ISC/17/SHARKWG-1/5

Stock Assessment for Blue Shark in the North Pacific Ocean Using Stock Synthesis[1]

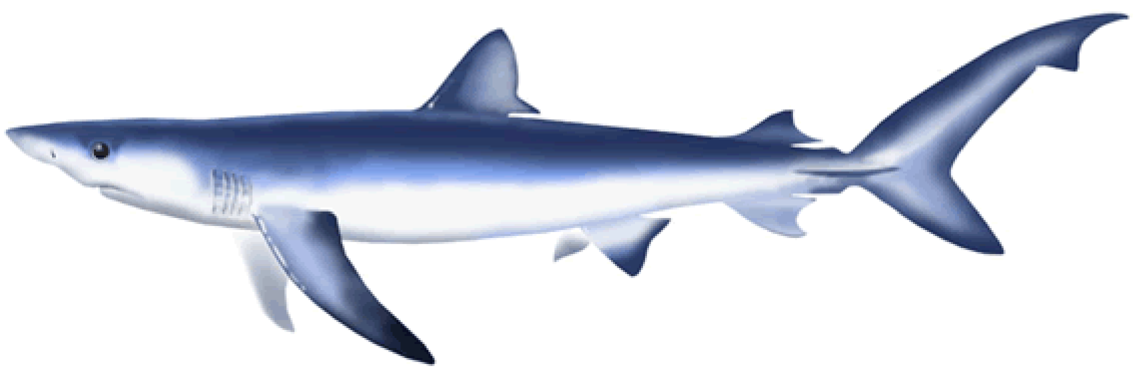
Felipe Carvalho[2], Mikihiko Kai[3], Tim Sippel[4], Kevin Piner[4]

[2]NOAA Pacific Islands Fisheries Science Center Honolulu, HI 96818 USA

[3]National Research Institute of Far Seas Fisheries, Fishery Research Agency Shizuoka, Japan

[4]NOAA Southwest Fisheries Science Center La Jolla, CA 92037 USA

Email: felipe.carvalho@noaa.gov



[1] Working document submitted to the ISC Shark Working Group Workshop, 17-24 March 2017, NOAA Southwest Fisheries Science Center, La Jolla, California U.S.A.

Document not to be cited without author's permission.

Table of Contents

Summary

This paper presents an age-based statistical catch-at-length stock assessment using Stock Synthesis (SS) for blue sharks in the North Pacific Ocean (NPO). In the previous blue shark stock assessment undertaken by the ISC in 2014 it was identified that the resulting stock status conclusions were extremely sensitive to the shape of the Low Fecundity Stock Recruitment function (LFSR), and consequently further investigation of how the biology of blue sharks can be modelled within the SS modelling framework was recommended for future assessments. For the stock assessment presented here numerical simulations were used to estimate probable values of stock- recruitment steepness for a Beverton-Holt stock-recruitment curve. The estimated value for steepness was then used for determining which combinations of the LFSR parameters are most representative of biological responses expected for blue shark. A suite of diagnostics were utilized to evaluate model convergence and fit to data sources of three scenarios based on alternative assumptions about stock productivity. Overall, all three models showed similar trends of spawning stock biomass (SSB) and management reference points over time. Estimates of SSB declined from 1971-1991, followed by an increase until the mid-2000's. More recently, SSB slightly declined from 2009-2013, followed by an increase in the final two years (2014-2015). Over the course of the modelled time series, estimated fishing mortality increased abruptly in the late 1970s and early 1980's with a peak around 1989, in response to higher catches. For the last two decades, fishing mortality (F) showed a declining trend, with recent values being close to those observed in the beginning of the time series. The historical trajectory of stock status revealed that north Pacific blue shark had experienced some levels of depletion and overfishing in previous years showing that the stock moved through the orange (overfishing) zone in the Kobe plot. However, in the last two decades, the stock condition returned into the Kobe green zone (no overfishing, not overfished).

1 Background

This paper presents one of two stock assessment approaches being developed by the ISC Shark Working Group (WG) for blue shark in the North Pacific Ocean (NPO). The WG agreed to use an age-based statistical catch-at-length stock assessment conducted using Stock Synthesis (SS) (Methot and Wetzel, 2013) (version 3.24F) to assess the status of the stock, as well as use a Bayesian Surplus Production (BSP) model to compare with the results from the SS model and the previous assessment conducted by the WG in 2014.

2 General assessment approach

In the previous blue shark stock assessment undertaken by ISC in 2014 using SS (Rice et al., 2014), it was identified that the resulting stock status conclusions were extremely sensitive to the shape of the Low Fecundity Stock Recruitment (LFSR) function, and consequently further investigation of how the biology of blue sharks can be modelled within the SS modelling framework was recommended. For the stock assessment presented here, Kai and Fujinami (2017) conducted a numerical simulation using the method of Mangel et al. (2010) to estimate probable values of stock- recruitment steepness for a Beverton-Holt stock-recruitment curve (Beverton and Holt, 1957) for the North Pacific blue shark. The estimated value for steepness was then used for determining which combinations of the LFSR parameters are most representative of biological responses expected for blue shark.

3 Biological inputs and assumptions

Blue sharks have a pan-Pacific distribution, and genetic evidence of distinct population structure within the Pacific has not been found (Taguchi et al., 2015). Conventional tagging in the eastern, central and western North Pacific regions has resulted in recoveries within each neighboring North Pacific region, providing evidence of wide movement throughout the North Pacific (Sippel et al., 2011). However, tagging data have not demonstrated movement across the equator (Sippel et al., 2011; Stevens et al., 2010). Consensus within the WG supports a single stock within the North Pacific, distinct from the South Pacific, although more information is needed to further explore the potential for size and sex segregation in the North Pacific as proposed by Nakano (1994). In addition to assumptions regarding stock structure, the other critical information on the biology of blue shark necessary for the SS assessment relates to sex-specific growth, natural mortality, maturity and fecundity. Biological assumptions and parameter values used in the SS models are summarized in Table 1.

3.1 Growth

The standard assumptions made concerning age and growth in the SS model are (i) the lengths-at-age are normally distributed for each age-class; and (ii) the mean lengths-at-age follow a von Bertalanffy growth curve. For any specific model, it is necessary to assume the number of significant age-classes in the exploited population, with the last age-class being defined as a “plus group”, i.e. all fish of the designated age and older. For the results presented here, 24 yearly age-classes have been assumed.

Sex-specific estimates of growth (Fujinami et al., 2016a) and length-weight parameters (Yamamoto et al., 2016) were assumed in the assessment – No attempt was made to estimate growth due to the uninformative nature of the size data to track cohorts through time.

