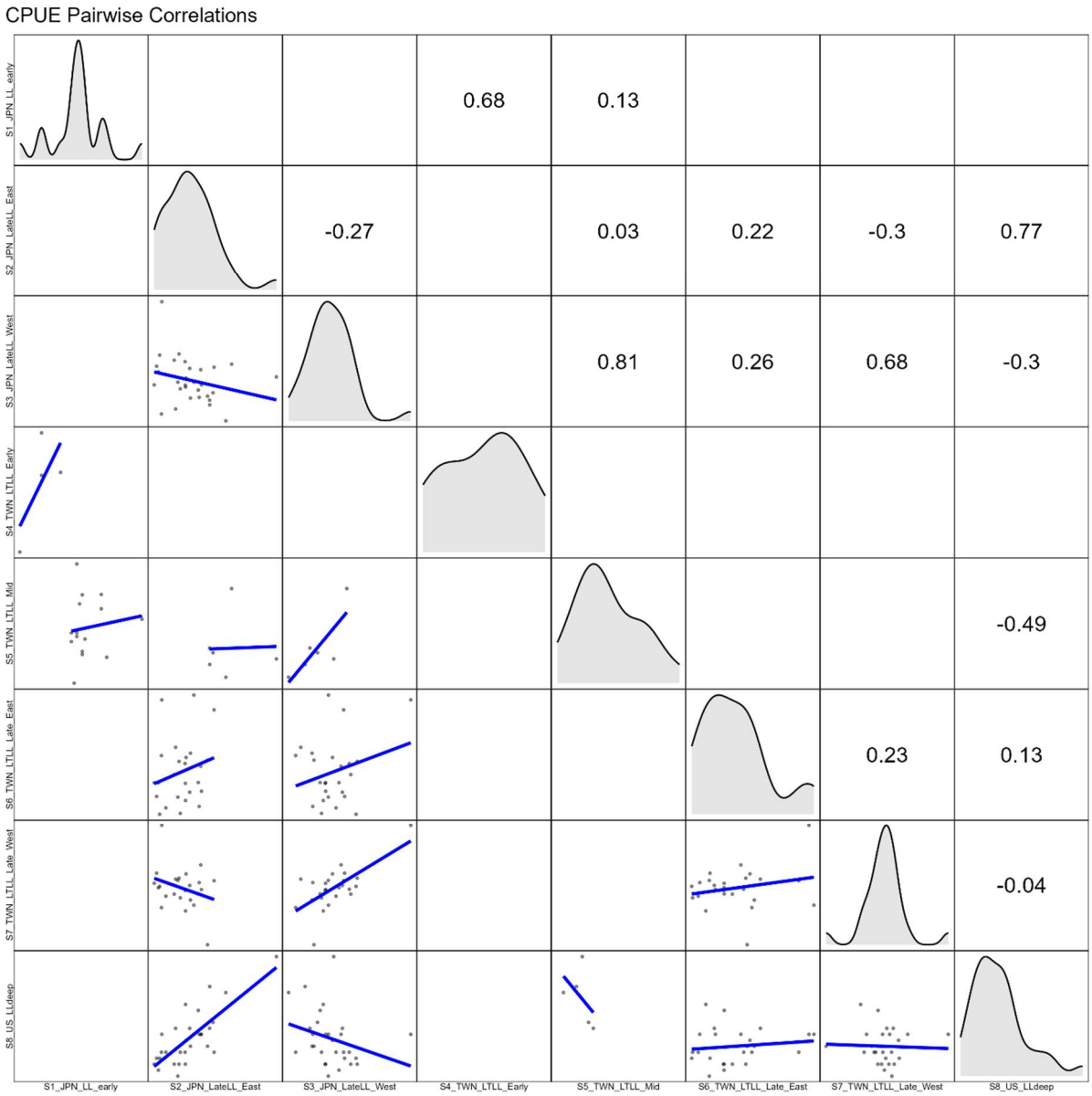


Figure 11. Pairwise scatterplots with blue regression lines (lower left), correlation coefficients (top right), and the range of observations to illustrate correlations among all CPUE indices (central diagonal).

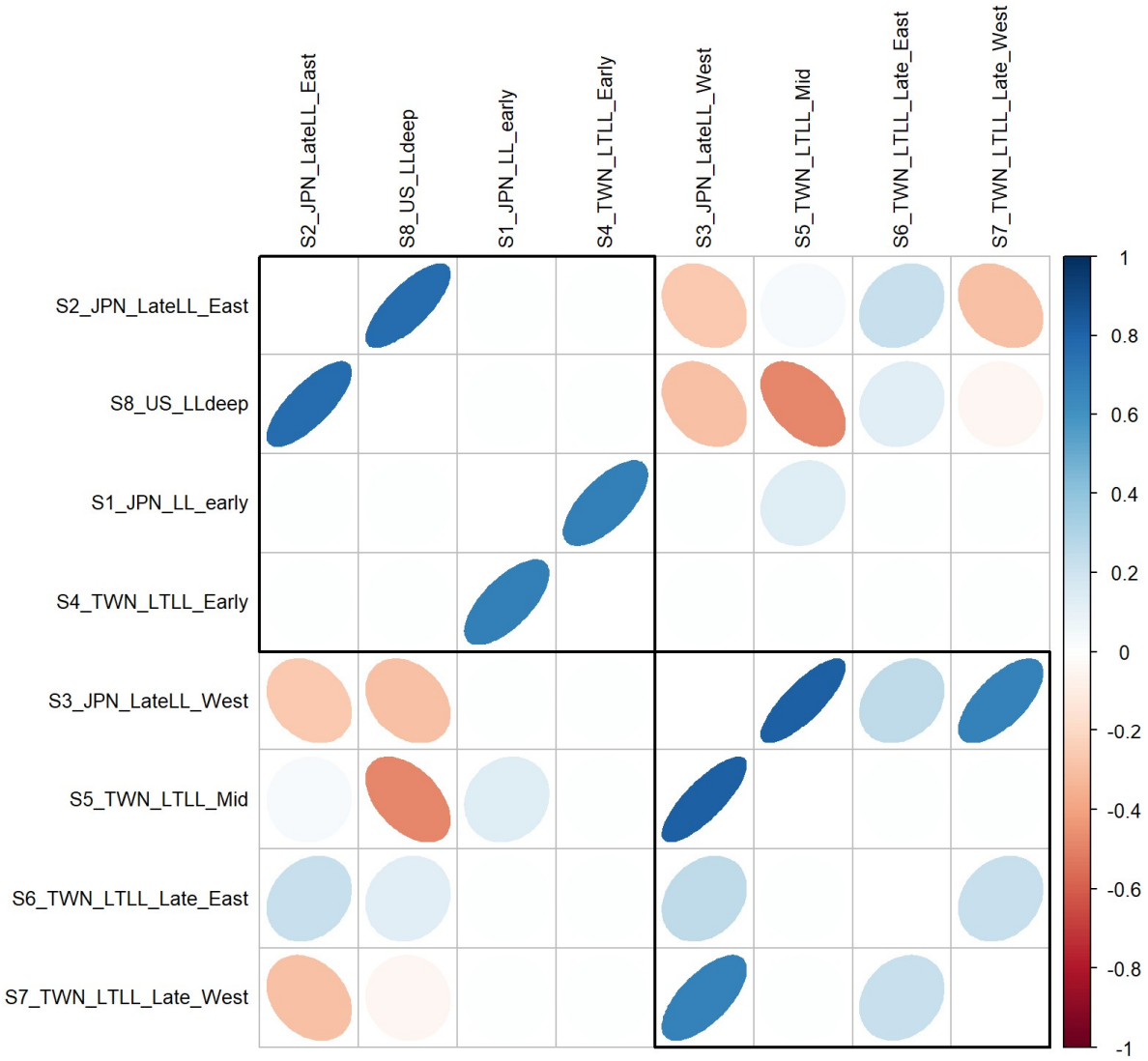


Figure 12. Plot of the correlation matrix for CPUE indices. Blue indicates a positive correlation and red negative. The order of the indices and the rectangular boxes are chosen based on a hierarchical cluster analysis using a set of dissimilarities for the indices being clustered.

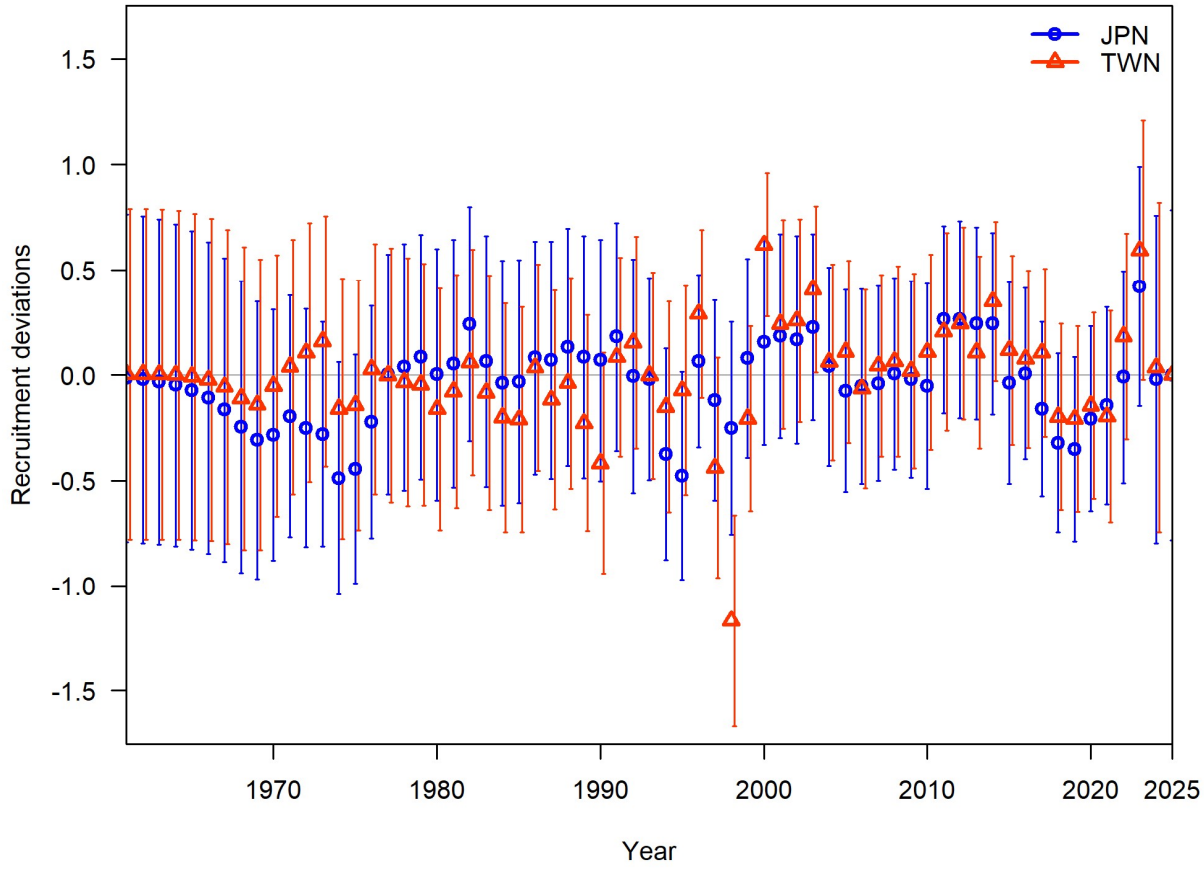
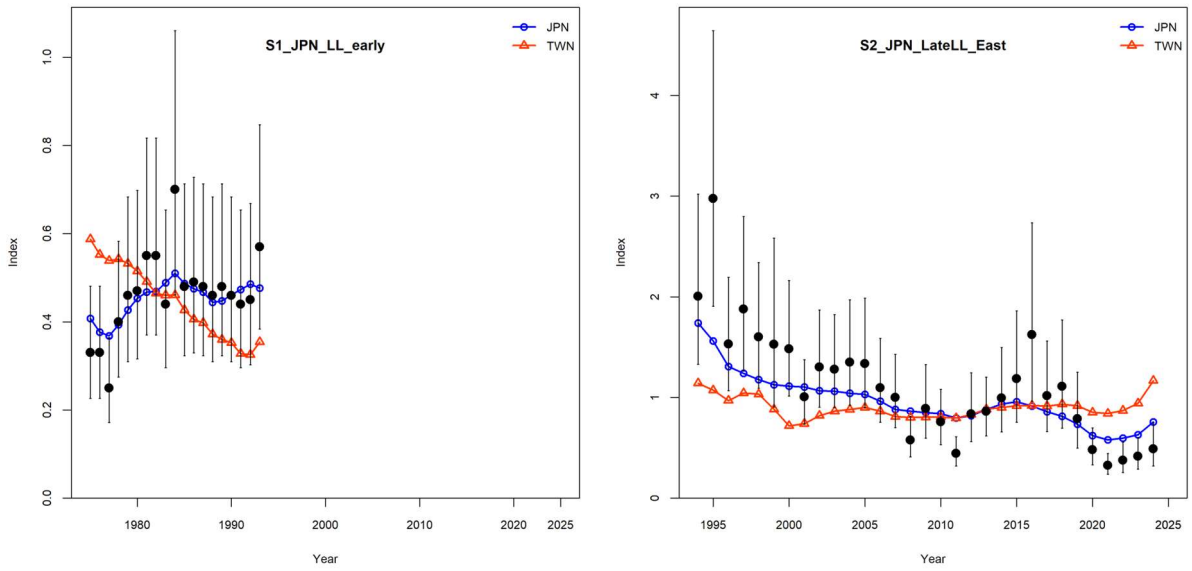


Figure 13 Recruitment deviations of the Japan (blue) and Taiwan (red) preliminary base-case models, with 95% confidence intervals.

Figure 14. Recruitment deviations of the Japan (blue) and Taiwan (red) preliminary base-case models.



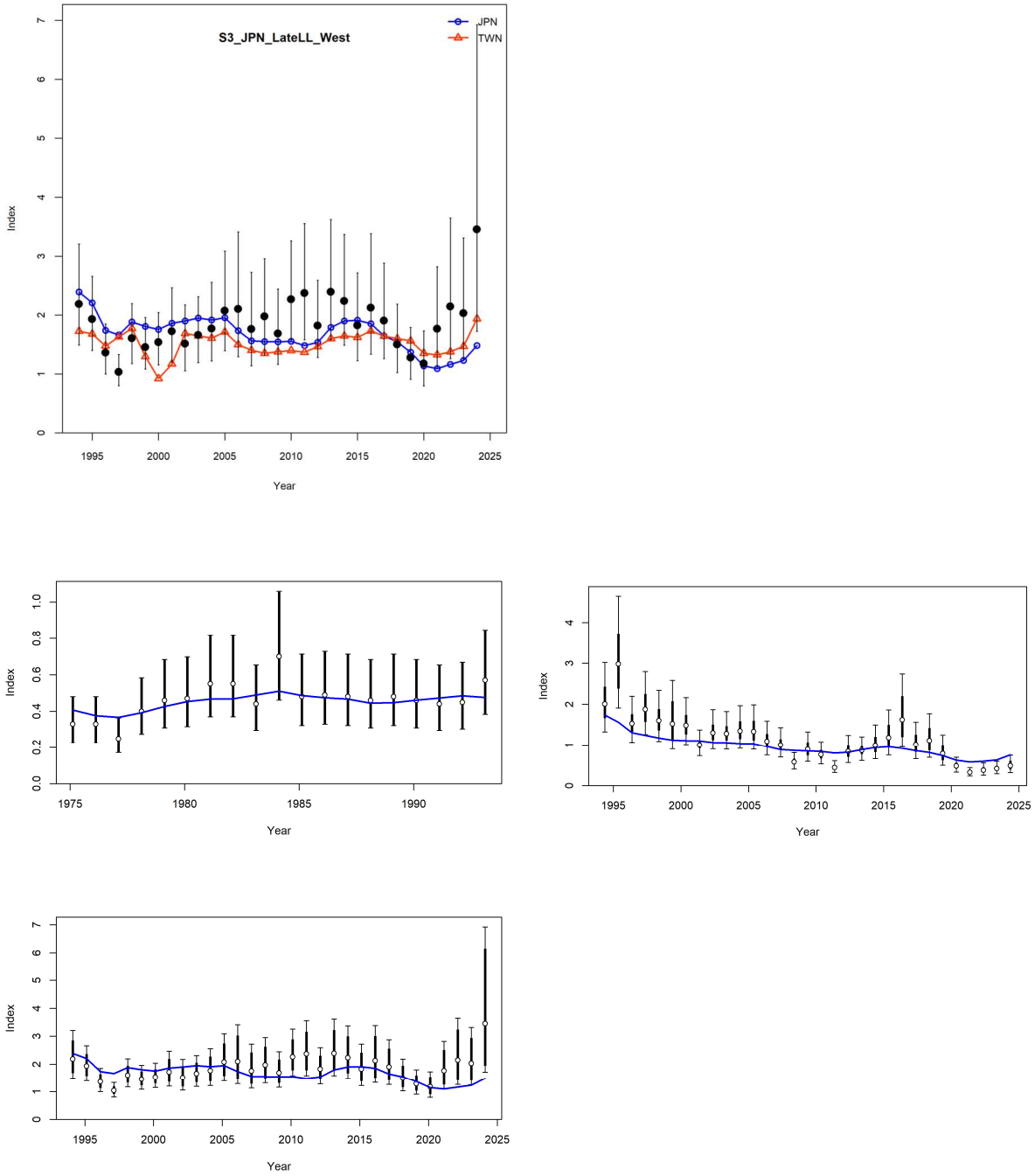


Figure 15. Japan CPUE indices (points) and fits (blue line) of the Japan model. Error bars are 95% confidence intervals of the adjusted variance of the index.