A CV of 0.25 was used to model variation in length-at-age. All lengths reported from the assessment relate to pre-caudal length (PCL).

3.2 Natural mortality

Age and sex-specific natural mortality ogives were considered in the assessment. They were calculated based on the Method II proposed by Walters et al. (2016) and described in Semba and Yokoi (2016).

3.3 Maturity and fecundity

It is critically important to measure spawning potential in the correct units for stock assessment purposes. This assessment considered a single maturity ogive and did not consider age/length- specific changes in fecundity in the final set of model runs. In Section 5.2 we describe potential relationships between pre-recruit survival and spawning potential (essentially the spawner recruitment relationship) that were examined in the assessment.

For the purpose of computing the stock spawning biomass (SSB), we assume a logistic maturity schedule based on length with the age-at-50% maturity for females equal to 156.6 cm (Fujinami et al., 2016b). There is no information which indicates that sex ratio differs from parity throughout the lifecycle of blue shark. In this assessment the term stock spawning biomass (SSB) is a relative measure of stock spawning potential (the mature female population).

4 Data compilation

4.1 Spatial and temporal stratification

The assessment was based on a single North Pacific stock, bounded by the equator in the south, Asia in the west, and North and Central America in the east.

4.2 Temporal stratification

An annual (Jan 1-Dec 31) time-series of fishery data for 1971-2015 were used for the assessment (Figure 1).

4.3 Definition of fisheries

The WG estimated catches of many fisheries from different nations and member sources in an effort to understand the nature of fishing mortality. While the BSP assessment only considered a single catch series, the SS model used the 18 fisheries defined in Table 2. The primary sources of catch were from longline and drift gillnet fisheries, with smaller catches also estimated from purse seines, trap, troll, and recreational fisheries (Figure 2). As in the previous assessment, highest catches came from Japan and Taiwan, with newly available Mexican fishery data providing a relatively small, but important source of catch.

4.4 Catch

4.4.1 Japan

Offshore and distant water longline catches were estimated using two time-series of standardized CPUEs (1976-1993 and 1994-2015). Since the landings of sharks are frequently underestimated due to the lower catches when compared to other teleost species such as tunas and billfishes, total catches including retained and discard/released catches were estimated using a product of the yearly changes in standardized CPUEs and fishing effort. The former CPUE was estimated by Hiraoka et al. (2013a) and the latter CPUE was updated by Kai and Shiozaki (2016). The former and latter catches were converted to biomass using the mean weight by season and area (Hiraoka et al., 2013a). The estimation methods and estimated catch amount can be found in Kai et al. (2014) and Kai (2016), respectively.

Japanese coastal fishery catches (coastal and other longline, drift net, set-net, bait fishing, others) were updated from 1994 to 2014 (Kai and Yano, 2016). Most of the Japanese shark catch data were reported in species aggregated form as "sharks", thus the ratio of the catch of blue shark among those of sharks by fishing gear were calculated using available species specific landing data, and used to estimate the catch of blue shark. The Japanese coastal fishery catches prior to 1994 were provided in Yokawa (2012) and Kimoto et al. (2012).

4.4.2 Taiwan

Taiwan small scale longline catches were updated in Liu et al (2016a). Large scale longline catch was estimated in two areas (0-25 degrees north of equator; and northwards of 25 degrees) using catch rates multiplied by effort in the two separate areas (Tsai and Liu, 2016).

4.4.2 Republic of Korea

The Korean annual reports for the 2010 and 2011 WCPFC SC meetings indicated that the catch of major shark species reported in logbooks includes only blue and "other" sharks (reported as "porbeagle" sharks but since corrected to "other" sharks, Y. Kwon pers. comm.). Observer records for one year showed that 65% of the catches of major shark species was comprised of blue shark for one year. The Korean annual report to the WCPFC in 2010 indicated that the average CPUE of blue shark caught by Korean longliners was 0.07 (number/100 hooks) based on observer data. Using the annual aggregated shark catch and effort data submitted to the ISC, and an average blue shark size of 30 kg, the average size caught in a comparable Japanese longline fishery, estimated CPUE by year in number of blue sharks per 1000 hooks caught by Korean longliners ranged from 0.0 to 0.89 which is comparable to the average CPUE obtained by the Korean observer data. For this assessment, Korean blue shark catch was assumed to be equal to North Pacific species-aggregated shark catch reported to the ISC (various shark species, code SHK). Beginning in 2013, a small amount of shark catch was reported as blue shark, which was added to the species-aggregated shark catch for the assessment time series. Kwon et al. (2017) developed an independent estimate of Korean longline blue shark catch for the period 1973-2015. Catch estimates were derived by applying area-specific CPUE based on observer data to Korean longline fishing effort recorded in logbooks. Careful review of the catch estimation methodologies and time series was not possible in time for the assessment; however, the magnitude and trends in the catch time series were quite similar to those developed by the SHARKWG.

4.4.3 China

China longline species-specific catch and effort were available for 2007-2015 and effort data were available back to 2001. The mean annual CPUE for 2007-2015 was applied to effort data for 2001-2006 to estimate catch for those years. It was assumed that effort of Chinese longliners in the North Pacific was minimal prior to 2001.

4.4.4 Canada

Blue shark bycatch in Canadian fisheries were estimated from a combination of observer and logbook records from 1979-2015 for groundfish, salmon, sardine, albacore, hake and squid fisheries (King and Surry, 2016). Minor adjustments to previous estimates were based on newly available information.

4.4.5 USA

Blue shark catch in US fisheries including the Hawaii-based longline fleet, as well as west coast drift gillnet, recreational, albacore troll fleets and small longline fisheries were provided in Kohin et al (2016). Estimation methods were consistent with those used in the 2014 assessment, except the discard mortality rate estimate used for the Hawaii based longline fishery was updated, and catches from the albacore troll fishery (less than 1 mt annually) had not been previously estimated.

4.4.6 Mexico

Total blue shark catches were calculated from artisanal, commercial longline, and historical drift gillnet fisheries. Catches were sourced from annual fishery statistics yearbooks of SAGARPA (the Mexican fishery authority - provided by INAPESCA) from five Mexican States (Baja California, Baja California Sur, Sinaloa, Nayarit and Colima), published articles and reports (including grey literature) (Castillo-Geniz et al., 2017; Sosa-Nishizaki and Castillo-Geniz, 2016).

4.4.7 IATTC members

IATTC provided estimates of blue shark bycatch in tuna purse seine fisheries in the north EPO. The methods were the same as for the last stock assessment. The number of blue sharks caught from 1971-2015 was estimated from observer bycatch data, and observer and logbook effort data. Some assumptions regarding the relative bycatch rates of blue sharks were applied based on their temperate distribution and catch composition information. Estimates were calculated separately by set type, year and area. Small purse seine vessels, for which there are no observer data, were assumed to have the same blue shark bycatch rates by set type, year and area, as those of large vessels. Prior to 1993, when shark bycatch data were not available, blue shark bycatch rates assumed to be equal to the average of 1993-1995 rates were applied to the available effort information by set type, area and year. Numbers of sharks were converted to tons by applying an average annual weight estimate derived from blue sharks measured through the IATTC observer program.

4.4.8 SPC

Blue shark longline catches for non-ISC member countries in the WCPFC area north of the equator were estimated from SPC observer data holdings. Catches during 1995-2010 were estimated based from standardized CPUE values for each 5 x 5 degree cell multiplied by the effort reported in that cell summed on an annual basis. The non-ISC countries represented in the dataset include 12 countries, many of them that likely fish only south of the equator, thus it is believed that the north Pacific blue shark catch of non-ISC member countries represented in the WCPFC database is attributed to Federated States of Micronesia, Kiribati, Marshall Islands, Papua New Guinea and Vanuatu. Total dead removals are assumed to be the same as longline catches. For 2011-2014 the reported effort in the North Pacific (publically available Category 1 data; https://www.wcpfc.int/node/4648) was multiplied by the 2000-2010 average CPUE based on the estimated catch for non-ISC members divided by total effort data for the North Pacific.

4.5 Abundance indices

Indices of relative abundance were developed with fishery data from five nations or information sources. Four of these abundance indices were updated with data available since the 2014 assessment or reused from that assessment, and one new index was developed by Mexico.

4.5.1 Japan

Abundance indices from Japanese offshore and distant water shallow-set longline fisheries were updated for the period from 1994 to 2015 (Kai and Shiozaki, 2016). Catch per unit of effort (CPUE) was standardized using a generalized linear model with negative binomial error distributions. A continuous time series of data was used to standardize the CPUE without separation of the data after the Tsunami in 2011. The abundance indices before 1994 were estimated by Hiraoka et al. (2013b).

4.5.2 Taiwan

The abundance index was developed with a delta lognormal generalized linear model using observer data from the Taiwanese large scale longline fisheries for the 2014 assessment was updated through 2015 (Tsai and Liu, 2016).

4.5.3 USA

The abundance indices developed with delta lognormal generalized linear models using observer data from the Hawaii deep-set and shallow set longline fisheries for the 2014 assessment were updated through 2015 (Carvalho, 2016).

4.5.4 Mexico

An abundance index was developed with generalized linear models using observer data from Mexican longline fishery in Pacific (Fernandez-Mendez et al., 2016).

4.5.5 SPC

The same relative abundance index developed with longline observer data during 1993-2009 for the 2014 assessment was included (Rice and Harley, 2014).

4.6 Catch-at-length

Length composition data were provided for different fisheries from Japan, Taiwan, South Korea, China, USA and Mexico. The request was for sex-specific data in the observed measurement units (FL – fork length, TL – total length, DL – dorsal length, AL – alternate length) which were subsequently converted to precaudal length (PCL) using agreed conversion equations. The conversion equations used to prepare data for this assessment were (Carvalho and Sippel, 2016);

PCL = (FL x 0.894) + 2.547

PCL = (TL x 0.748) + 1.063

PCL = (AL x 2.462702) + 12.7976

PCL = (DL x 2.56) + 9.97

The data were also requested with vessel coordinates where samples were taken in order to investigate spatially-explicit size and sex structure. Some data were provided with exact coordinates and some data were summarized into spatial blocks (1x1, 5x5, or 20x10) (Sippel et al., 2016). Some of the conversion equations used in the spatial analyses differed from those listed above.

4.6.1 Japan

Japan provided blue shark size data from the following fishery data sources; Kinkai shallow longlines (Hiraoka et al., 2011), research and training longline vessels (Ohshimo et al., 2014, 2016), small scale longlines (Kimoto et al., 2012), the longline observer program, and drift gillnets (Yokawa, 2012).

Size data from longline gear comprised 97% of all Japanese size data and it was divided into shallow set longline (shallow LL) and deep set longline (deep LL) based on operational patterns (i.e., night or daytime), the number of hooks per basket (HPB), and location of sets.

Size data categorized as “Kinkai shallow” included data from shallow-set research and training vessels, shallow-set observer longlines, small scale shallow-set longlines, and Kinkai-shallow longliners, fishing at night and targeting sharks and swordfish.

Size data categorized as “Kinkai deep” included data from deep-set research and training vessel, deep-set observer longlines, and deep-set small scale longlines fishing during the day for tunas. and size data from the other Japanese fisheries were categorized as “Enyo-deep”.

Size data from the large mesh drift gillnet fishery was also provided.

4.6.2 Taiwan

Size data of small scale longline collected by observers from 2014-2015 were reported by Liu et al. (2016a). Size data of large scale longline collected by observers from 2004 to 2014 were reported Liu et al (2016b).

4.6.3 South Korea

Lengths measured by observers on South Korean longline vessels were provided from 2005-2008 and 2013-2014. The majority of sampling prior to 2008 was from the WCPO, but starting in 2008 sampling effort moved to the EPO (Kim et al., 2016).

4.6.4 China

No documentation about the collection of size data from Chinese fisheries was provided. However, 2146 lengths measured in FL, TL or PCL were provided during 2009-2015 from observers on Chinese longline vessels.

4.6.5 USA

Collection of composition data by observers in the Hawaii-based longline fisheries (deep and shallow set) has been described in Walsh and Teo (2012) and Sippel et al. (2014), from observers in the US West

Coast drift gillnet fishery (Teo et al., 2012), and a small-scale scientific survey of juvenile sharks in Southern California (Runcie et al., 2014).

4.6.6 Mexico

Size data were collected by observers opportunistically deployed in Mexico's Ensenada and San Carlos based longline fleets during 2006-2014 (Castillo-Geniz et al., 2017).

5 Population and fishery dynamics

The model partitions the population into 24 yearly age-classes in the North Pacific Ocean. The last age-class comprises a "plus group" in which mortality and other characteristics are assumed to be constant. The population is "monitored" in the model at yearly time steps, extending through a time window of 1971-2015. The main population dynamics processes are as follows:

5.1 Abundance indices

CPUE series are critical to every assessment and candidate standardized abundance indices were developed from catch and effort data of Japanese, Taiwanese, Mexican, and US longline fisheries, and longline fisheries in the tropical north Pacific subject to the SPC observer program (Figure 3). It is well known that bias and uncertainty in the assessment results can occur if multiple indices with confounding trends are used in the same assessment. A suite of criteria were therefore used by the WG to select indices for the base case and sensitivity runs from the candidate indices. Key criteria include data quality, spatio-temporal coverage of data, potential changes in regulations and/or fishing operations, and the adequacy of diagnostics from model-based standardizations.