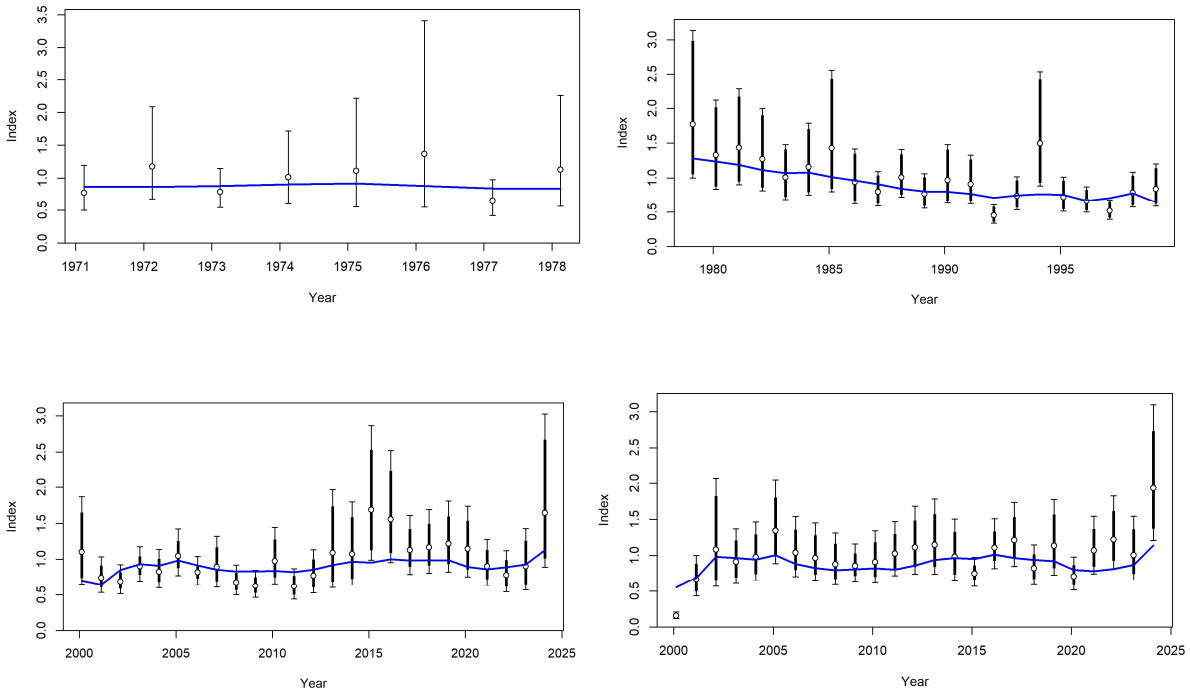


Figure 16. Taiwan CPUE indices (points) and fits of the Taiwan (blue line) models. Error bars are 95% confidence intervals of the adjusted variance of the index.

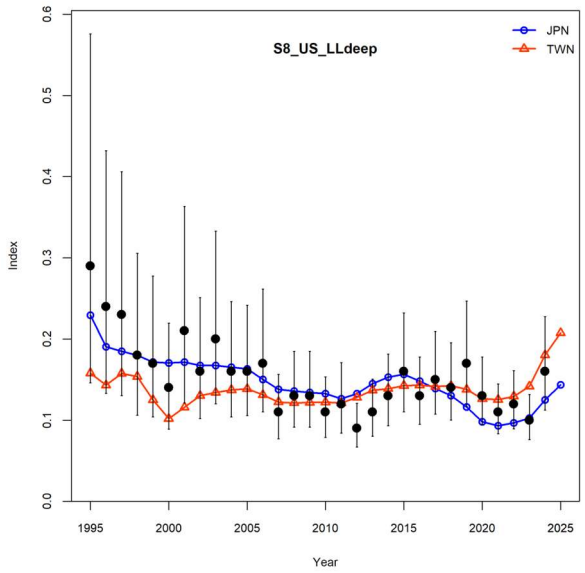


Figure 17. U.S. CPUE indices (points) and fits of the Japan (blue) and Taiwan (red) models. Error bars are 95% confidence intervals of the adjusted variance of the index.

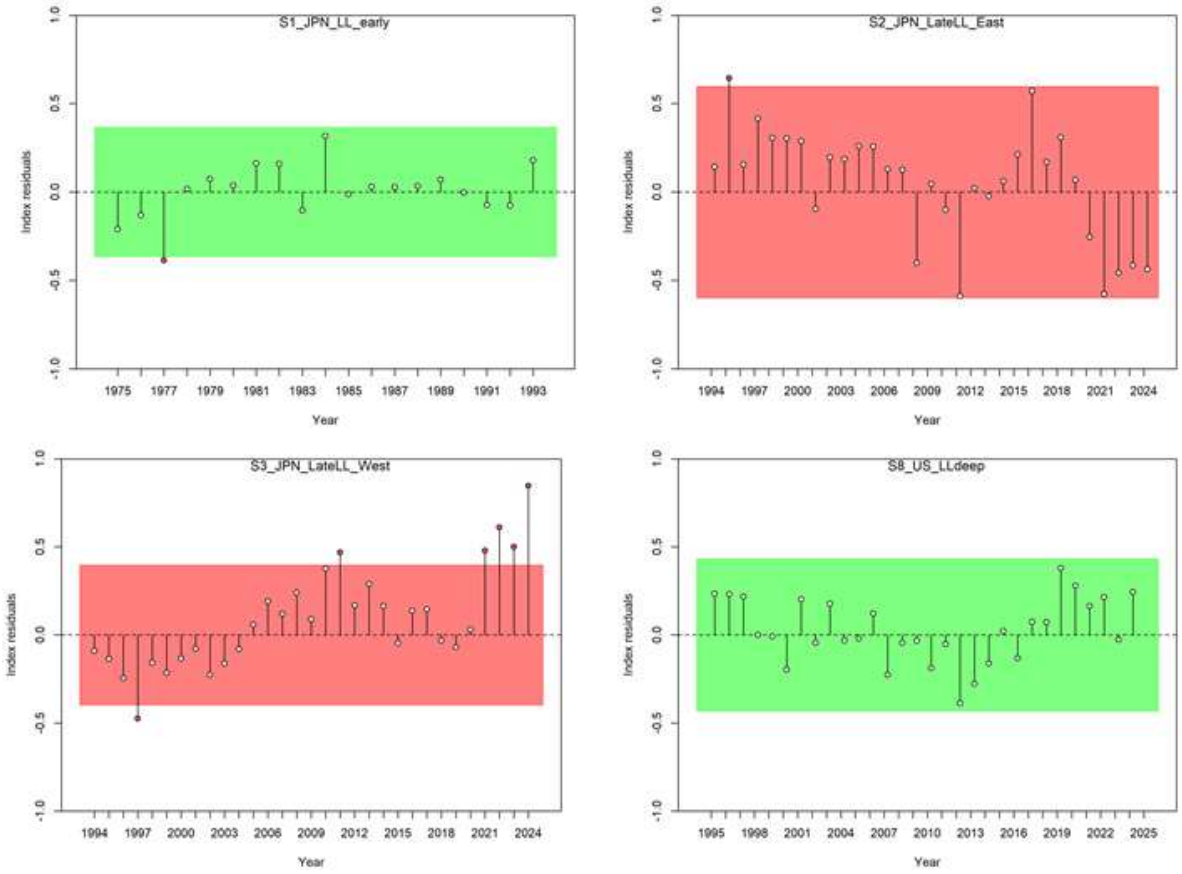


Figure 18. Runs test for the CPUE indices included in the Japan model. Red indicates the index fit fails the runs test, green indicates it passes.

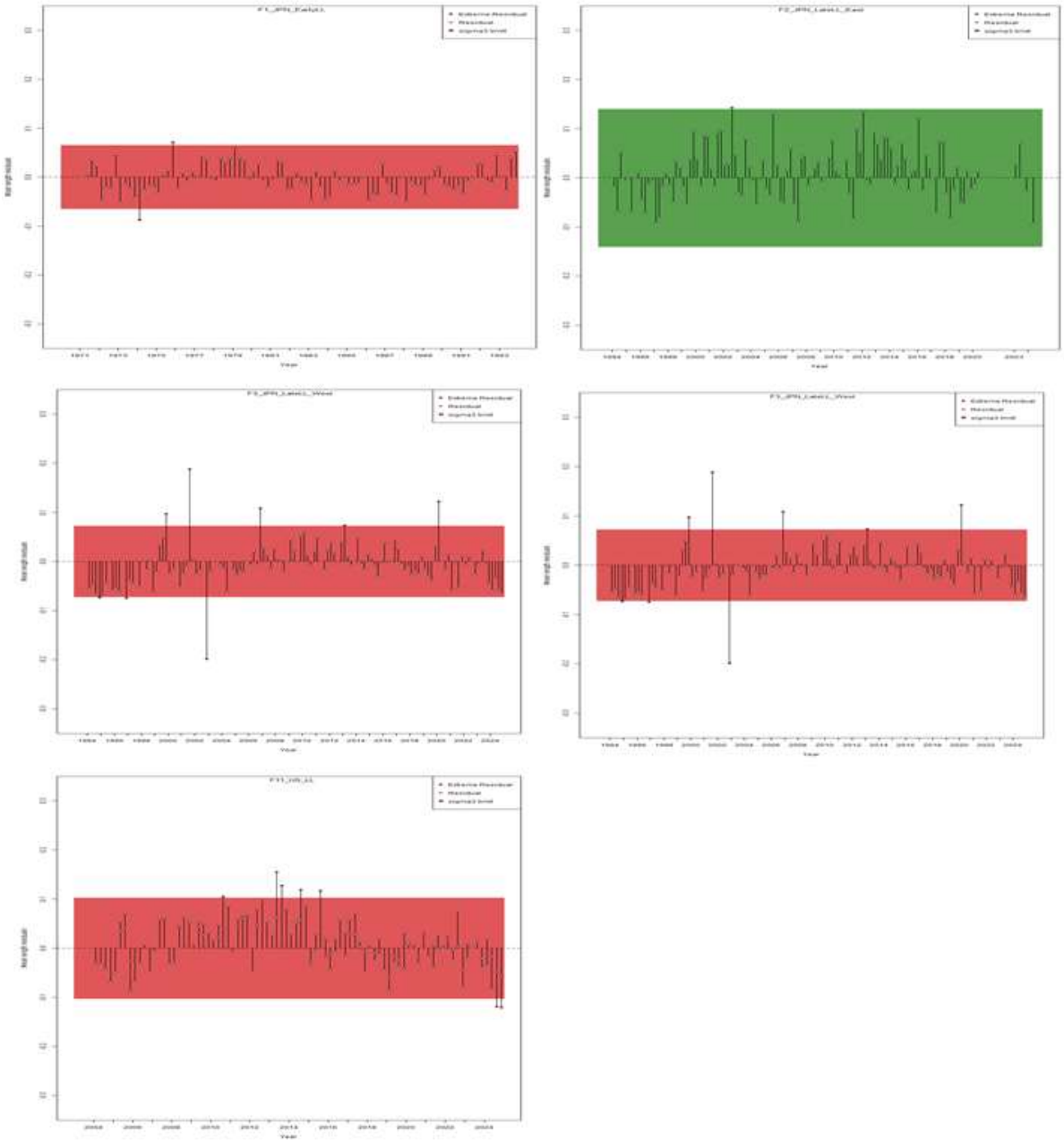


Figure 19. Runs test for the length composition data for each catch fleet included in the Japan model. Red indicates the index fit fails the runs test, green indicates it passes.

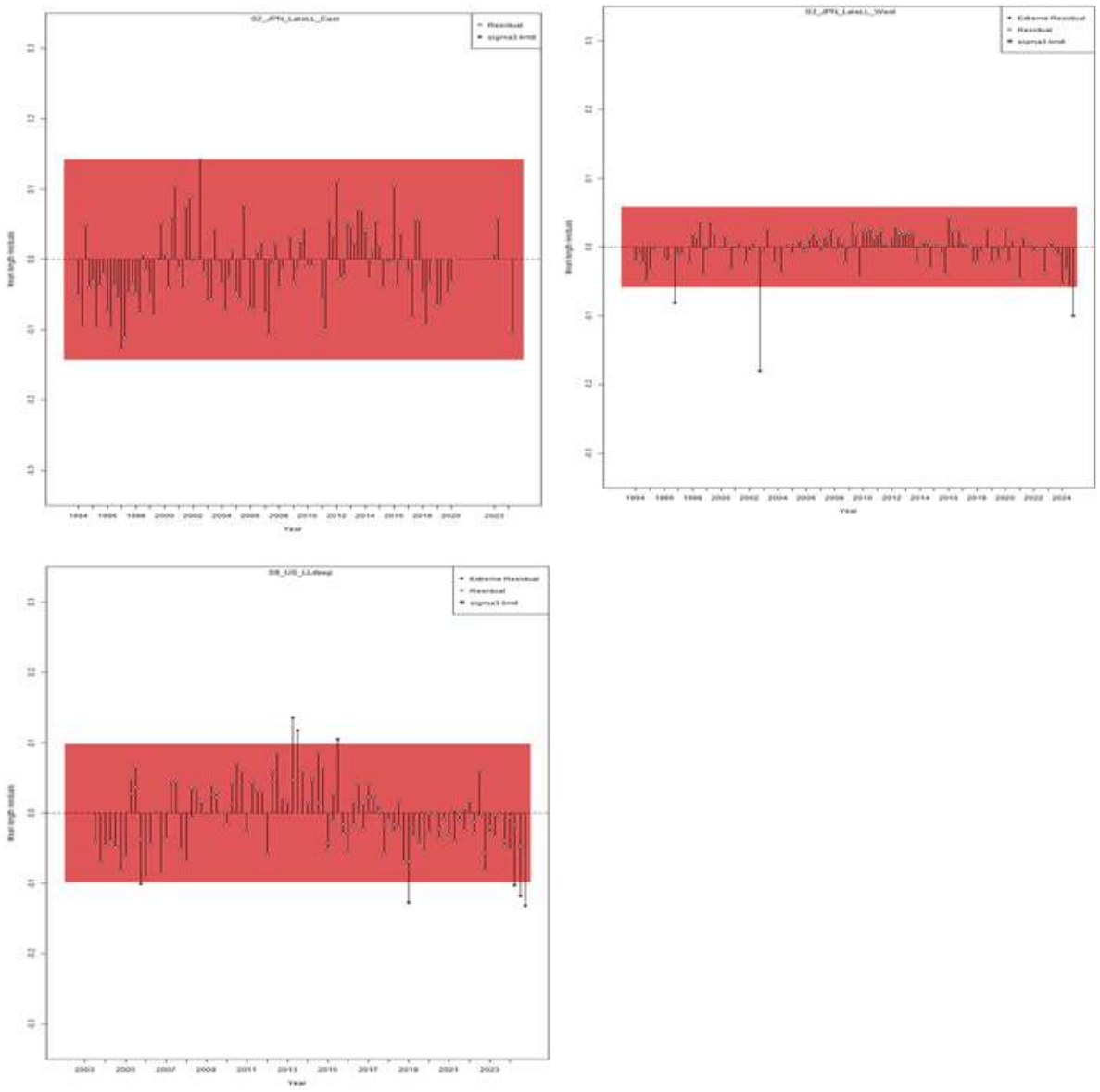


Figure 20. Runs test for the length composition data for each CPUE fleet included in the Japan model. Red indicates the index fit fails the runs test, green indicates it passes.