5.2 Recruitment and the Low-Fecundity Spawner-Recruitment Relationsip (LFSR)

In this model "recruitment" is the appearance of age-class 1 fish. The results presented in this WP were derived using one recruitment episode per year, which is assumed to occur at the start of each year. Annual recruitment deviates from the recruitment relationship were estimated, but constrained reflecting the limited scope for compensation given estimates of fecundity. As in the previous ISC blue shark stock assessment, a survival based spawner-recruitment function was used (Taylor et al., 2013) which we refer to as the Low Fecundity Spawner Recruitment relationship (LFSR).

Recruitment (R_y) in each year is then defined as:

$$R_y = S_y B_y \qquad \text{Equation 1}$$

Where B_y is the spawning output in year *y* and S_y is the pre-recruit survival given by:

$$S_y = \exp\left(-z_0 + (z_0 - z_{min})\left(1 - \left(\frac{B_y}{B_0}\right)^{\beta}\right)\right) \qquad \text{Equation 2}$$

Where:

$z_0 = -\log\left(\frac{B_y}{B_0}\right)$, where R_0 is the recruitment at equilibrium, resulting from the exponential of the estimated $\log(R_0)$ parameter, and $B_0$is the equilibrium spawning output.

$z_{min} = z_0(1 - s_{Frac})$ is the limit of the pre-recruit mortality as depletion approaches 0, parameterized as a function of s_{Frac} (which represents the reduction in mortality as a fraction of z_0); and, *Beta* (β) is a parameter controlling the shape of density-dependent relationship between spawning depletion and pre-recruit survival. During the previous ISC blue shark stock assessment no information regarding the stock recruitment relationship was available.

In the way that the LFSR is set up in SS, values of β < 1 has survival increasing fastest at low spawning output (concave decreasing survival (Figure 4A), whereas β > 1 has the increase in survival occurring fastest closer to the unfished equilibrium (convex decreasing survival (Figure 4B)). As observed by Rice et al. (2014) it is unlikely that blue shark survival would decrease fastest at low stock size; instead it is reasonable to expect that for a low-fecundity species, offspring survival would decrease faster due to competition when the population approaches carrying capacity (β > 1). Then, Rice et al. (2014) considered a wide range of LFSR shapes which gave similar productivity to that assumed in the production model developed simultaneously at that time. The selected values were 0.1, 0.3 and 0.5, for s_{Frac} and 1, 2 and 3 for β.

Kai and Fujinami (2017) applied the simulation method developed by Mangel et al. (2010) to estimate probable values of stock- recruitment steepness (h) for a Beverton-Holt stock-recruitment curve for North Pacific blue shark. Results indicated that the mean steepness (h) was μ_h = 0.67 with a standard deviation of σ = 0.073. We did not attempt to estimate β or s_{Frac} inside the stock assessment model because it is a task harder than estimating h as an extra parameter is involved. However, using equations from Taylor et al. (2013), s_{Frac} could be parameterized in terms of h, Z_0, and β (equation 3).

$$S_{frac} = \log(5 * h)/(Z_0 * (1 - 0.2^{\beta})) \quad \text{Equation 3}$$

Using equation 3 and life history information provided by Kai and Fujinami (2017) we calculated s_{Frac} under the selected values of 1, 2, and 3 for β. The resulted values for s_{Frac} were 0.467, 0.391, and 0.378 for β fixed at 1, 2, and 3, respectively (Figure 5).

5.3 Initial population state

It is not assumed that the blue shark population was in an unfished state of equilibrium at the start of the model (1971) as significant longline fishing occurred in the region from the 1950s and in Japanese coastal waters prior to that. As in the previous ISC blue shark stock assessment this stock assessment assumed an initial equilibrium catch of 40,000 mt.

This value represent approximately 100% of the first four years estimated catch. For this approach we had to choose a selectivity to assign this catch to. The selectivity estimated for one of the Japanese fleets (F4 JPN_KK_SH) was selected as it dominated catches in the early years and its selectivity was not extreme towards small or large fish.

The population age structure and overall size in the first year is determined as a function of the estimate of the first years recruitment (R1) offset from virgin recruitment (R0), the initial ‘equilibrium’ fishing mortality discussed above, and the initial recruitment deviations. As the size data were found to be uninformative about initial depletion and recruitment variation only a small number of initial recruitment deviates were estimated.

5.4 Selectivity curves

A double-normal functional form was assumed for all selectivity curves and an offset on the peak and scale was estimated for sex-specific differences in selectivity that were evident in the data. Selectivity is fishery-specific and temporal variations in selectivity were captured by the time blocks employed for F8 (2011; 2006-2014; 2015), F14 (1990-2005; 2006-2015), and F16 (1995-2005; 2006-2015). A cubic spline was used for fitting to size composition data for F17, since it was not possible to obtain model solutions using the double-normal functional form due to extreme peaks in the size-composition data. The parameterization of the cubic spline function estimates a starting and ending gradient and a selectivity value at each node using a smoothing function to connect the nodes (cubic spline selectivity curve). Given its flexibility, the benefit of this function is not just to increase additional process but also reduce the potential misfit of size compositions without introducing too many highly-correlated nodes. Selectivity patterns of fisheries without size composition data were mirrored to (assumed equal to) the selectivity patterns of fisheries with similar operations and areas for which a selectivity pattern was estimated. Mirrored selectivity patterns were based on expert opinions of members of the working group (Table 3).

5.5 Parameter estimation and uncertainty

Model parameters were estimated by maximizing the log-likelihoods of the data plus the log of the probability density functions of the priors, and the normalized sum of the recruitment deviates estimated in the model. For the catch and the CPUE series we assumed lognormal likelihood functions while a multinomial was assumed for the size data. The maximization was performed by an efficient optimization using exact numerical derivatives with respect to the model parameters (Fournier et al. 2012). Estimation was conducted in a series of phases, the first of which used arbitrary starting values for most parameters. The SS control file "BSH.ctl" documenting the phased procedure, initial starting values and model assumptions are available from the lead author.

The Hessian matrix computed at the mode of the posterior distribution was used to obtain estimates of the covariance matrix. This was used in combination with the Delta method to compute approximate confidence intervals for parameters of interest.