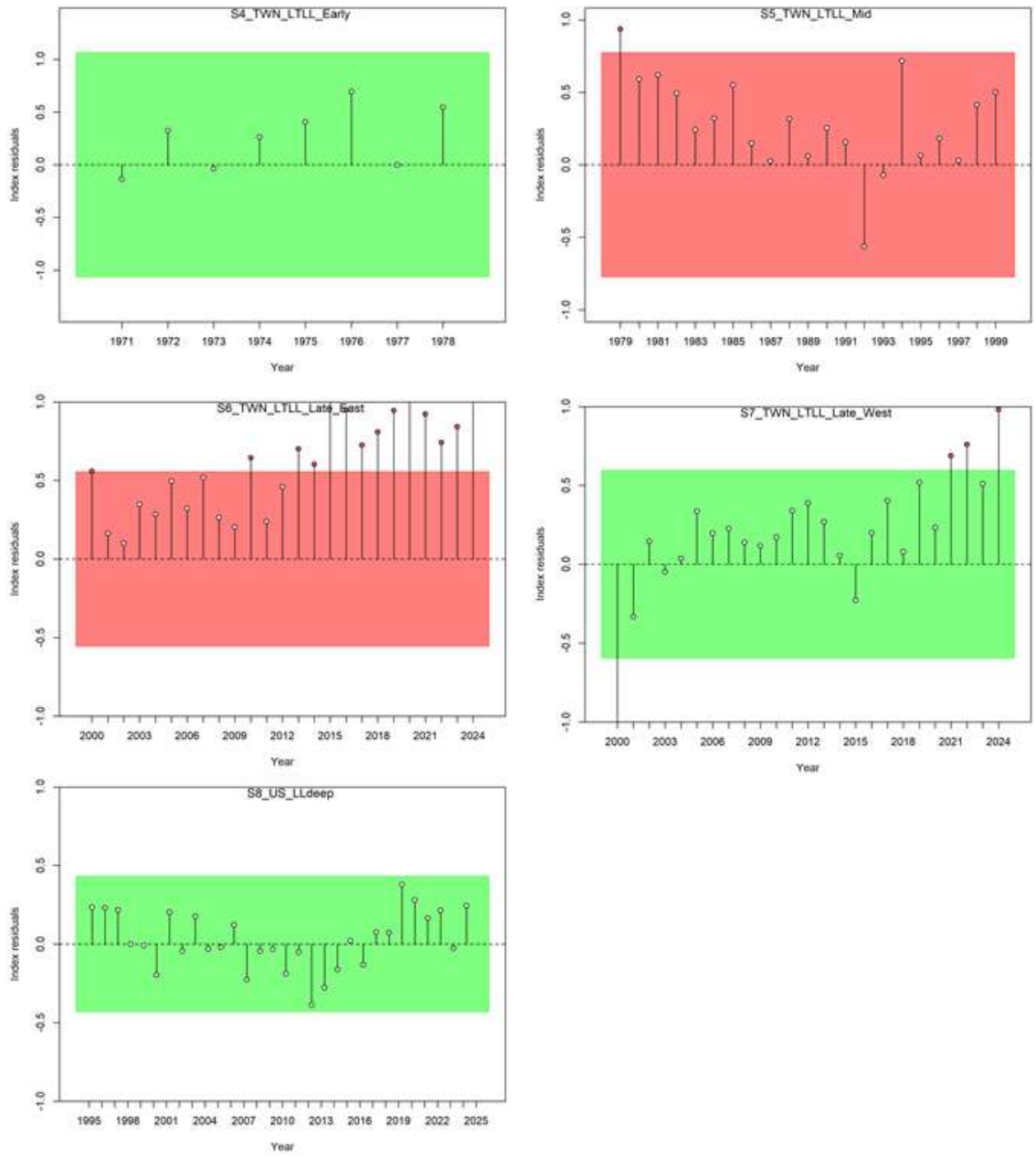


Figure 21. Runs test for the CPUE indices included in the Taiwan model. Red indicates the index fit fails the runs test, green indicates it passes.

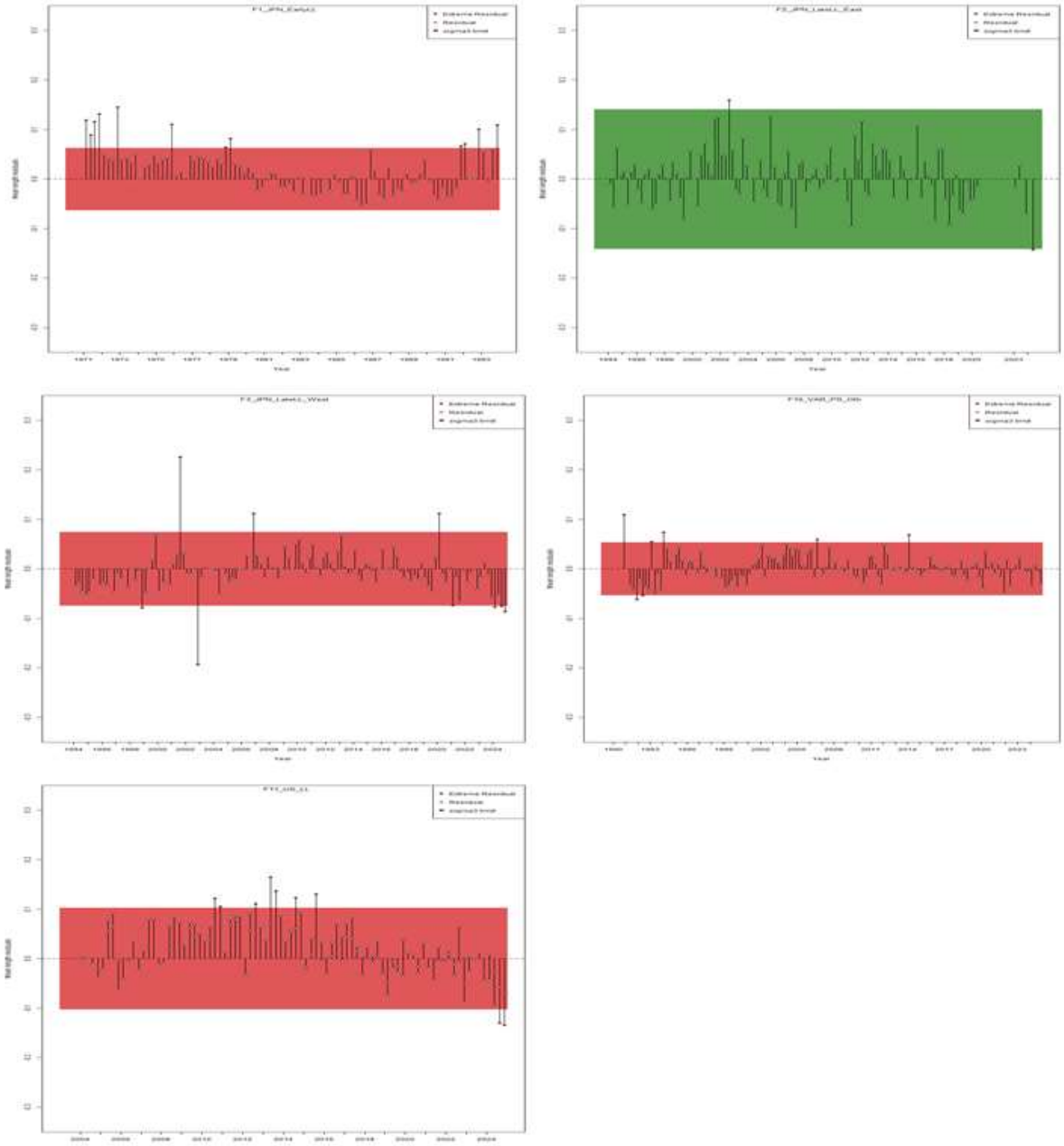


Figure 22. Runs test for the length composition data for each catch fleet included in the Taiwan model. Red indicates the index fit fails the runs test, green indicates it passes.

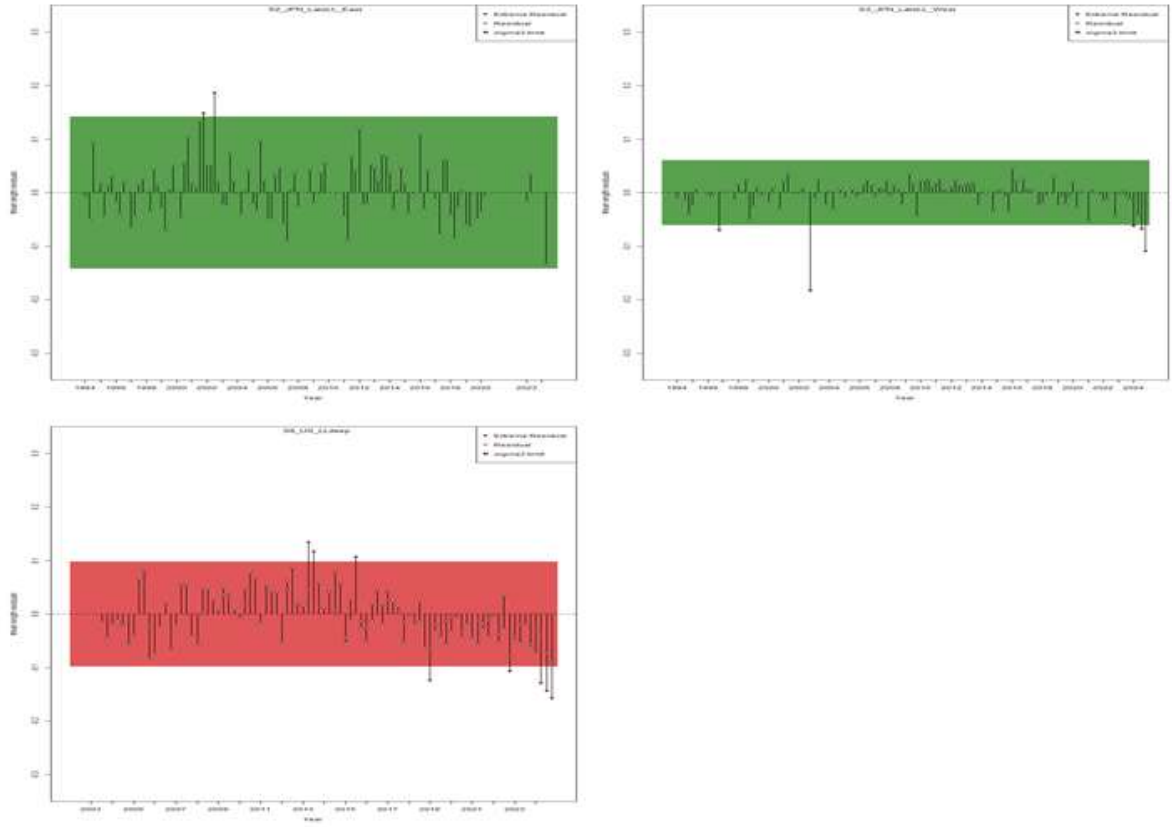


Figure 23. Runs test for the length composition data for each CPUE fleet included in the Taiwan model. Red indicates the index fit fails the runs test, green indicates it passes.

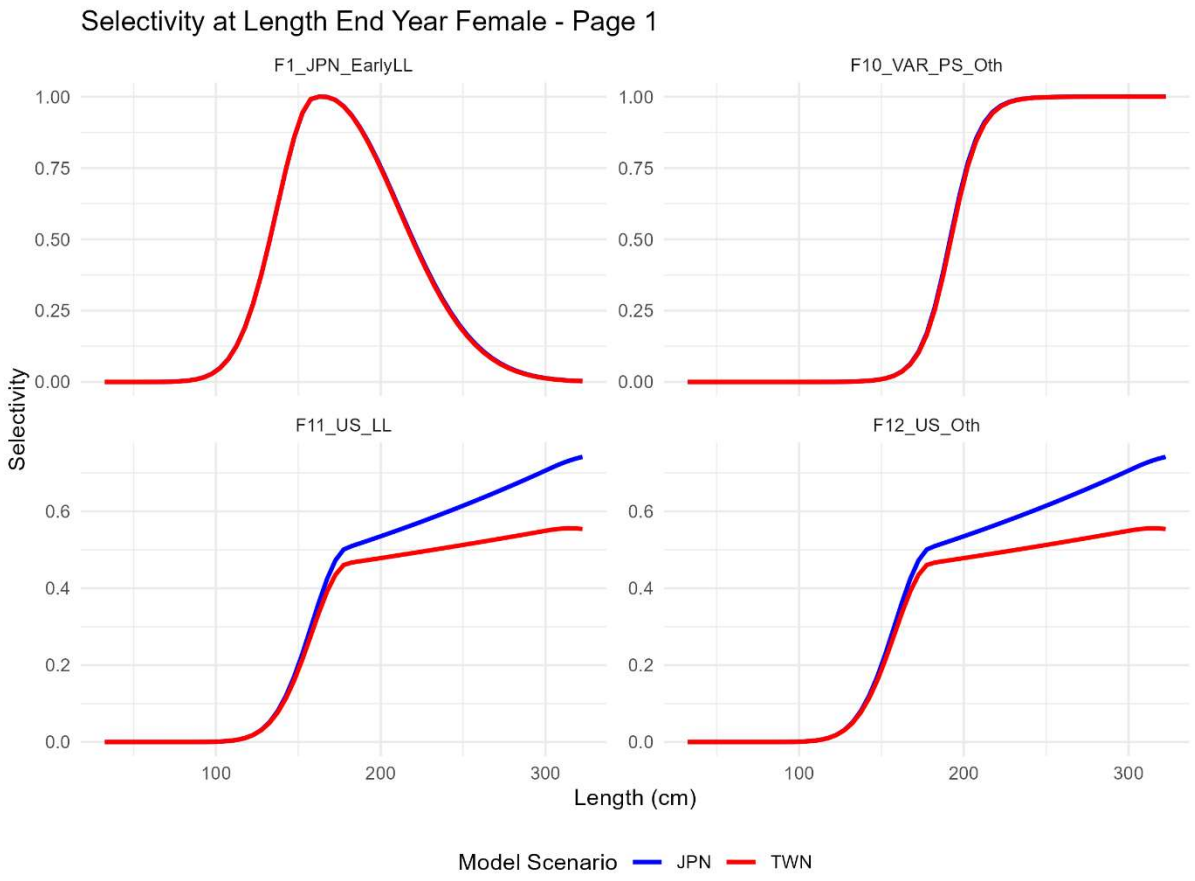


Figure 24. Selectivity at length of the Japan (blue) and Taiwan (red) models.

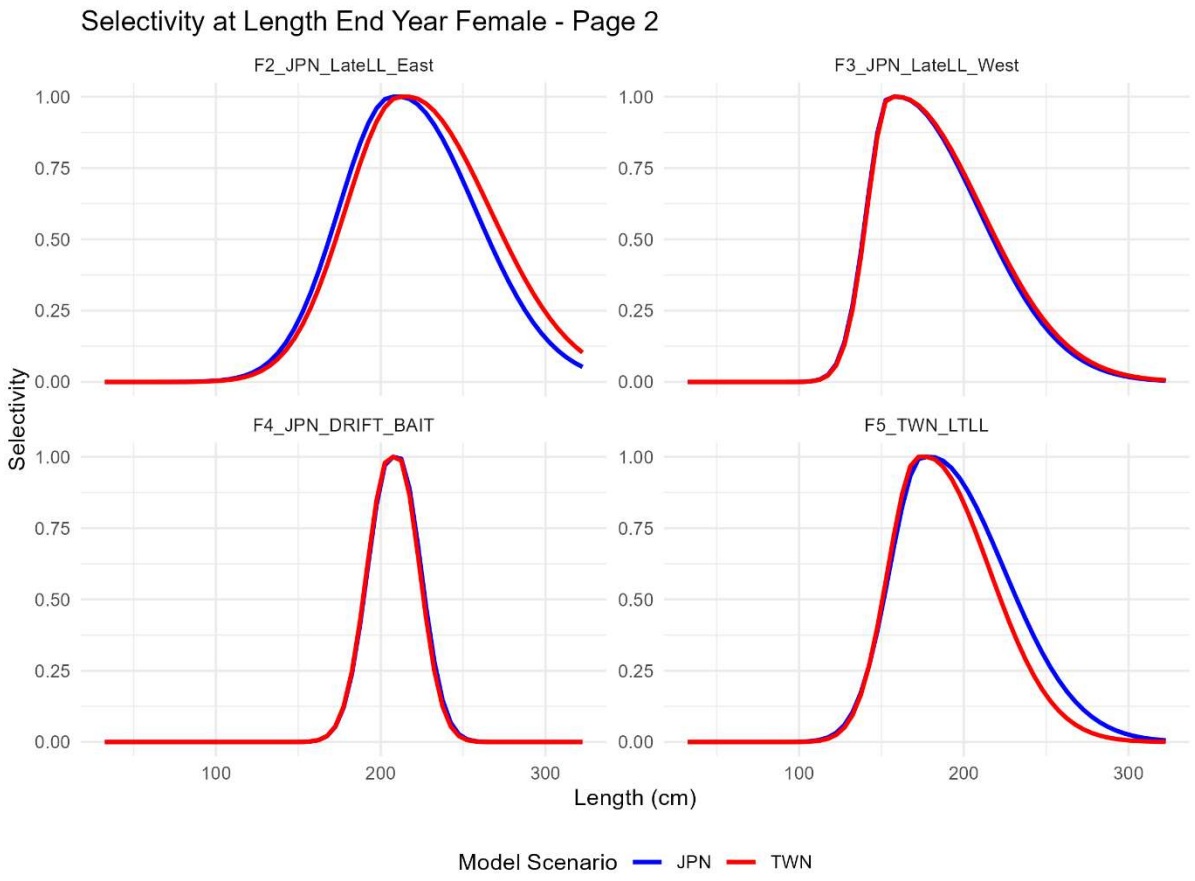


Figure 25. Selectivity at length of the Japan (blue) and Taiwan (red) models.