5.6 Data weighting

Some size and sex composition data of catch were available. Many of the time series suffered from low sample sizes and inconsistencies across years. We assumed an annual sample size proportional to the number of fishing trips, with a max of 100, for each record as:

$ESS_{j,y}$ is the annual effective sample size for the fleet j in year y, and it is calculated by (equation 4):

$$ESS_{j,y} = max\ (n, 100) \qquad \text{Equation 4}$$

Where n is the number of fishing trips.

It is well known that the results of fishery stock assessments based on integrated model can be sensitive to the values used to weight each of the data types included in the objective function. The weight given to each data point in a stock assessment model is determined by a measure of the assumed size of the error associated with that point: typically a coefficient of variation (CV) for abundance indices, and a sample size for composition data. If we change the data weighting, we change the balance between the different data sets, and thus change the parameter estimate. Punt (2016) provided a comprehensive review and a comparison of various iterative re-weighting methods for length composition data. The iterative re-weighting approach attempts to reduce the potential for particular data sources to have a disproportionate effect on total model fit, while creating estimates of uncertainty that are commensurate with the uncertainty inherent in the input data.

In this stock assessment we conducted a two stage Francis (2011) data weighting approach. In stage one we assume a minimum average standard error (SE; on the natural log scale) for each CPUE series. In stage two, the McAllister and Ianelli-2 method (using the harmonic mean) is applied to estimate the effective sample size of each length composition data from the residuals of the Stock Synthesis model fit to the data.

Stage 1. The relative CPUE to its mean were assumed to have log-normally distributed errors with standard error (SE) in log-space (log(SE)) which was approximated as sqrt ($\log(1+CV^2)$). The log (SE) of each CPUE were estimated by the statistical model in the standardization process. The estimated log (SE) only capture observation error within the statistical model but it does not reflect the inherent process error between the unobserved vulnerable population and the observed CPUE. We therefore assumed a minimum average log (SE) for each CPUE of 0.1. If the average log(SE) for each CPUE was smaller than 0.1, the estimated log (SE) was scaled to 0.1 by adding a constant value to the time series of estimated log (SE). If the average estimated log (SE) was larger than 0.1 the values were not changed.

Stage 2. After an initial model run with the input CVs adjusted for each CPUE as described above, the input sample sizes for the length composition data for fleets F1, F3, F4, F5, F7, F8, F14, F16, and F17 were adjusted one time with variance adjustment multiplication factors so that the sample size entered for each length composition data set was equal to the effective sample size obtained using the McAllister and Ianelli (1997) method.

6 Model diagnostics

There are limited diagnostics available for assessing the goodness of fit and identifying model misspecification in integrated fishery stock assessment models (Carvalho et al., 2017).

6.1 Residual analysis

Residuals are examined for patterns to evaluate whether the model assumptions have been met (e.g., Wang et al., 2009). Many statistics exist to evaluate the residuals for desirable properties. One way is to calculate, for each abundance index, the standard deviation of the normalized (or standardized) residuals divided by the sampling (or assumed) standard deviation (SDNR) (Breen et al., 2003; Francis, 2011; Carvalho et al., 2017). The SDNR is a measure of the fit to the data that is independent of the number of data points. A relatively good model fit will be characterized by smaller residuals (i.e. close to zero) and a SDNR close to 1. In addition, the root-mean-square-error (RSME) was used as a goodness-of-fit diagnostic, with relatively low RMSE values (i.e., RMSE < 0.2) being indicative of a good fit.

6.2 Age-structured production model (ASPM)

The ASPM diagnostic intends to evaluate the influence of data sets on absolute abundance (Maunder and Piner, 2015; Carvalho et al., 2017). This diagnostic may also be used to determine if a stock is recruitment-driven, fishery-driven, or a combination of both. Here we use the ASPM to determine whether information on temporal recruitment variability is needed to interpret the information about absolute abundance contained in the index of relative abundance. To conduct the ASPM diagnostic we follow the protocol provided in Mint-Vera et al. (2017) as follows:

1) run the SS base case model;
2) fix selectivity parameters at the maximum likelihood estimate (MLE) from the base case model,
3) turn off the estimation of all parameters except the scaling parameters and the parameters representing the initial conditions (a parameter for the equilibrium recruitment and a parameter for the equilibrium fishing mortality), set the recruitment deviates to zero (early recruitment and model period recruitments), and set the recruitment bias correction to zero (in order to achieve this in SS V3.24f the estimation phase of the recruitment deviates needs to be set to a large number, e.g. 50, and the maximum estimation phase needs to be set to a smaller value, e.g. 10);
4) fit the model to the indices of abundance only;
5) compare the estimated trajectory to the one obtained in the fully integrated model.

6.3 R0 profile

Likelihood profiles are used to check that a solution has actually been found (a minimum likelihood exists) and to evaluate the information content of the data. It is not uncommon for indices to contain insufficient information to estimate the parameters of a stock assessment model. Indices may also be conflicting and fitting therefore involves weighting averages of contradictory trends. This generally produces parameter estimates intermediate to those obtained from the data sets individually. Likelihood profiles on the average recruitment R_0 by data component were plotted to evaluate the information in each series in relation to the estimated parameters.

6.4 Stock assessment strategy

The development of a stock assessment model is comprised of the model processes, data and statistical methods for comparing data to predictions. Systematic misfit to data or conflict between data within an assessment model should be considered as a diagnostic of model misspecification.

Unacceptable model fit (i.e. model predictions do not match the data) can be detected by either the magnitude of the residuals being larger than implied by the observation error, or trends in residuals indicating systematic misfit. Data conflicts occur when different data series, given the model structure, provide different information about important aspects of the dynamics. Unacceptable model misfit or conflict between data can be dealt with by either data weighting or model process changes/flexible model parametrization.

Because it is difficult to determine the underlying cause of the model misfit and conflict, we often assume that some data are more reliable than other data for determining particular aspects of the population dynamics (Francis 2011). Our goal is to create a dynamic model of all the available data that fits the data well and is internally consistent. Internal consistency implies all data are fit as well as their observational errors and trends in residuals are minimized. Important aspects of the dynamics (scale, trend and relative scale) should be derived from the most trusted data sources.

Our modelling approach can be summarized as the following steps:

1) Selection of the data and estimation of the true sampling error;
2) Development of the initial model with original sampling error;
3) Determine if CPUE indices have information on scale and prioritize data;
4) Run stock assessment model;
5) Apply model diagnostics;
6) Modify or add additional process based on diagnostics and complete steps 4 to 6 again until internally consistent model is achieved;
7) Re-weight the data as needed.