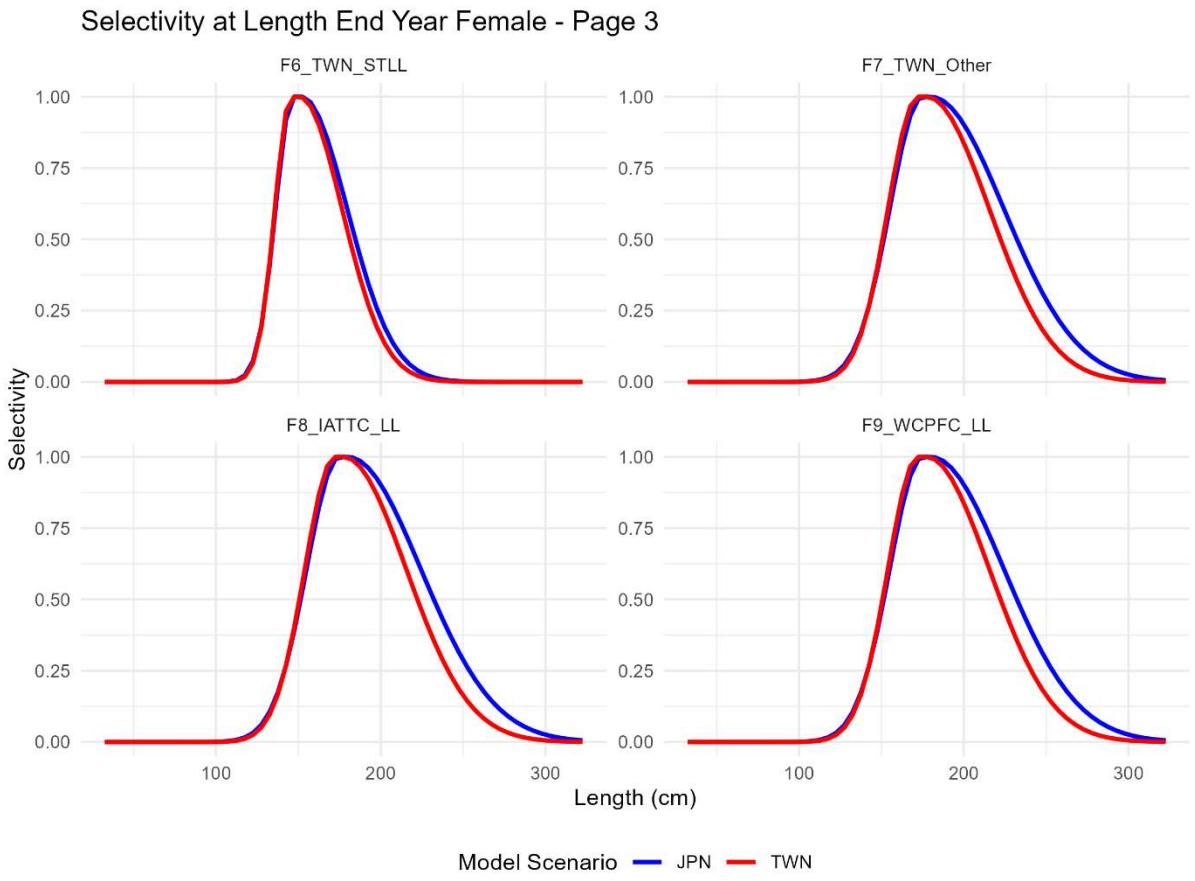


Figure 26. Selectivity at length of the Japan (blue) and Taiwan (red) models.

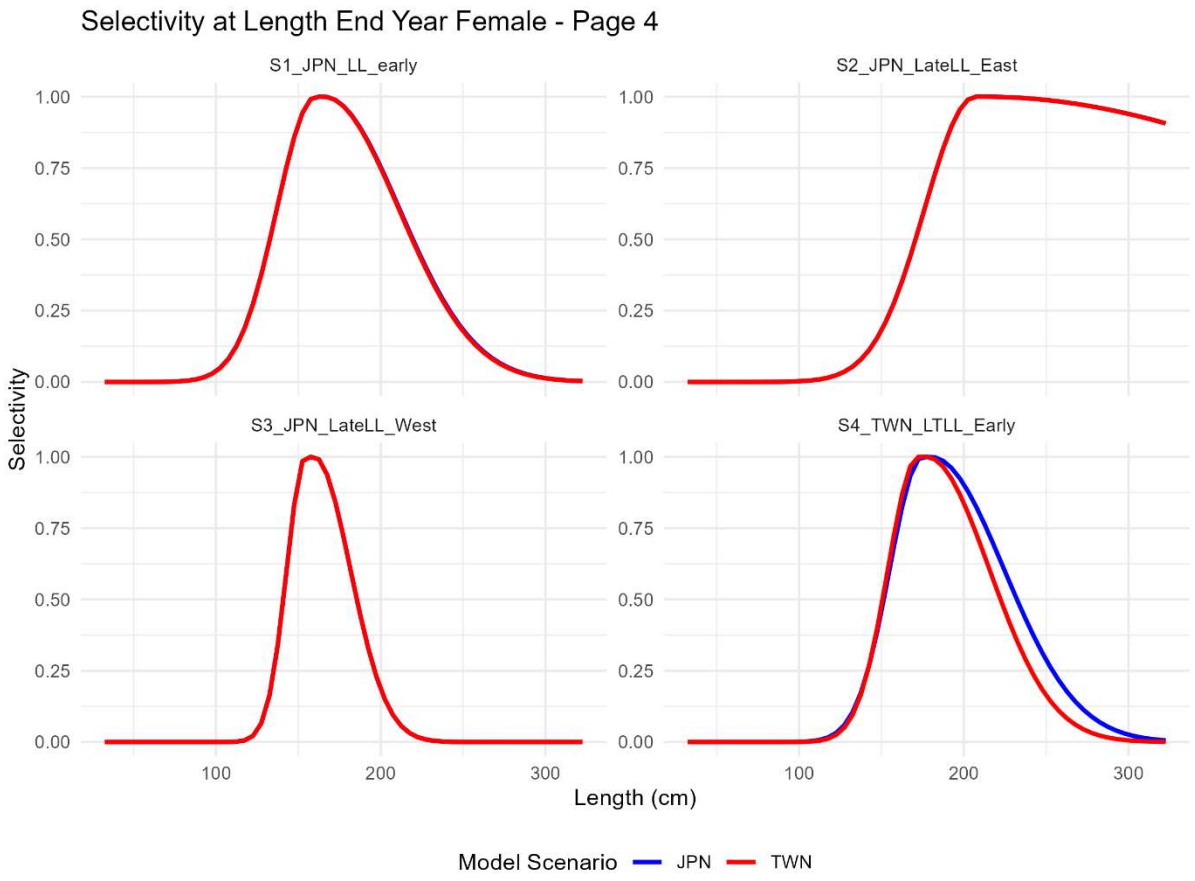


Figure 27. Selectivity at length of the Japan (blue) and Taiwan (red) models.

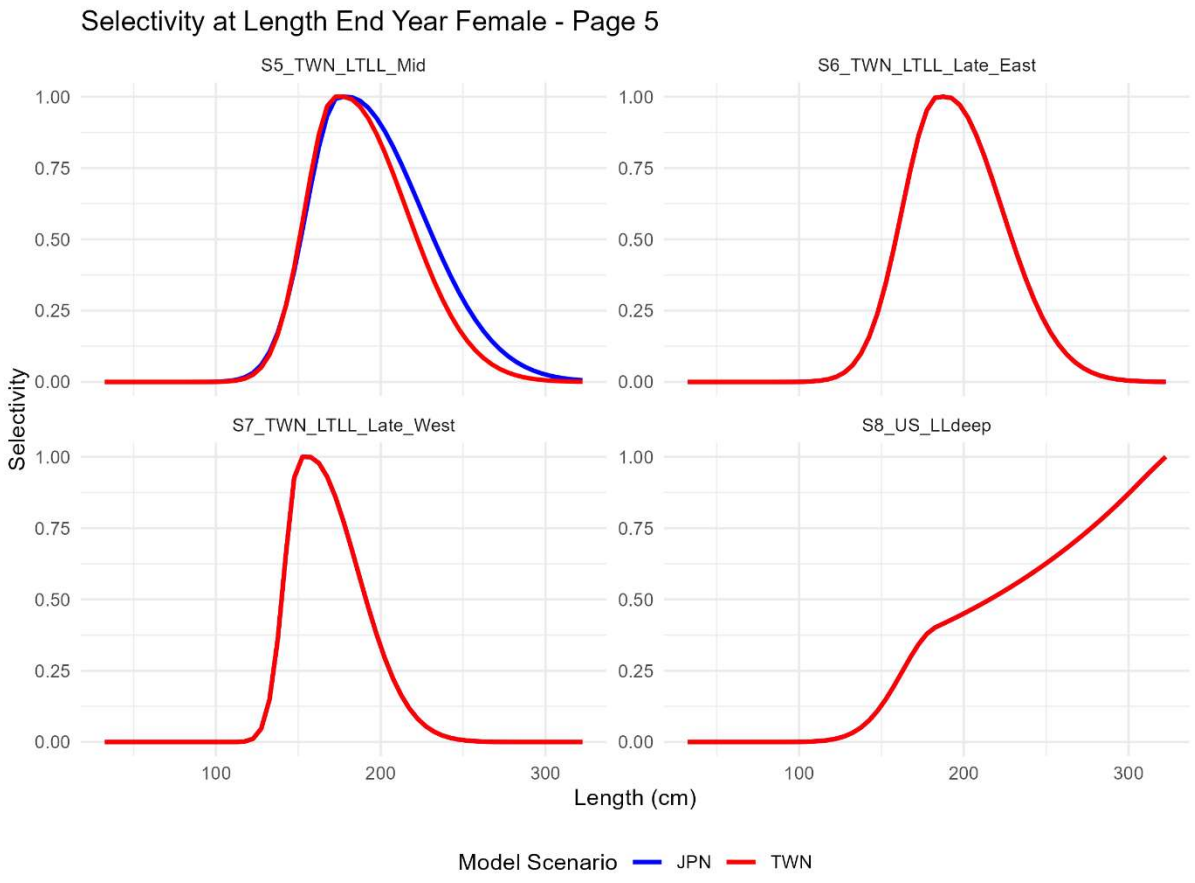


Figure 28. Selectivity at length of the Japan (blue) and Taiwan (red) models.

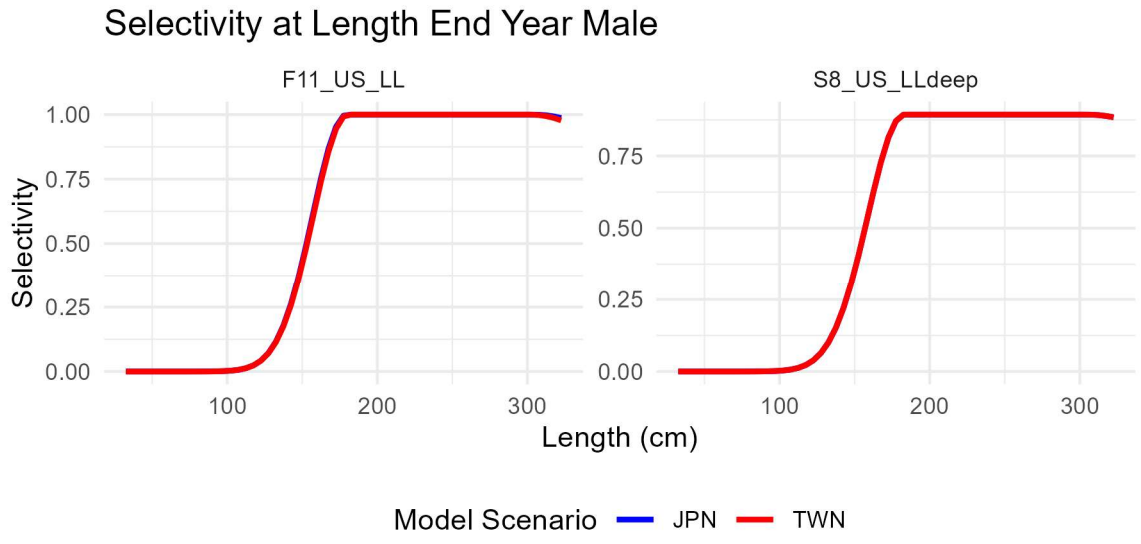


Figure 28a. Male selectivity at length of the Japan (blue) and Taiwan (red) models.

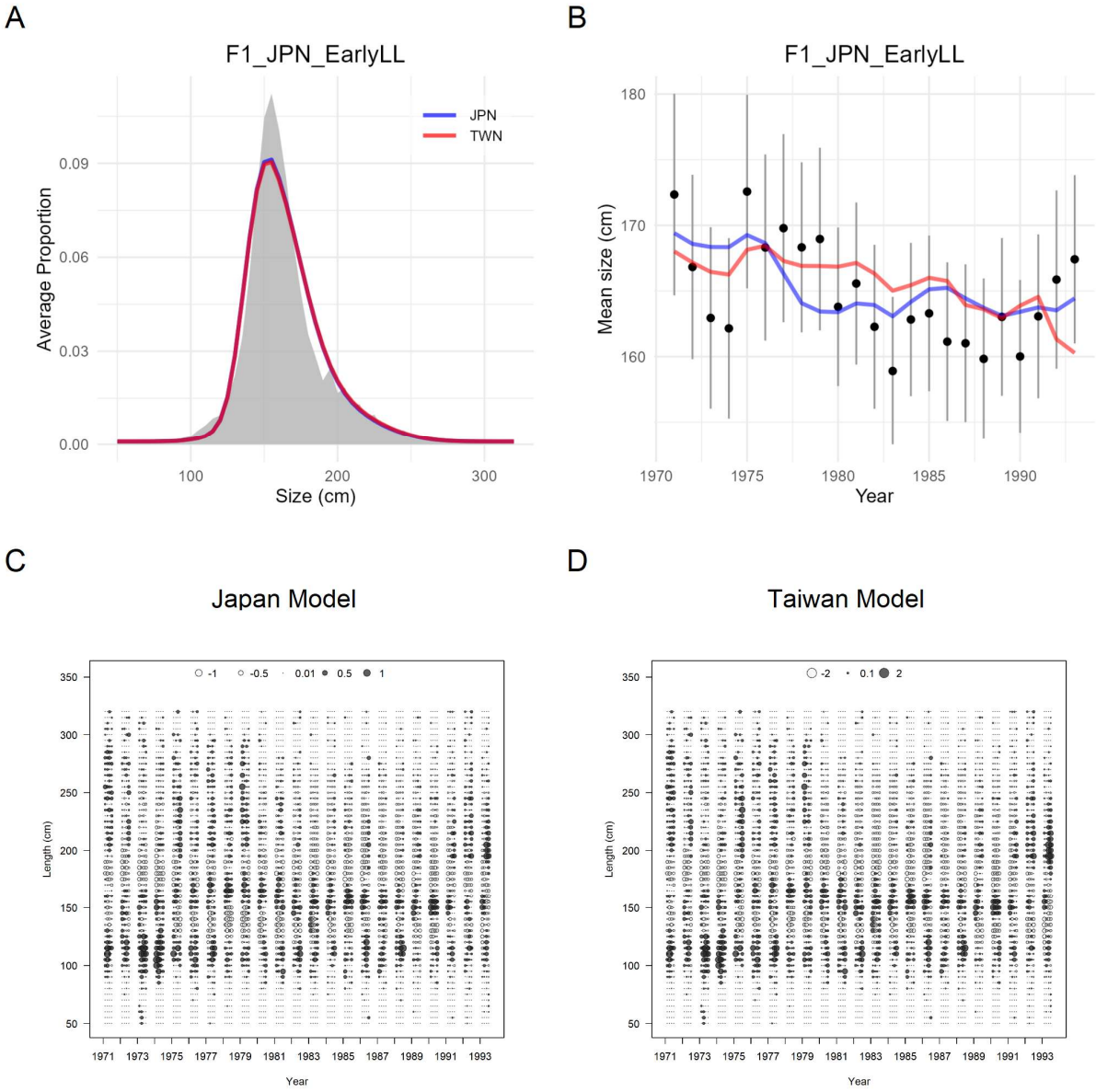


Figure 29. Model fits to F1_JPN_EarlyLL size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis 3, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

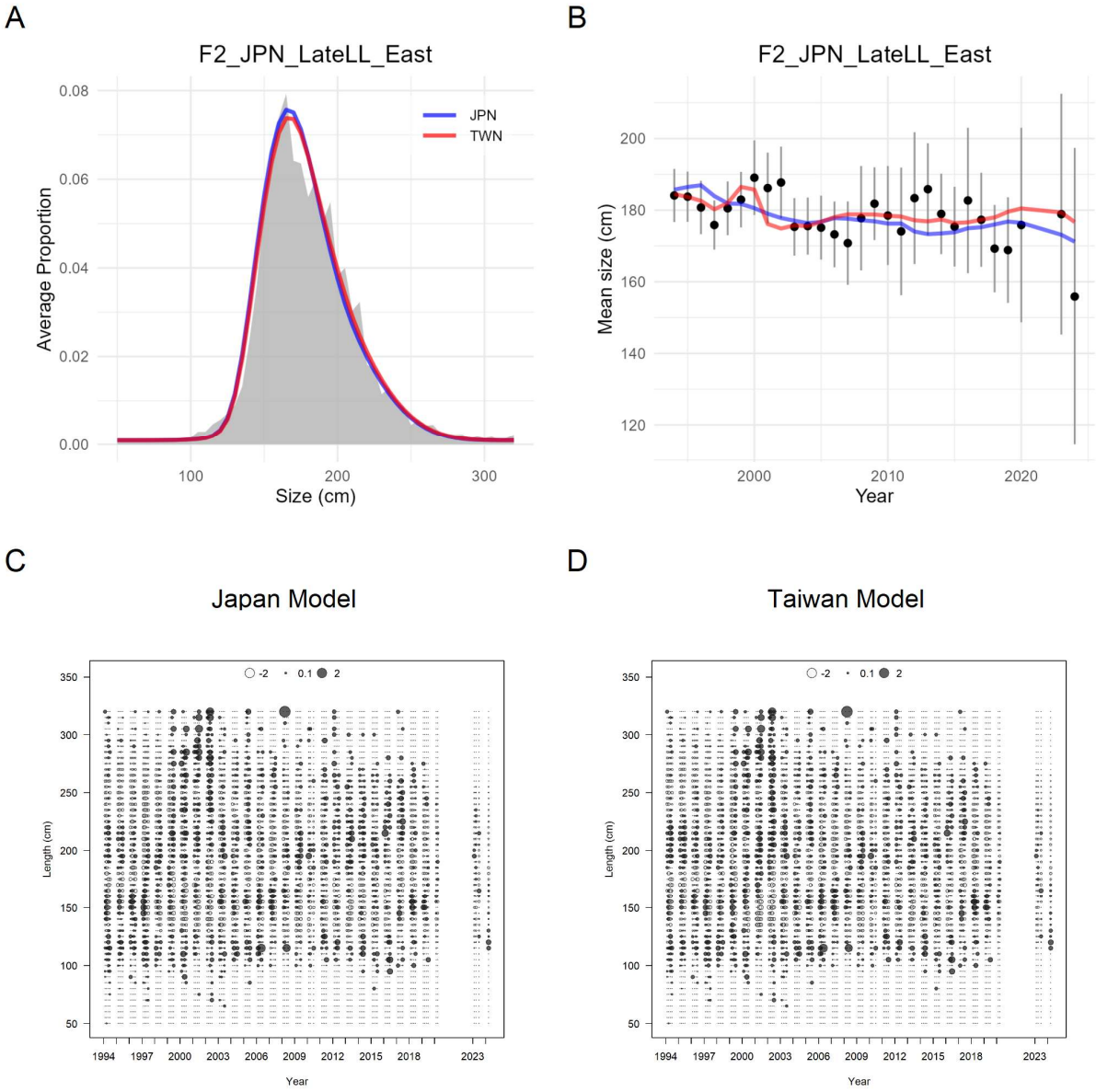


Figure 30. Model fits to F2_JPN_LateLL_East size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis 3, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

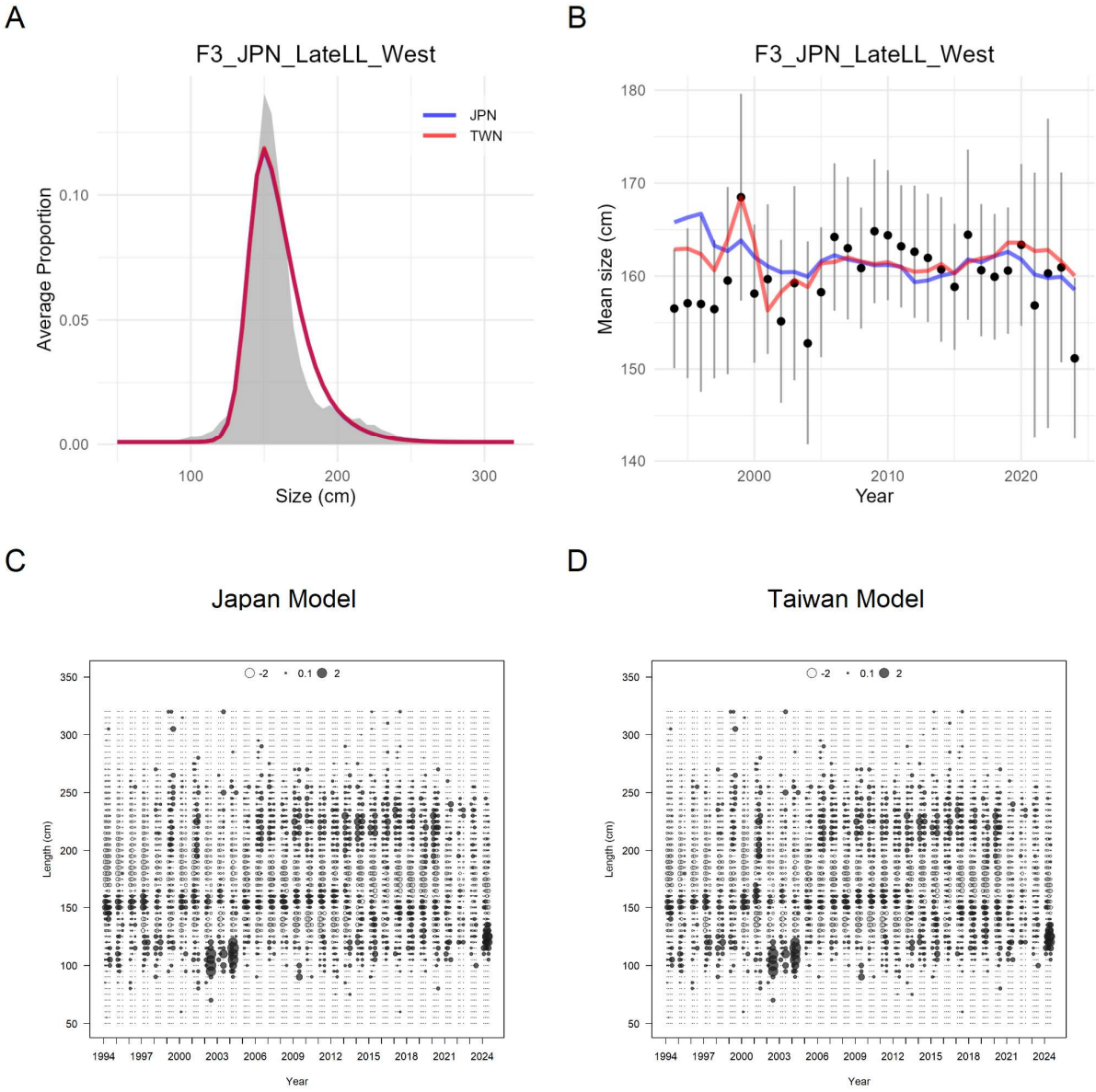


Figure 31. Model fits to F3_JPN_LateLL_West size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

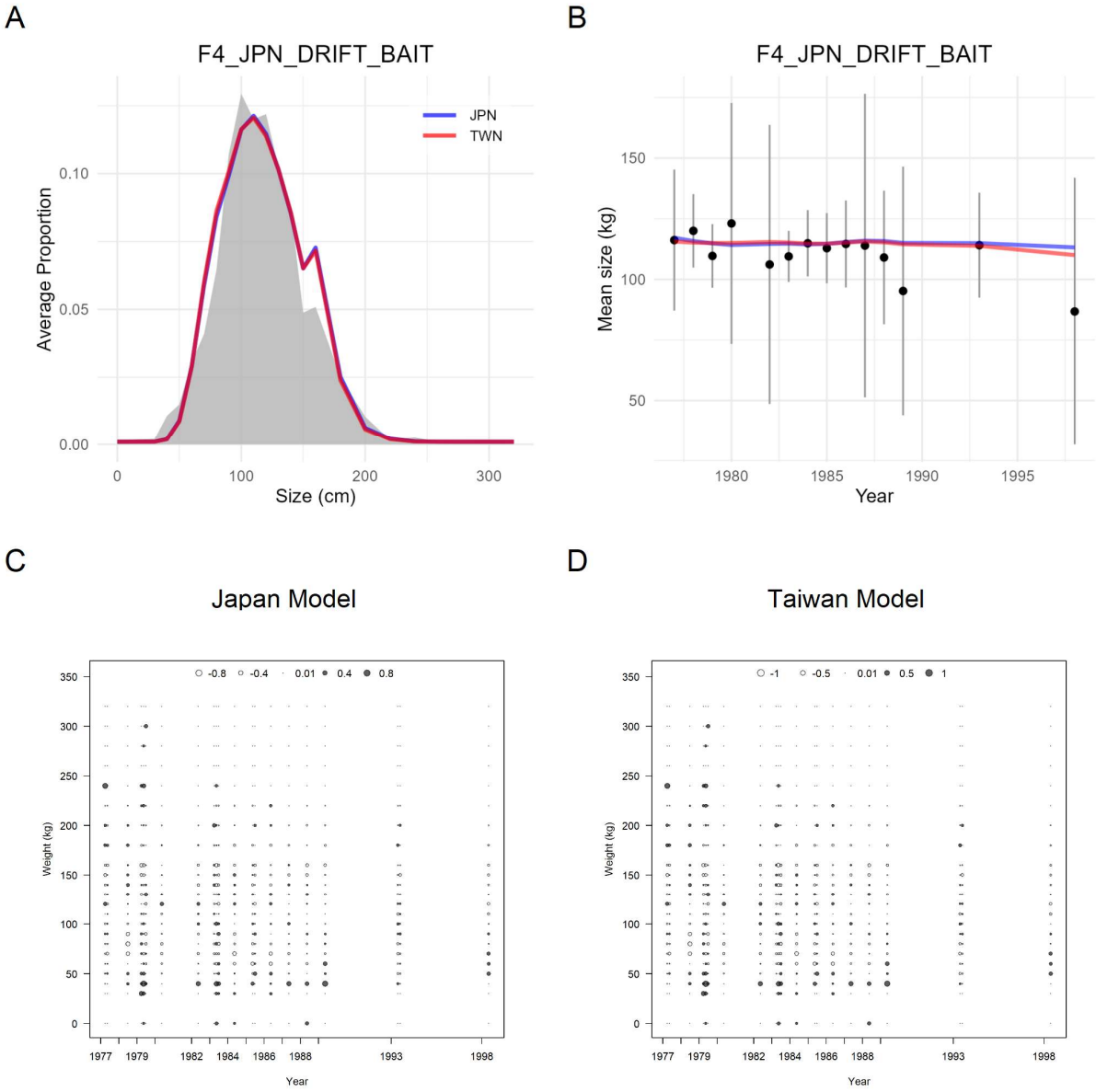


Figure 32. Model fits to F4_JPN_DRIFT_BAIT size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

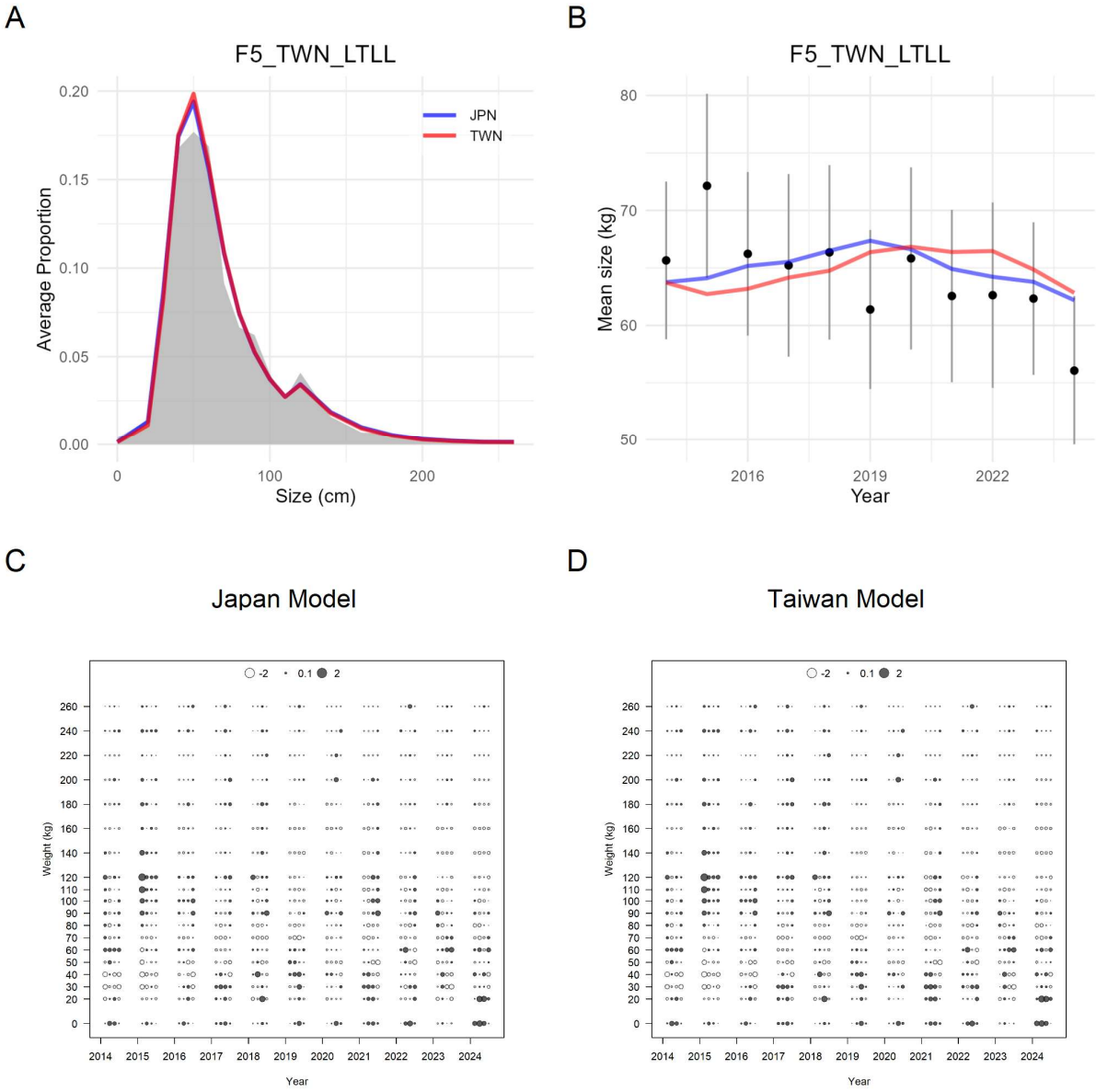


Figure 33. Model fits to F5_TWN_LTLL size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

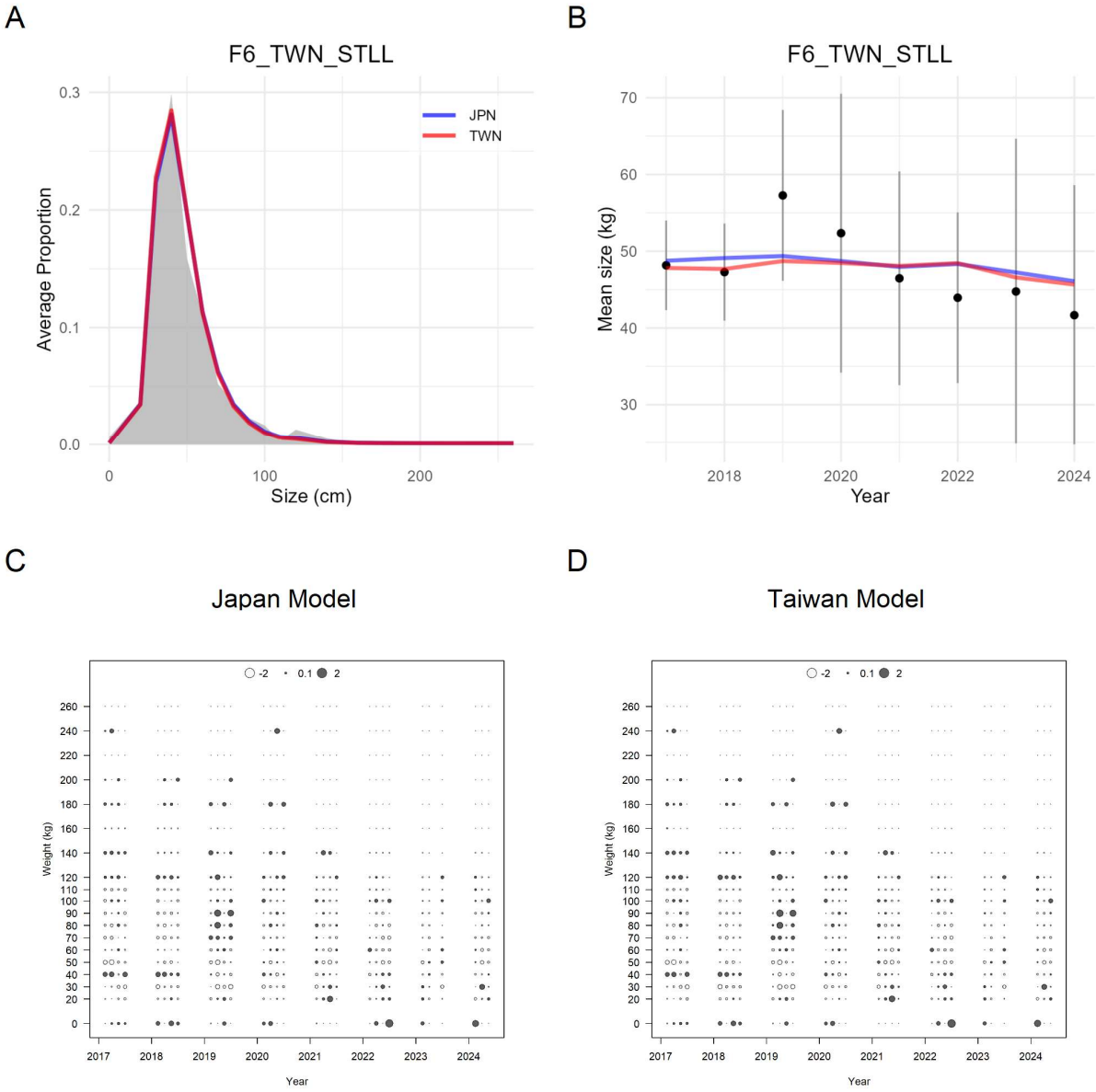


Figure 34. Model fits to F6_TWN_STLL size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