The models selected for presentation in this WP used: the CPUE series recommended by the WG (JPN early and JPN late); the best practice approach for weighting size frequency data to ensure that the data don't overwhelm the abundance indices; sigma r of 0.3; initial catch fixed at 40,000 mt, and three combinations for the parameters of the LFSR (s_{Frac} = 0.467 and Beta=1; s_{Frac} = 0.391 and Beta=2; and s_{Frac} = 0.378 and Beta=3). This gave a total of three model runs.

In order to examine the effects of assuming an alternative stock recruitment relationship, we conducted a sensitivity analysis using the Beverton-Holt stock recruitment relationship available in SS. The steepness parameter h was fixed at 0.67, based on Kai and Fujinami (2017).

7 Results

In this section we focus on providing information to assist the WG on selecting the best case model. Here we present key results and diagnostics for all three models developed.

7.1 Model convergence

All estimated parameters in the three models were within the set bounds, and the final gradients of the models indicated that the models had converged onto a local or global minimum.

Convergence to a global minimum was examined by randomly perturbing the starting values of all parameters by 10 percent and by randomly assigning the estimated phase. Improved fit would confirm that the models had not converged to the global solution. There is no evidence of substantial differences in the scaling parameter (R_0) and total likelihood showing a better fit in all three models. Based on these results, it is concluded that the models are relatively stable with no evidence of lack of convergence to the global minimum.

7.2 Model fit diagnostics

The performance of the models was assessed by comparing input data with predictions for two data types: abundance indices and size compositions. Abundance indices provide direct information about stock trends and composition data inform about strong and weak year classes and the shape of selectivity curves (Francis 2011).

The model fits to the CPUE indices by fishery are provided in Figures 6 to 8 and Table 4. The fit to the CPUE indices were summarized into two groups: (1) those in which indices contributed to the total likelihood (S5_JPN_EARLY and S6_JPN_LATE), and (2) those in which indices did not contribute to the total likelihood (S1_HW_DP; S3_TAIW_LG; S9_SPC; S10_MEX). Results showed that both abundance indices S5_JPN_EARLY and S6_JPN_LATE had RMSE < 0.2 and SDNR values < 1, which indicate that the models fit those CPUE indices well. However, all the other indices had values for RMSE > 0.2 and SDNR > 1, which indicate that those indices were not consistent with the data included in the model.

The models fit the length modes in data aggregated by fishery and season fairly well given the estimated effective sample sizes (effN) (Figures 9 to 11), and the results of the estimated selectivity patterns were consistent with the assumed selectivity patterns (Figures 12 to 14).

Figure 15 presents the results of the likelihood profiling on log (R_0) for each data component for the model using s_{Frac} = 0.391 and Beta = 2 (the other two models showed very similar results, and are available upon request). Detailed information on the changes in negative log-likelihoods (NLL) among the various fisheries' data are shown in Tables 6 and 7. Changes in NLL each data component indicated how informative that data component was to the overall estimated model fit. Ideally, relative abundance indices should be the primary sources of information on the population scale in a model (Francis, 2011). The changes in NLL of abundance indices showed a reasonably concave shape and the minimum value (0) was close to that of total likelihood $\log(R_0)$ = 11.3.

S5_JPN_EARLY and S6_JPN_LATE showed the largest changes in NLL across values of $\log(R_0)$ among abundance indices (Table 5). The changes in NLL was also high for S3_TAIW_LG, although this index was not included in the total likelihood of the model. S5_JPN-Early showed the largest change in likelihood across values of R_0, while S10_MEX showed the lowest change in likelihood across values of R_0 among all indices. The MLE estimate for log(R_0) of S5_JPN_EARLY and S6_JPN_LATE matched a local minimum of 11.3 observed in the fleet combined likelihood profile for index data.

Overall, the changes in log-likelihoods among the nine length composition data sources were smaller than those from the abundance index, over the range of log(R_0) values (Table 6). Two out of the nine fleets (JPN kinkai shallow and JPN enyo-deep) had minimum relative negative log-likelihoods that occurred at 11.3.

There was a significant level of agreement between the length composition data and the abundance indices based on log(R_0) likelihood profiles. In other words, the generalized-size composition data did not stop the model from fitting abundance data.

7.3 Biomass

Overall, all three models showed similar trends of SSB and management reference points over time. Estimates of SSB declined from 1971 to 1991, followed by an increase between 1995 and 2007 (Figure 16). More recently, SSB slightly declined from 2009-2013, followed by an increase in the final two years (2014-2015). Recruitment (age-1 fish) estimates varied around a mean of 37,000 (in 1000's) (Figure 17). Over the course of the modelled time series, estimated fishing mortality increased abruptly in the late 1970s and early 1980's with a peak around 1989, in response to higher catches. For the last two decades F showed a declining trend, with recent values being close to those observed in the beginning of the time series (Figure 18). SS provides estimates of the MSY-related quantities, these and other quantities of interest for management for all three models are provided in Table 7.

Degrees of stock depletion and overfishing for the three stock assessment models were illustrated using the "Kobe plot"(Figure 19). Compared to MSY-based reference points, the current spawning biomass (average for 2013-2015) was well above SB_{MSY}, and the current fishing mortality (average in 2013-2015) was well below F_{MSY}. The historical trajectories of stock status revealed that North Pacific blue shark had experienced some levels of depletion and overfishing in previous years showing that the trajectories moved through the orange (overfishing) zone in the Kobe plots. However, in the last two decades, the stock condition returned into the Kobe green zone (not overfished, no overfishing). This stock assessment defaulted to evaluating stock status according to SB_{MSY} and F_{MSY}, but management bodies have not yet set biological reference points for this stock.

The ASPM produces similar estimates of abundance to the fully integrated model using s_{Frac} = 0.391 and Beta = 2 suggesting that there is information about absolute abundance in the indices of relative abundance and how it is depleted by the catch (Figure 20).

8 Sensitivity analysis

For the sensitivity run, which used the Beverton-Holt stock recruitment relationship, comparisons of spawning stock biomass and fishing intensity trajectories were completed, as well as a Kobe plot was produced (Figures 21). When compared to the models using the LFSR, the sensitivity run produced similar trends in SB and fishing mortality over time. However, the sensitivity run produced a slightly more optimistic stock status in the terminal stock assessment year (Figure 21).

9 Discussion

All of the models are expected to describe the status and trends of blue shark in the North Pacific. The models exhibited little to no conflict in the R_0 profiles. The gradients of likelihood resulting from size-composition data is low, and therefore the CPUE indices were influential in driving the model in the fitting process. As a result, the fits to the indices and size composition data were acceptable.