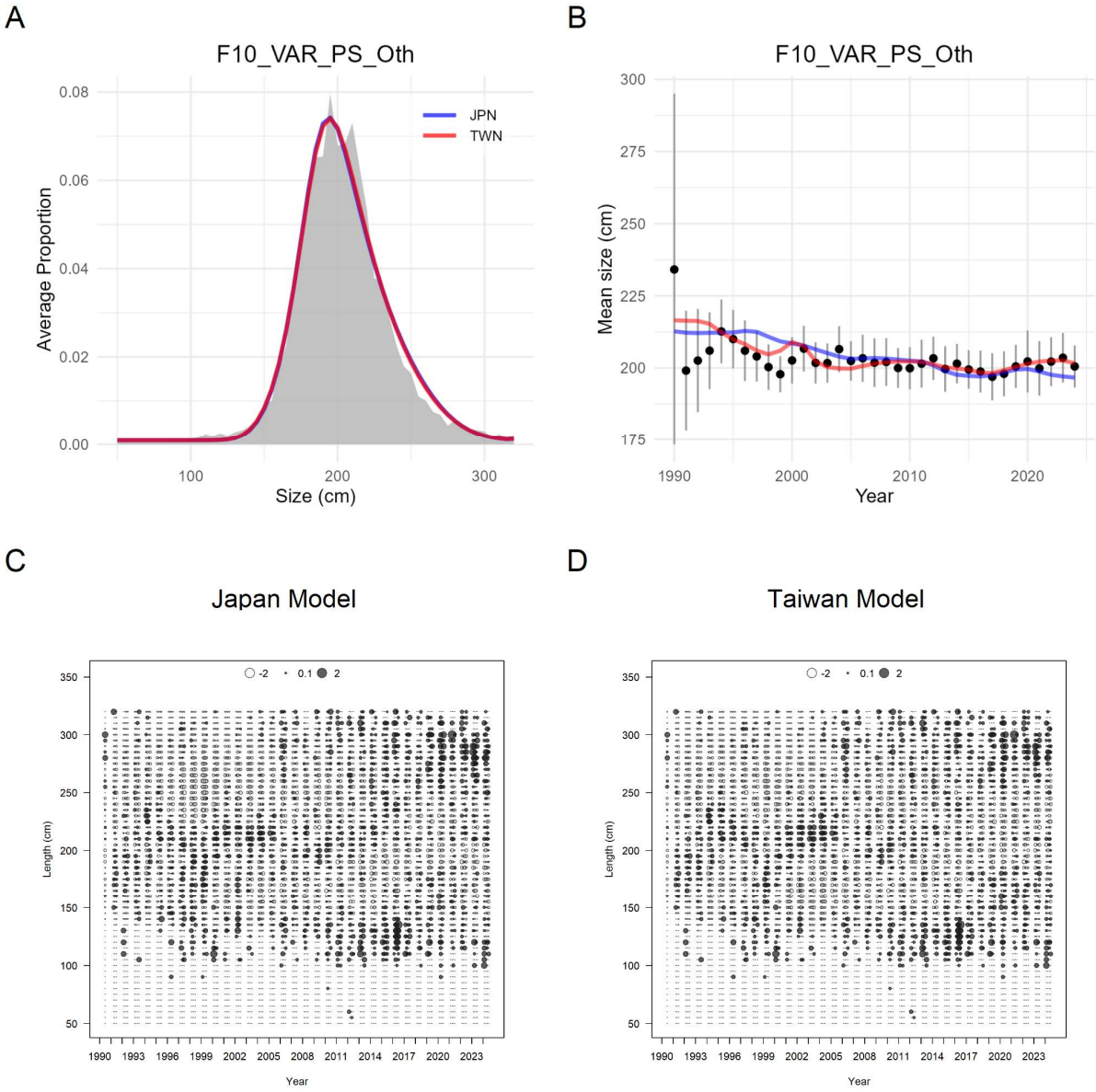


Figure 35. Model fits to F10_VAR_PS_Oth size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

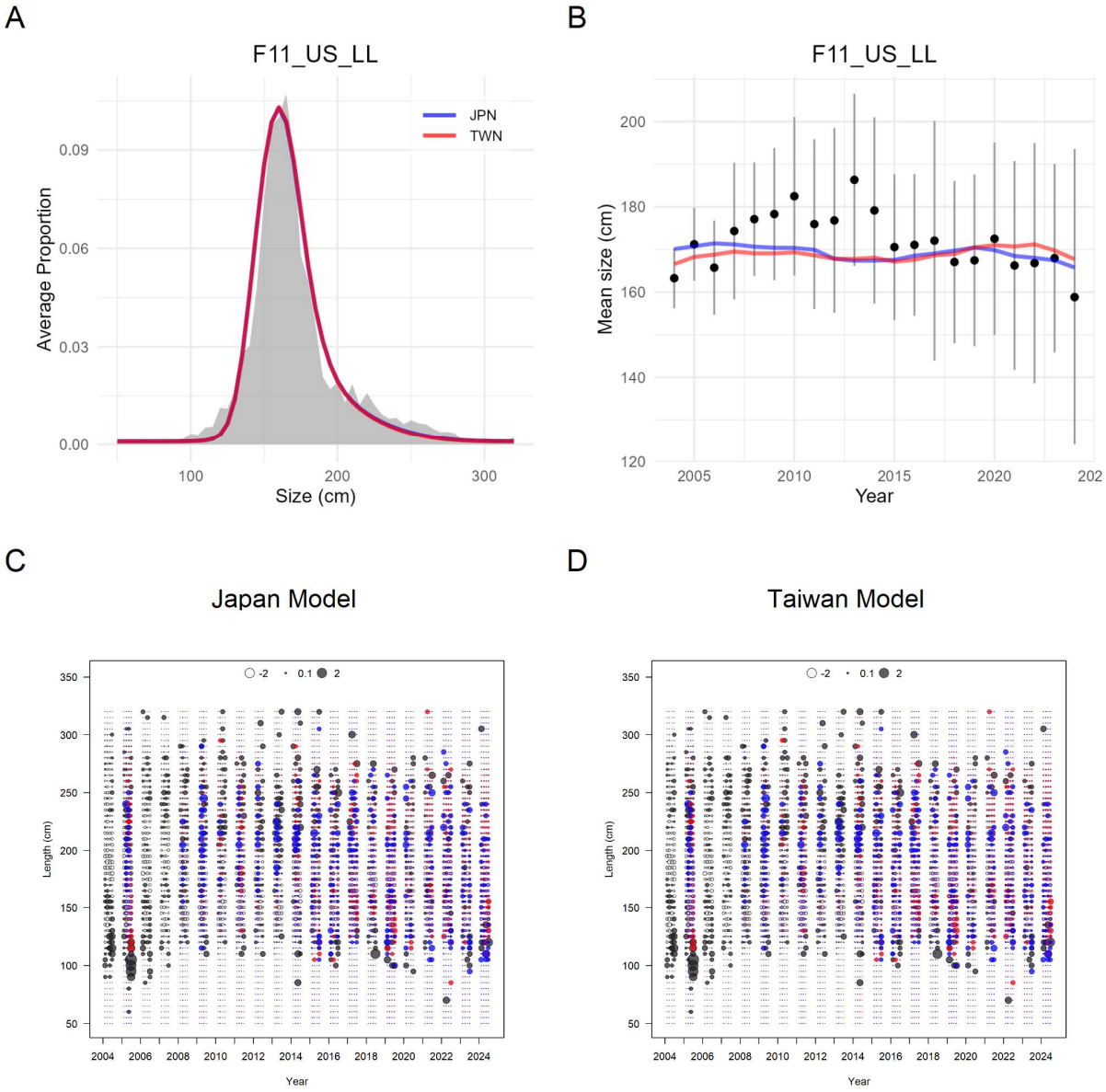


Figure 36. Model fits to F11_US_LL size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data. In (C) and (D) model residuals are color coded to indicate sex-specific length composition data and fits: unspecified (black), female (red), and male (blue).

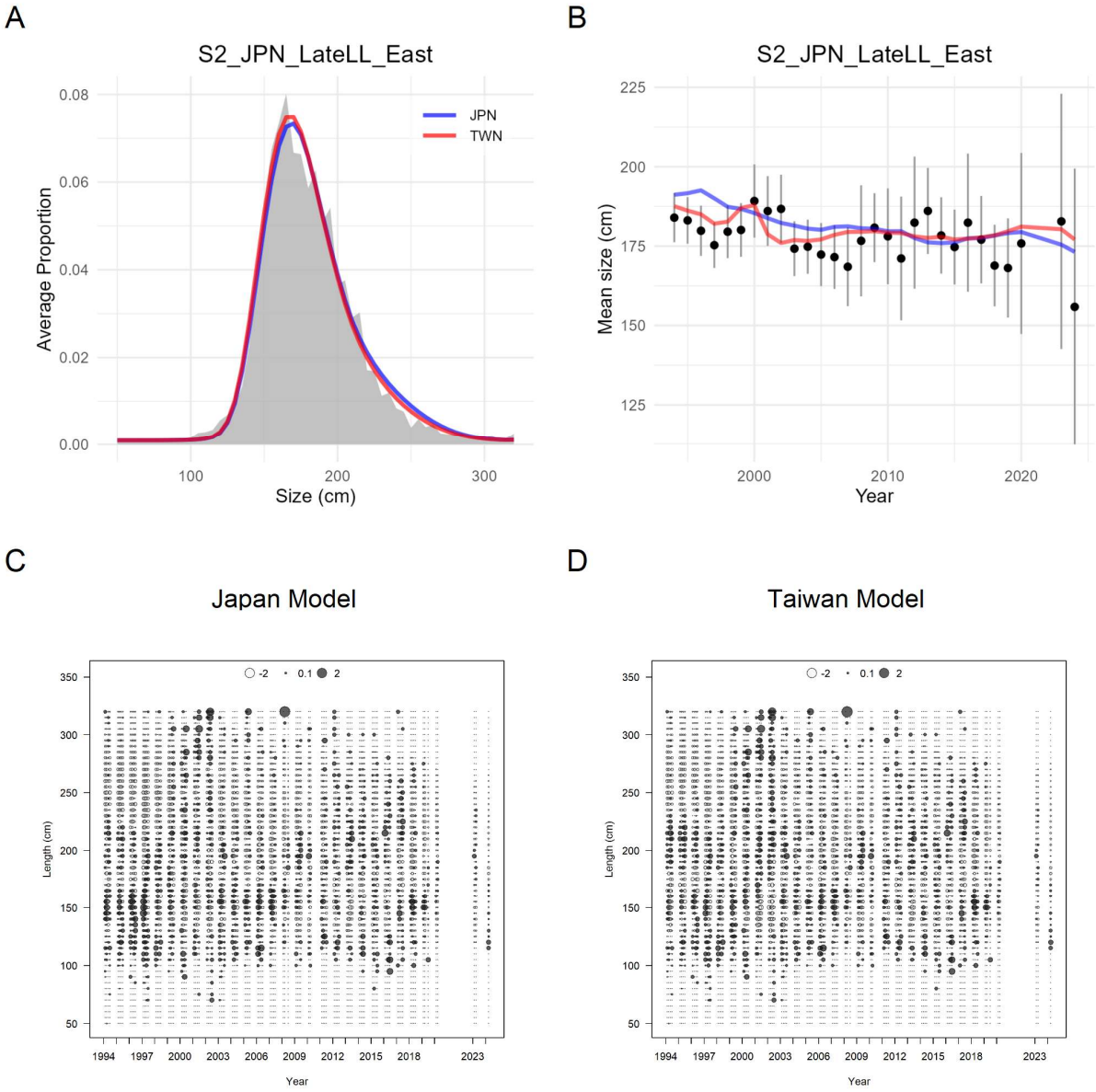


Figure 37. Model fits to S2_JPN_LateLL_East size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

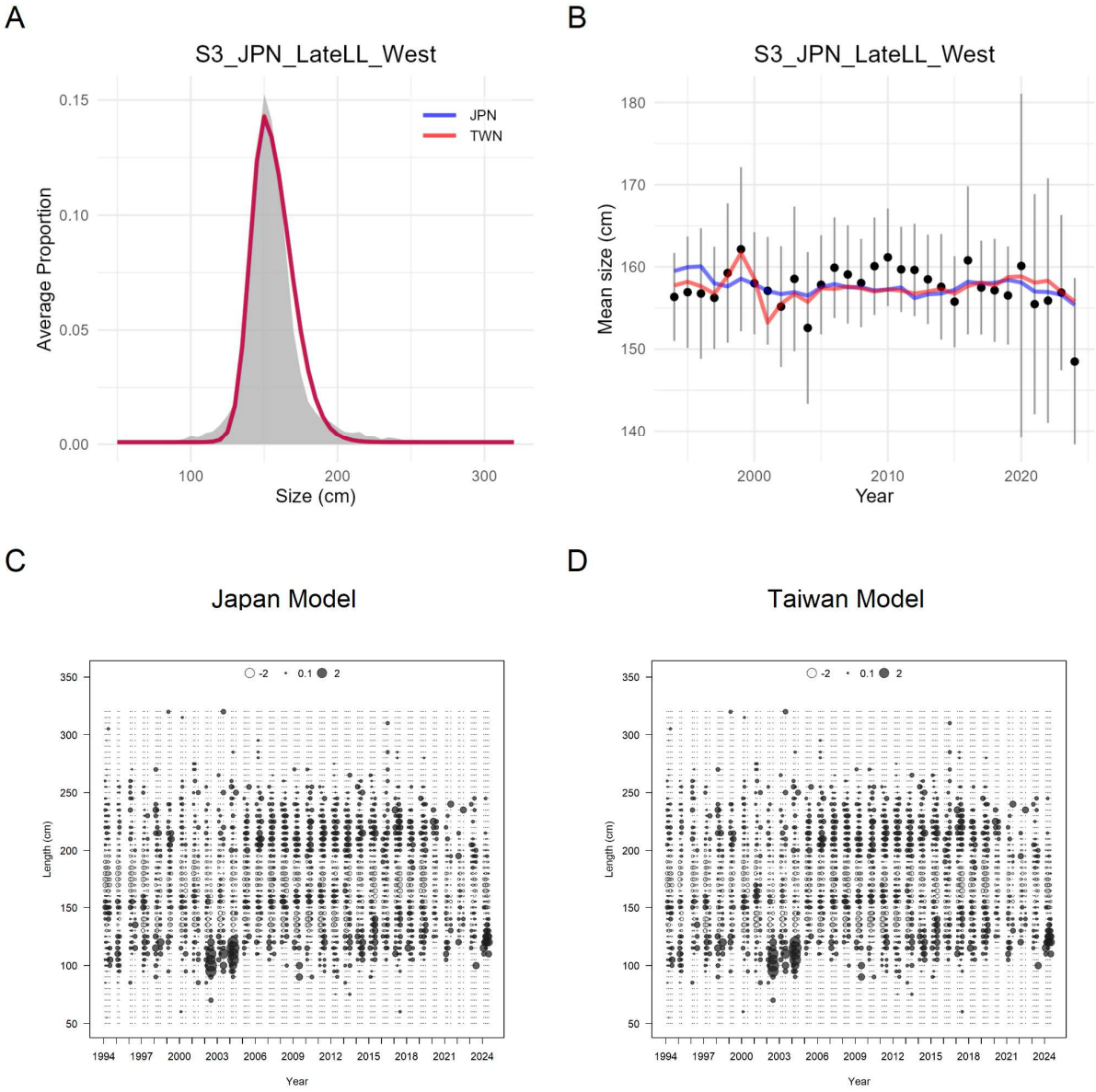


Figure 38. Model fits to S3_JPN_LateLL_West size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

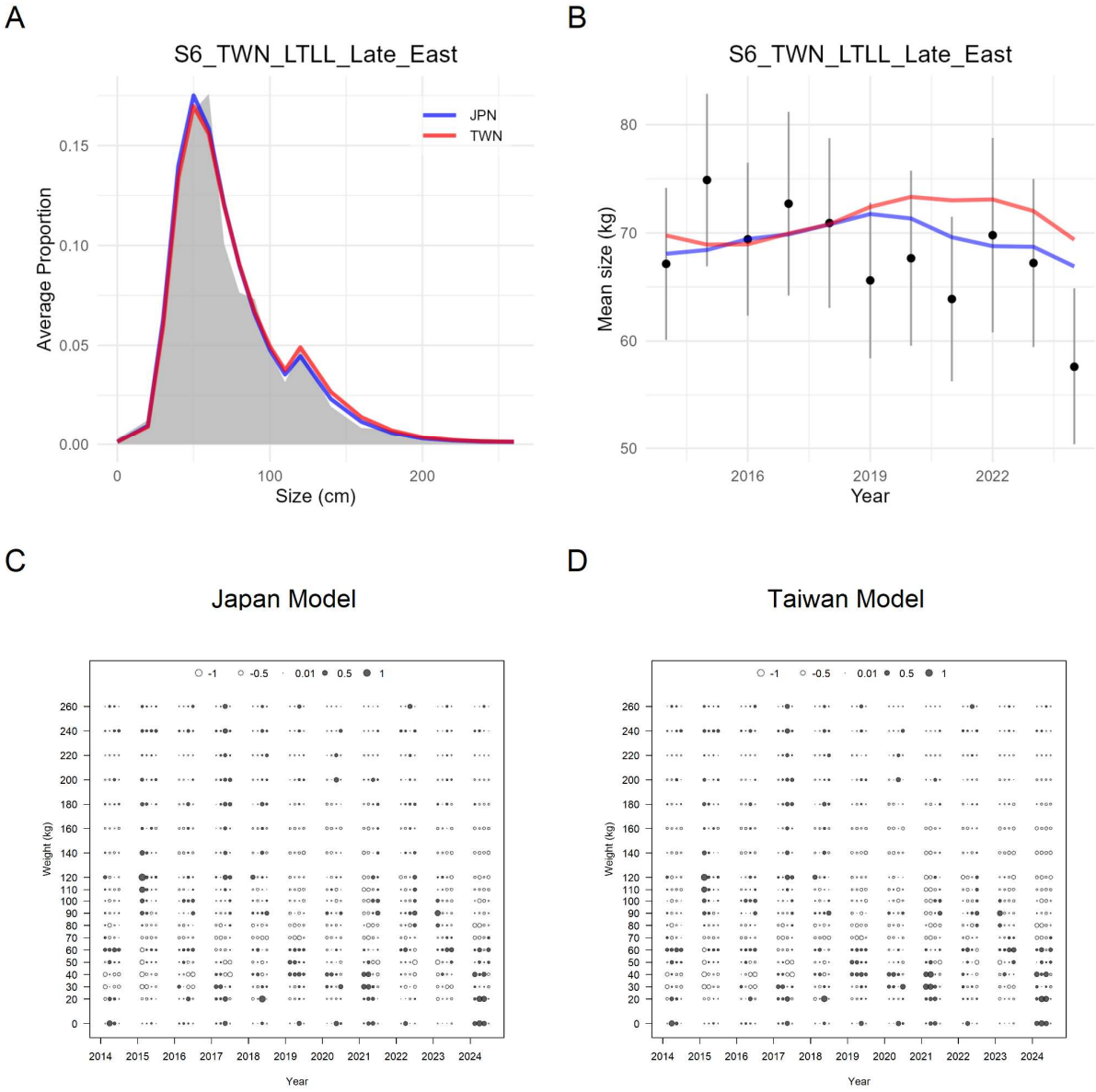


Figure 39. Model fits to S6_TWN_LTLL_Late_East size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

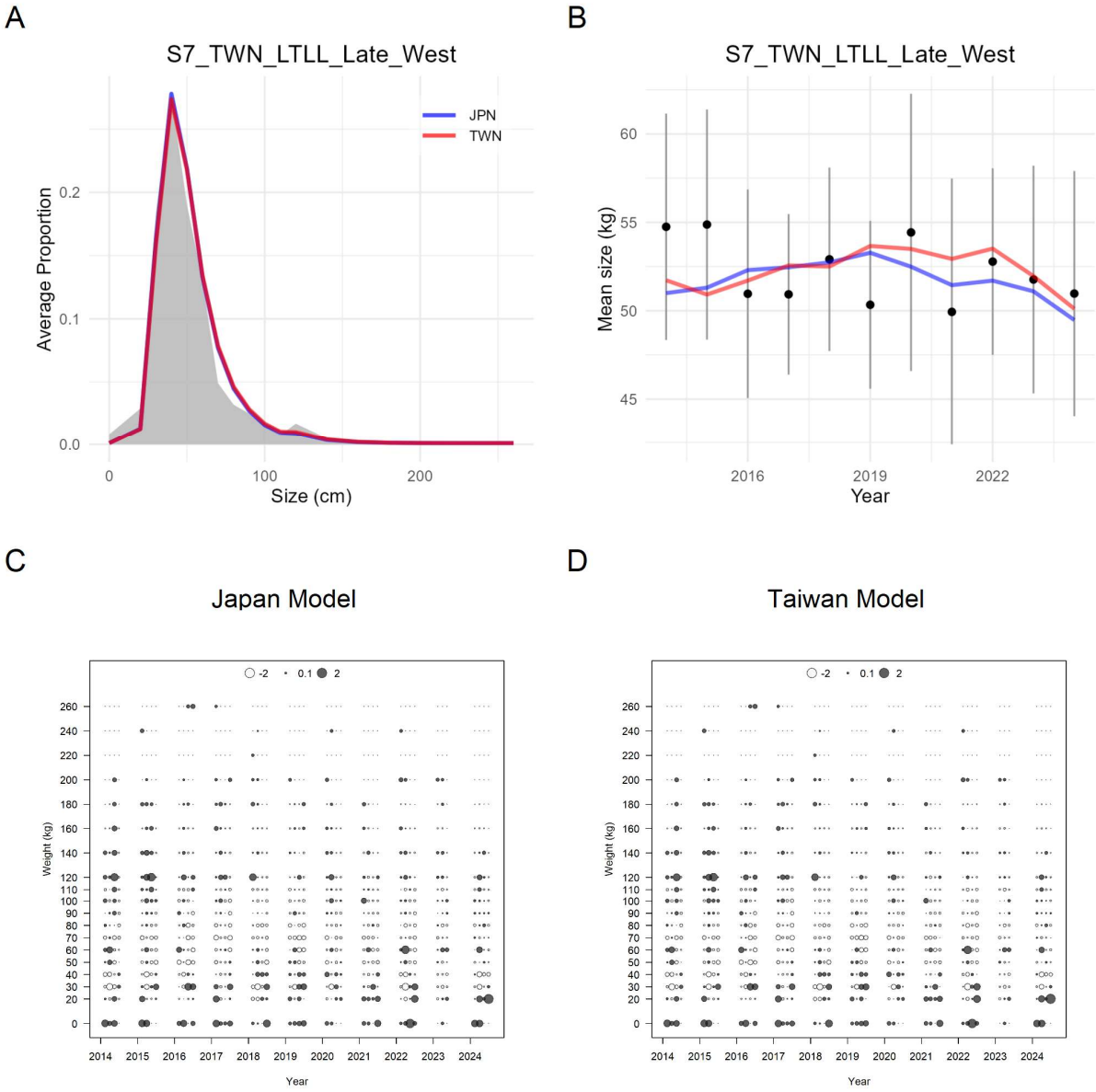


Figure 40. Model fits to S7_TWN_LTLL_Late_West size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data.

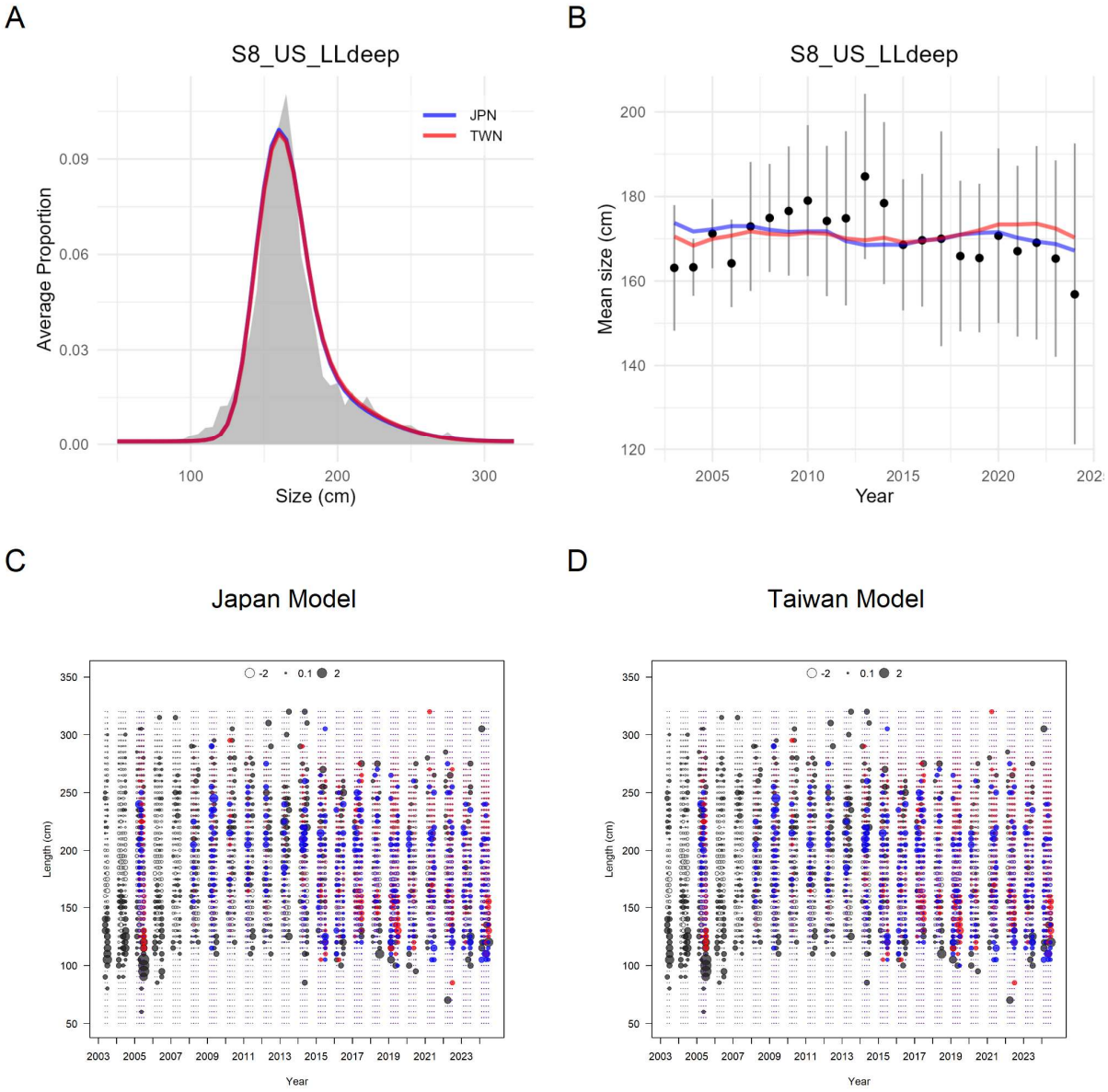


Figure 41. Model fits to S8_US_LLdeep size composition data (A) aggregated over all years, (B) annually, error bars are +/- 1 standard error based on the adjusted sample size in Stock Synthesis, (C) Pearson residuals of the Japan model fit, and (D) Pearson residuals of the Taiwan model fit to the annual size composition data. In (C) and (D) model residuals are color coded to indicate sex-specific length composition data and fits: unspecified (black), female (red), and male (blue).

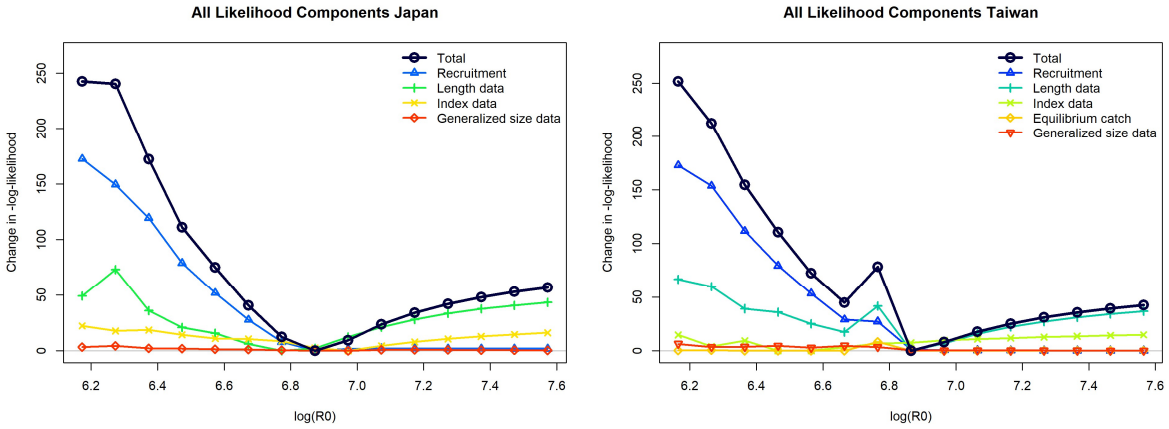


Figure 42. Overall likelihood component profiles of Ln(virgin recruitment) [log(R0)] for the Japan (left) and Taiwan (right) models.

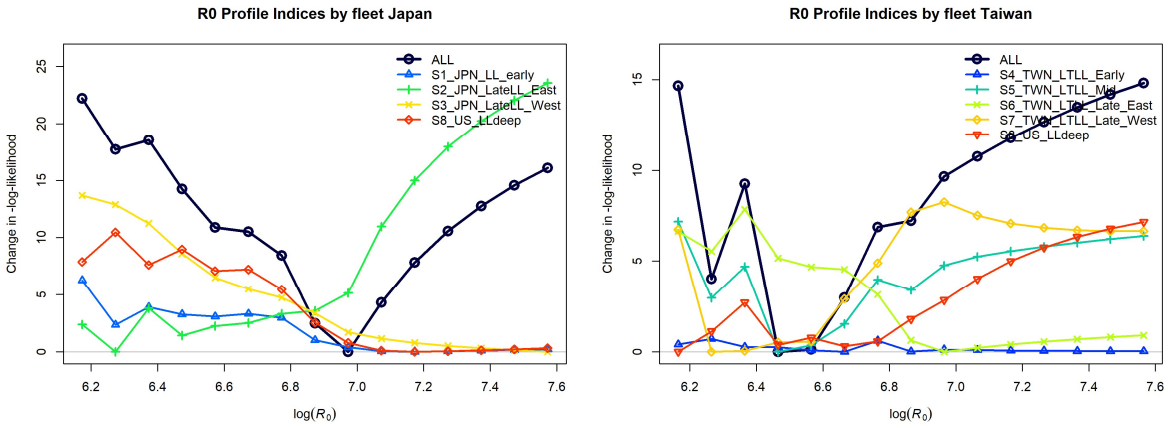


Figure 43. Index likelihood profiles of Ln(virgin recruitment) [log(R0)] for the Japan (left) and Taiwan (right) models.

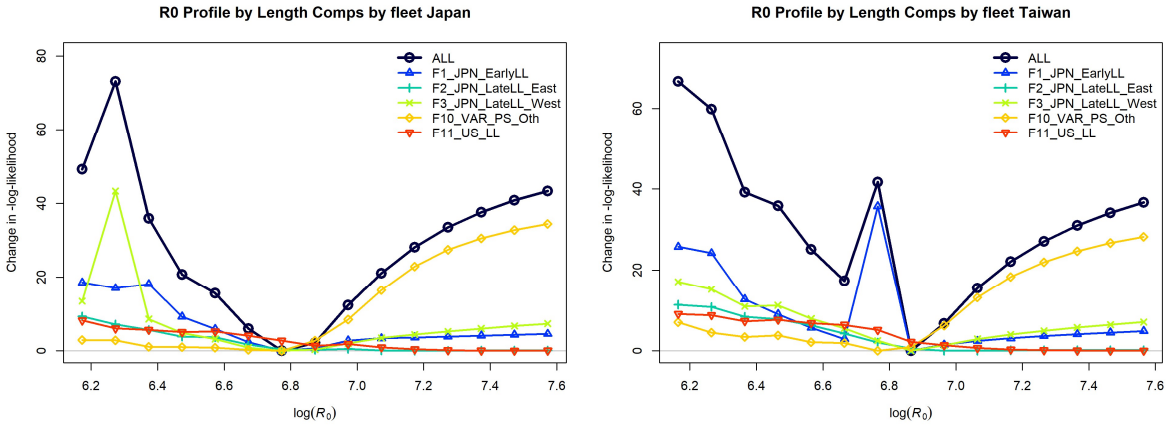


Figure 44. Length composition likelihood profiles of Ln(virgin recruitment) [log(R0)] for the Japan (left) and Taiwan (right) models.

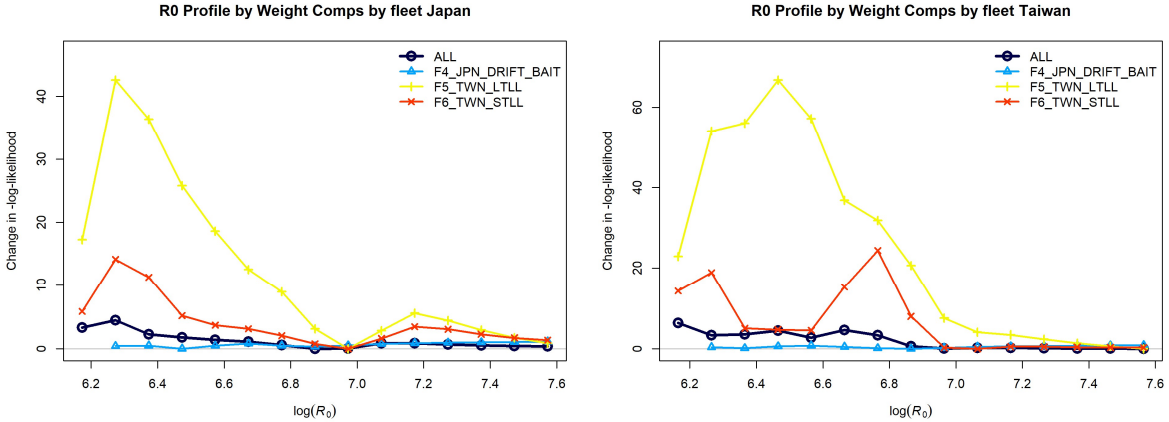


Figure 45. Size (weight) composition likelihood profiles of Ln(virgin recruitment) [log(R0)] for the Japan (left) and Taiwan (right) models.