10 Literature cited

Beverton, R., Holt, S., 1957. On the dynamics of exploited fish populations. Minist. Agric. Fish. Food Ser. 2 Fyshery Investig. 19.

Breen, P.A., Kim, S.W., Andrew, N.L., 2003. A length-based Bayesian stock assessment model for the New Zealand abalone Haliotis iris. Mar. Freshw. Res. 54, 619–634.

Carvalho, F., 2016. Standardized CPUE for Blue Shark (Prionace glauca) Caught by the Longline Fisheries Based in Hawaii (1994 – 2015) (No. ISC/16/SHARKWG-1/16). NOAA - Pacific Islands Fisheries Science Center, Busan, South Korea.

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

Carvalho, F., Sippel, T., 2016. Direct Estimates of Gear Selectivity for the North Pacific Blue Shark Using Catch-at-Length Data: implications for stock assessment (No. ISC/16/SHARKWG-1/13). NOAA - Pacific Islands Fisheries Science Center, Busan, South Korea.

Castillo-Geniz, L., Godinez-Padilla, J., Gonzalez-Ania, L., Haro-Avalos, H., Mondragon-Sanchez, L., Tovar-Avila, J., 2017. Size And Sex of the Blue Sharks Caught by the Mexican Longline Industrial Fleets Recorded by on board Observers in the Pacific 2006-2015 (No. ISC/17/SHARKWG-1/01). Instituto Nacional de Pesca (INAPESCA), La Jolla, CA.

Fernandez-Mendez, J., Gonzalez-Ania, L., Castillo-Geniz, L., 2016. Standardized catch rates for blue shark (Prionace glauca) in the 2006-2015 Mexican Pacific longline fishery based upon a shark scientific observer program (No. ISC/16/SHARKWG-1/25). Instituto Nacional de Pesca (INAPESCA), Busan, South Korea.

Francis, R.C., 2014. Replacing the multinomial in stock assessment models: A first step. Fish. Res. 151, 70–84.

Francis, R.C., 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68, 1124–1138.

Fujinami, Y., Semba, Y., Ijima, H., Tanaka, S., 2016a. Age and growth estimatuion of the blue shark, Prionace glauca, in the western North Pacific Ocean. (No. ISC/16/SHARKWG-1/02). National Research Institute of Far Seas Fisheries.

Fujinami, Y., Semba, Y., Okamoto, H., Ohshimo, S., Tanaka, S., 2016b. Reproductive biology of the blue shark, Prionace glauca, in the western North Pacific Ocean. (No. ISC/16/SHARKWG-1/01). National Research Institute of Far Seas Fisheries.

Hiraoka, Y., Kanaiwa, M., Yokawa, K., 2013a. Re-estimation of abundance indices and catch amount for blue shark in the North Pacific ISC/13/SHARKWG-1/03.

Hiraoka, Y., Kanaiwa, M., Yokawa, K., 2013b. Summary of estimation process of abundance indices for blue shark in the North Pacific ISC/13/SHARKWG-2/02.

Hiraoka, Y., Taguchi, M., Kanaiwa, M., Yokawa, K., 2011. The operation pattern of Japanese tuna longline fishery with the information for prefecture of vessels register and reporting rate in the North Pacific Ocean, 1994-2010 (No. ISC/11/SHARKWG-2/09). National Research Institute of Far Seas Fisheries.

Kai, M., 2016. Update of Japanese catches for blue shark caught by Japanese offshore and distant water longliner in the North Pacific (No. ISC/16/SHARKWG-1/11). National Research Institute of Far Seas Fisheries.

Kai, M., Fujinami, Y., 2017. Stock recruitment relationships of the North Pacific blue shark (No. ISC/17/SHARKWG-1/02). National Research Institute of Far Seas Fisheries.

Kai, M., Shiozaki, K., 2016. Update of Japanese abundance indices for blue shark caught by Japanese offshore and distant water shallow-set longliner in the North Pacific (No. ISC/16/SHARKWG-1/10). National Research Institute of Far Seas Fisheries.

Kai, M., Shiozaki, K., Ohshimo, S., Yokawa, K., Takahashi, N., Kanaiwa, M., 2014. Update of Japanese abundance indices and catch for blue shark Prionace glauca in the North Pacific. (No. ISC/14/SHARKWG-1/2). National Research Institute of Far Seas Fisheries.

Kai, M., Yano, T., 2016. Update of catches for North Pacific blue shark caught by Japanese coastal fisheries (No. ISC/16/SHARKWG-1/12). National Research Institute of Far Seas Fisheries.

Kim, D.N., Kwon, Y., Lee, S., Cho, H., Ku, J., Lee, M., 2016. Catch and size data of blue shark by Korean tuna longline fishery in the North Pacific Ocean (No. ISC/16/SHARKWG-1/23). Korean National Institute of Fisheries Science, Busan, South Korea.

Kimoto, A., Yano, T., Yokawa, K., 2012. Historical catch amount of blue shark caught by the Japanese coastal fisheries. Int. Sci. Comm. Tuna Tuna- Species North Pac. ISC/12/SharkWG-1/11, 1–8.

King, J., Surry, A., 2016. Blue Shark (Prionace glauca) bycatch statistics in Canadian Fisheries (No. ISC/16/SHARKWG-1/19). Canada Department of Fisheries and Oceans, Busan, South Korea.

Kohin, S., Sippel, T., Carvalho, F., 2016. Catch and Size of Blue Sharks Caught in U.S. Fisheries in the North Pacific (No. ISC/16/SHARKWG-1/15). NOAA - Southwest Fisheries Science Center, Busan, South Korea.

Kwon, Y., Kim, D.N., Lee, S., Cho, H., Ku, J., Lee, M., 2017. Fishery data and catch estimation of blue shark by Korean tuna longline fishery in the North Pacific Ocean (No. ISC/17/SHARKWG-1/6). Korean National Institute of Fisheries Science, La Jolla, CA.

Liu, K.M., Su, K., Chin, C.P., 2016a. Catch, size and distribution pattern of blue sharks by Taiwanese small-scale longline fleets in the North Pacific in 2001-2015 (No. ISC/16/SHARKWG-1/20). National Taiwan Ocean University, Busan, South Korea.

Liu, K.M., Su, K., Chin, C.P., 2016b. Size and spatial distribution of the blue shark, Prionace glauca, caught by Taiwanese large-scale longline fishery in the North Pacific Ocean (No. ISC/16/SHARKWG-1/21). National Taiwan Ocean University, Busan, South Korea.

Mangel, M., Brodziak, J., DiNardo, G., 2010. Reproductive ecology and scientific inference of steepness: a fundamental metric of population dynamics and strategic fisheries management. Fish Fish. 11, 89–104.