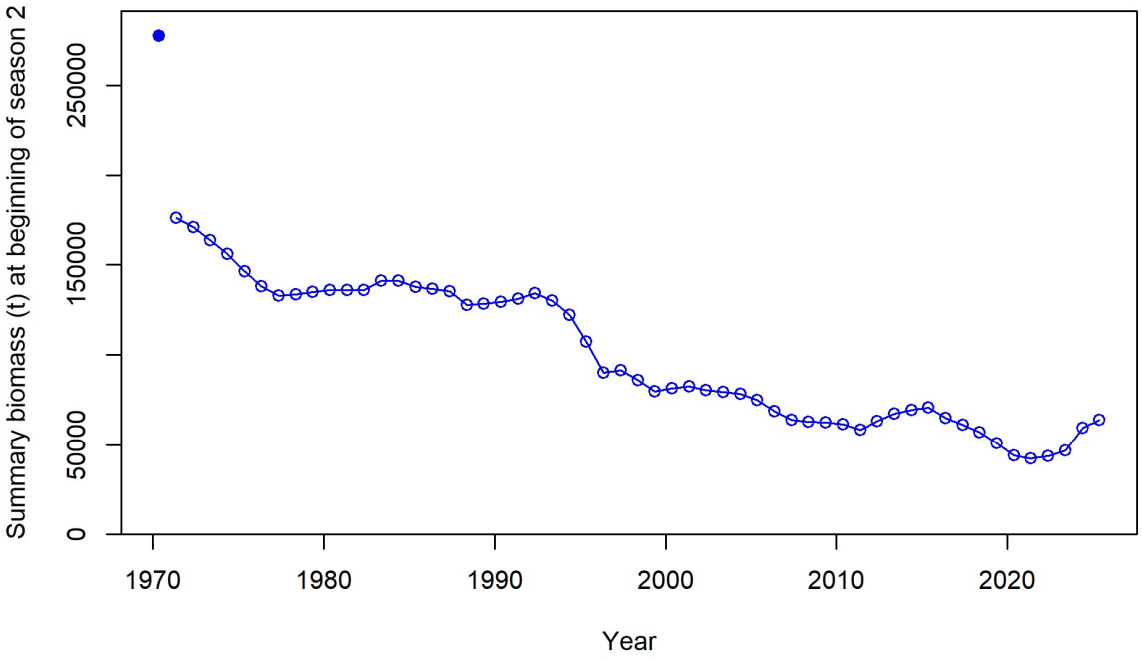


Figure 46. Biomass of age 1+ fish estimated at the beginning of season two from the Japan model.

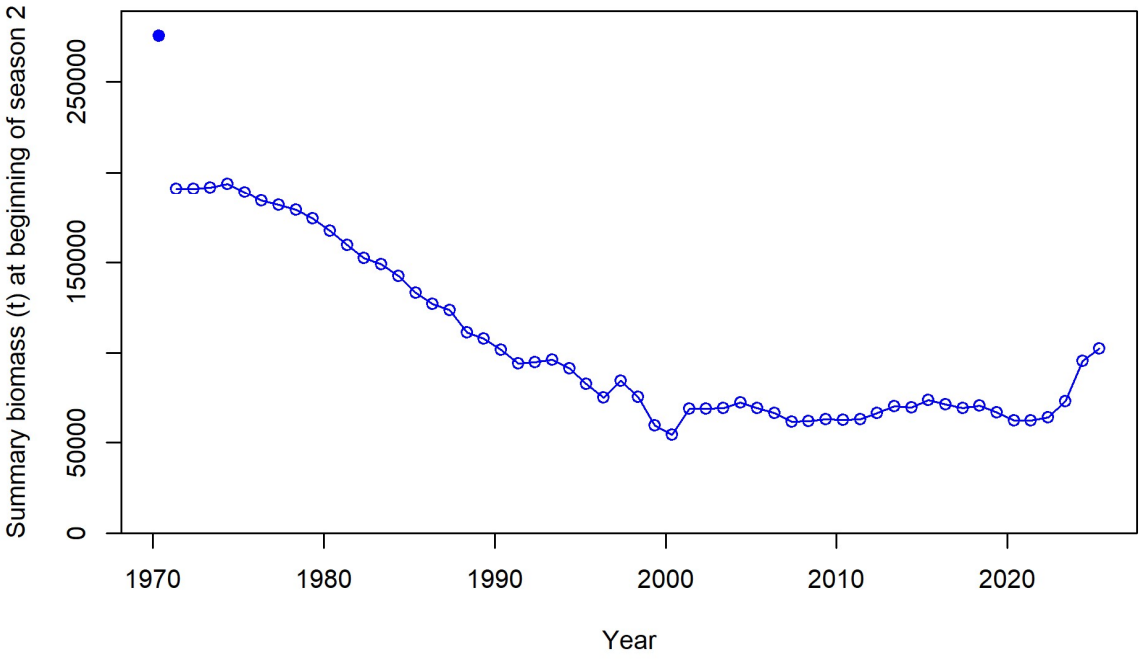


Figure 47. Biomass (t) of age 1+ fish at the beginning of season two estimated in the Taiwan model.

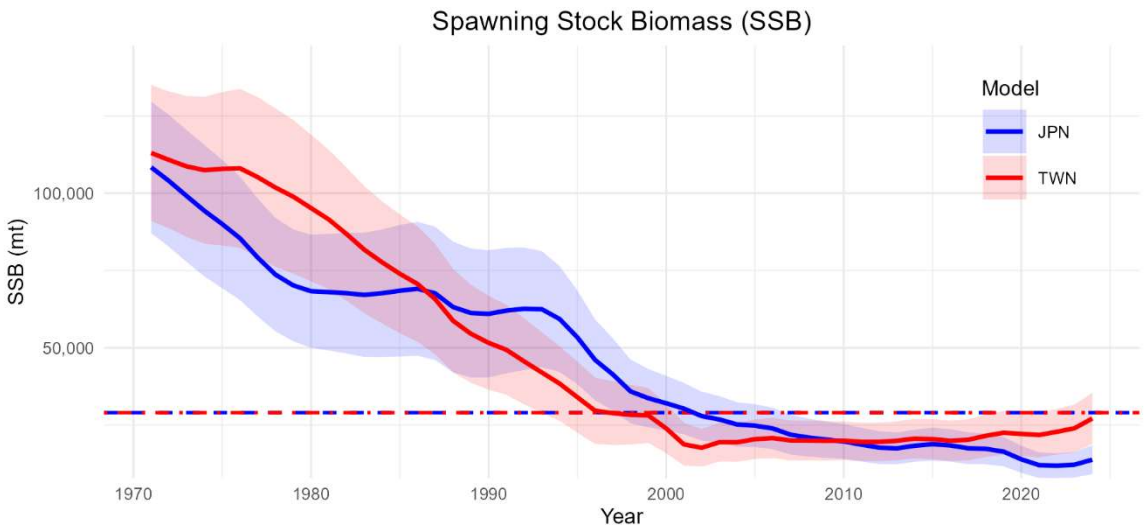


Figure 48. Estimated Female Spawning Stock Biomass (SSB) from the Japan and Taiwan models with 95% confidence intervals. SSB_{MSY} is indicated by the horizontal dashed line.

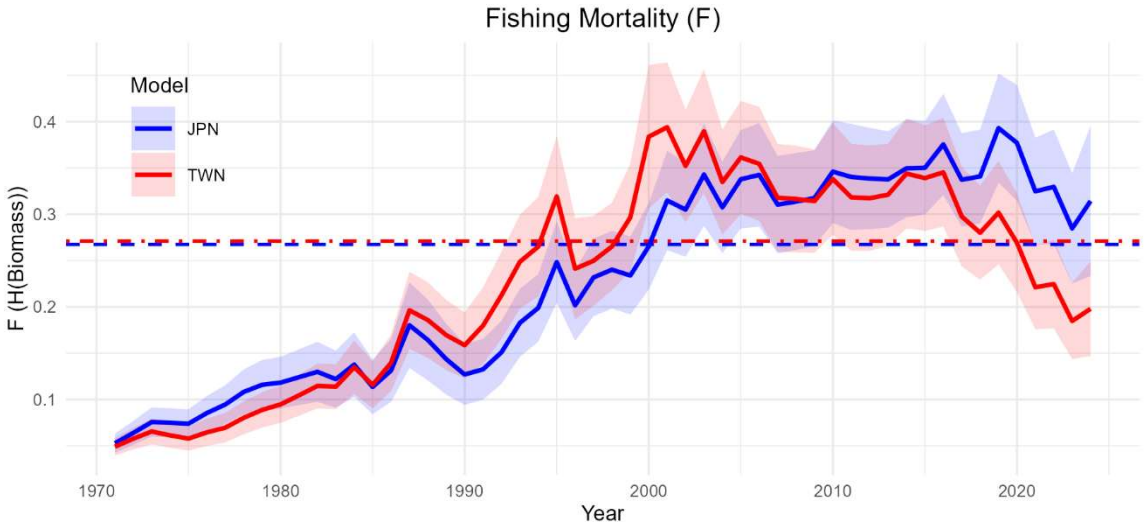


Figure 49. Estimated annual fishing mortality (as exploitation of biomass) from the Japan and Taiwan models with 95% confidence intervals. F_{MSY} is indicated by the horizontal dashed line.

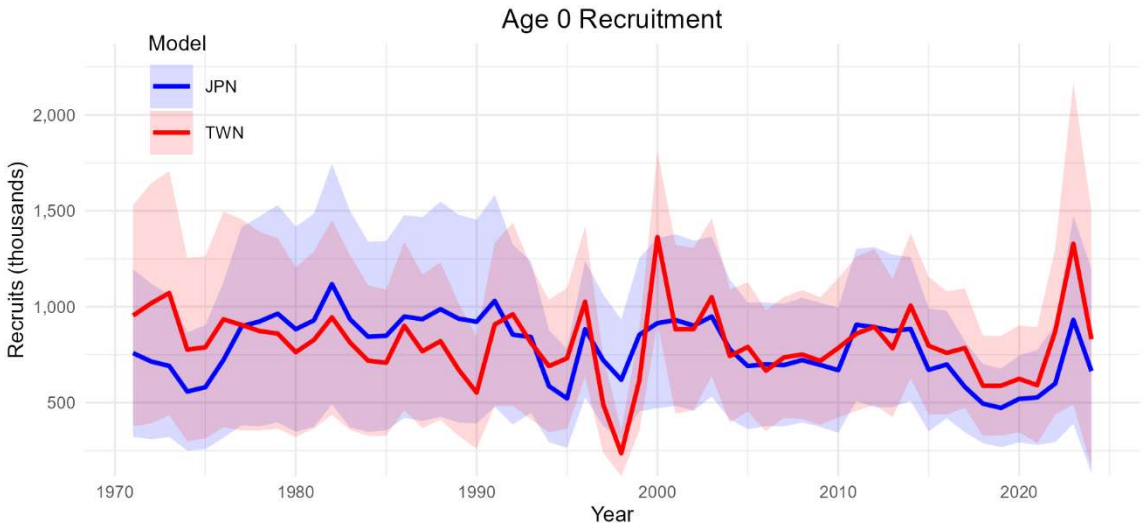


Figure 50. Estimated annual recruitment (thousands of age-0 fish) with 95% confidence intervals from the Japan and Taiwan models.

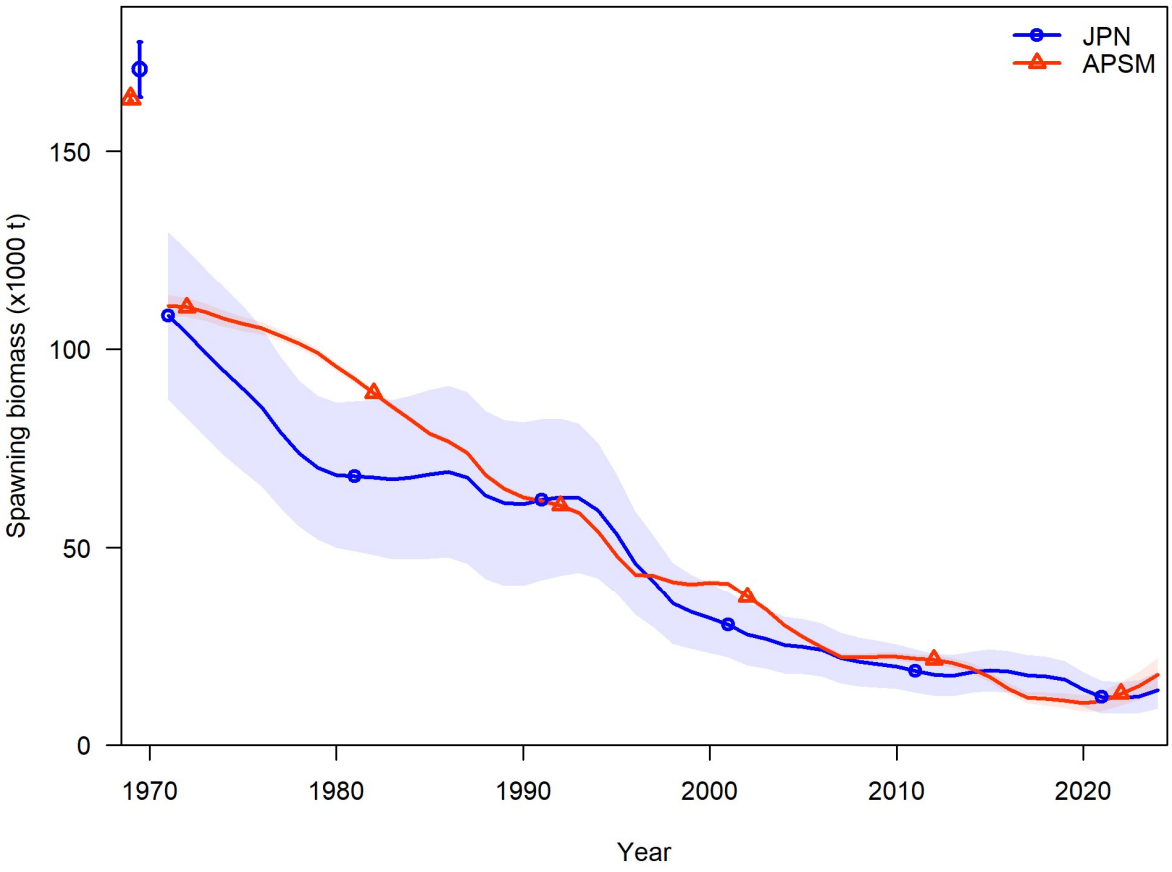


Figure 51. Plot of estimated spawning stock biomass (in metric tons) for the Japan model and the ASPM for the Japan model.

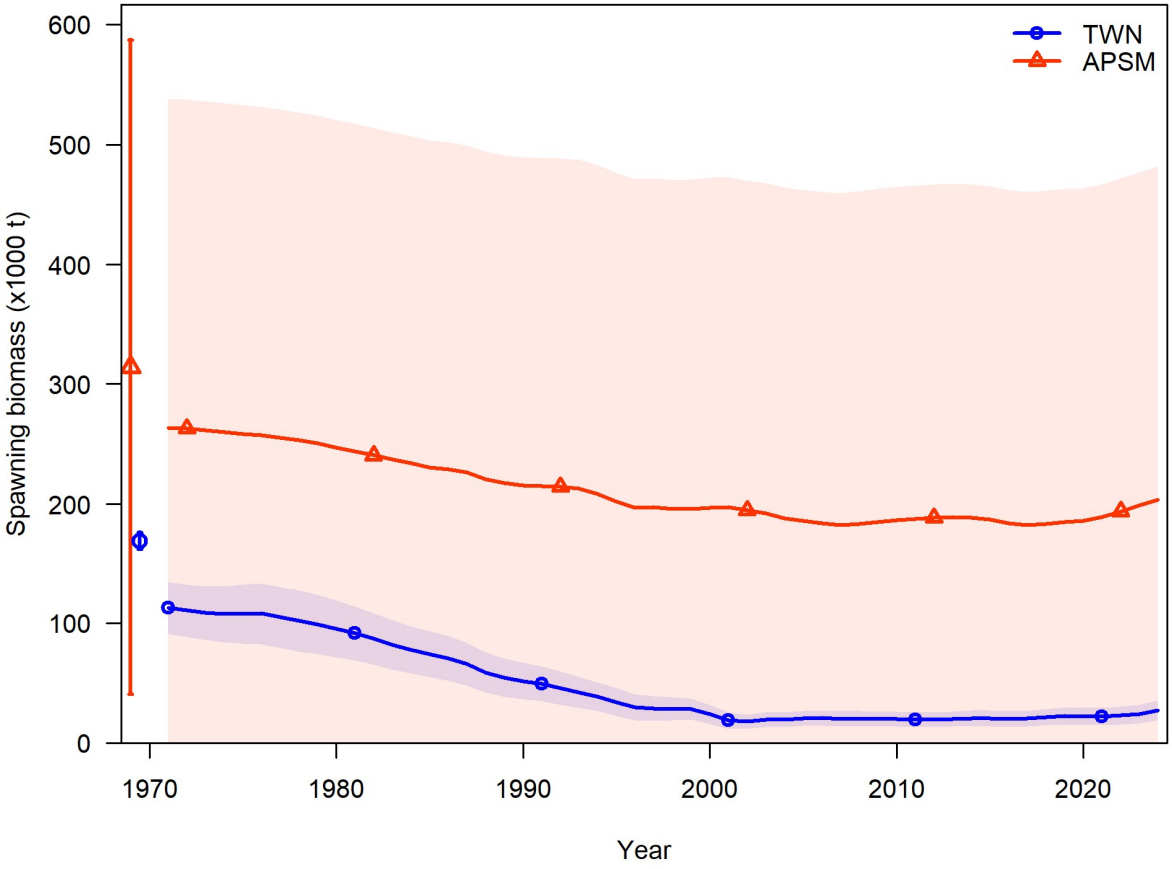


Figure 52. Plot of estimated spawning stock biomass (in metric tons) for the Taiwan model and the ASPM for the Taiwan model.