Maunder, M.N., Piner, K.R., 2015. Contemporary fisheries stock assessment: many issues still remain. ICES J. Mar. Sci. 72, 7–18.

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

Methot, R.D., Wetzel, C.R., 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99.

Minte-Vera, C.V., Maunder, M.N., Aires-da-Silva, A.M., Satoh, K., Uosaki, K., 2017. Get the biology right, or use size-composition data at your own risk. Fish. Res. doi:http://dx.doi.org/10.1016/j.fishres.2017.01.014

Nakano, H., 1994. Age, reproduction and migration of blue shark (Prionace glauca) in the north Pacific ocean. Bull. - Natl. Res. Inst. Far Seas Fish. (no.31) p. 141-256.

Ohshimo, S., Fujinami, Y., Shiozaki, K., Kai, M., Semba, Y., Katsumata, N., Ochi, D., Matsunaga, H., Minami, H., Yokawa, K., 2014. Distribution, body length and abundance of blue shark and shortfin mako shark in the Northwestern Pacific Ocean based on longline research vessels from 2000 to 2014 ISC/14/SHARKWG-3/04, 30.

Punt, A.E., 2016. Some insights into data weighting in integrated stock assessments. Fish. Res.

Rice, J., Harley, S., 2014. Standardization of blue shark catch per unit effort in the North Pacific Ocean based on SPC held longline observer data for use as an index of abundance (No. ISC/14/SharkWG-2/04). Secretariat of the Pacific Community, Keelung, Taiwan.

Rice, J., Harley, S., Maunder, M., Aires Da-Silva, A., 2014. Stock assessment of blue sharks in the North Pacific Ocean using Stock Synthesis. West. Cent. Pac. Fish. Comm. - 9th Sci. Committe Meet. WCPFC-SC9-2013/SA-WP-02.

Runcie, R., Xu, Y., Wraith, J., Kohin, S., 2014. Standardized abundance index of juvenile shortfin mako shark (Isurus oxyrinchus) based on a fishery-independent survey in the Southern California Bight (1994-2013) (No. ISC/14/SHARKWG-3/08). NOAA.

Semba, Y., Yokoi, H., 2016. Update of age and sex specific natural mortality of the blue shark (Prionace glauca) in the North Pacific Ocean (No. ISC/16/SHARKWG-1/06). National Research Institute of Far Seas Fisheries.

Sippel, T., Nasby-Lucas, N., Kohin, S., 2014. Description of the Hawaii Longline Observer Program (No. ISC/14/SHARKWG-1/05). NOAA.

Sippel, T., Semba, Y., Shiozaki, K., Carvalho, F., Castillo-Geniz, L., Tsai, W., Liu, K.M., Kwon, Y., Chen, Y., Kohin, S., 2016. Size and sex structure of blue sharks in the North Pacific Ocean (No. ISC/16/SHARKWG-1/14). NOAA - Southwest Fisheries Science Center, Busan, South Korea.

Sippel, T., Wraith, J., Kohin, S., Taylor, V., Holdsworth, J., Taguchi, M., Matsunaga, H., Yokawa, K., 2011. A summary of blue shark (Prionace glauca) and shortfin mako shark (Isurus oxyrinchus) tagging data available from the North and Southwest Pacific Ocean (ISC/11/SHARKWG-2/04 No. ISC/11/SHARKWG-2/04). NOAA.

Sosa-Nishizaki, O., Castillo-Geniz, L., 2016. Blue shark catches estimations for the Mexican Pacific (1976-2014) (No. ISC/16/SHARKWG-1/24). Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Busan, South Korea.

Stevens, J.D., Bradford, R.W., West, G.J., 2010. Satellite tagging of blue sharks (Prionace glauca) and other pelagic sharks off eastern Australia: depth behaviour, temperature experience and movements. Mar. Biol. 157, 575–591. doi:10.1007/s00227-009-1343-6

Taguchi, M., King, J.R., Wetklo, M., Withler, R.E., Yokawa, K., 2015. Population genetic structure and demographic history of Pacific blue sharks (Prionace glauca) inferred from mitochondrial DNA analysis. Mar. Freshw. Res. 66, 267–275.

Taylor, I.G., Gertseva, V., Methot, R.D., Maunder, M.N., 2013. A stock–recruitment relationship based on pre-recruit survival, illustrated with application to spiny dogfish shark. Fish. Res. 142, 15–21.

Teo, S., Sippel, T., Wells, D., Kohin, S., 2012. Preliminary time series for north Pacific blue and shortfin mako sharks from the U.S. West Coast drift gillnet fishery (No. ISC/12/SHARKWG-1/03). NOAA.

Tsai, W., Liu, K.M., 2016. Catch estimate and CPUE standardization of the blue shark based on observers' records of Taiwanese large-scale tuna longline fisheries in the North Pacific Ocean (No. ISC/16/SHARKWG-1/22). National Taiwan Ocean University, Busan, South Korea.

Wang, S.P., Chen, Y.R., Maunder, N.M., Nishida, T., 2009. Preliminary application of an age–structured assessment model to swordfish (*Xiphias gladius*) in the Indian Ocean. IOTC-WPB-2009-11

Walsh, W., Teo, S., 2012. Catch Statistics, Length Data and Standardized CPUE for Blue Shark Prionace glauca taken by Longline Fisheries based in Hawaii and California. (No. ISC/12/SHARKWG-1/02).

Walters, J., Sharma, R., Cass-Calay, S., Brown, C., 2016. Scaling natural mortality rate as a function of length or weight with an application to yellowfin tuna (No. SCRS/2016/116).

Yamamoto, A., Shibano, A., Semba, Y., Hiraoka, Y., Kanaiwa, M., Kobayashi, M., 2016. Conversion factors between specific body parts for blue shark in Pacific Ocean (No. ISC/16/SHARKWG-1/03). Bioresources Institute, Tokyo University of Agriculture.

Yokawa, K., 2012. Blue sharks caught by Japanese large mesh drift net fishery in the north Pacific in 1981 - 1993 ISC/12/SHARKWG-2/10.

Table 1. Key life history parameters and model structures used in the North Pacific blue shark stock assessment, including values, pertinent comments, and references.

Parameter	**Value**	**Comments**	**Source**
Gender	2	Two sex model	ISC (2014)
Natural mortality		Age-specific natural mortality Tmax=24	Semba and Yokoi (2016)
Reference age (a1)	1	Fixed parameter	Fujinami et al. (2016a)
Maximum age (a2)	20		
Female at first age	4		
Length at a1 (L1)	64.4 (Female)		
	68.2 (Male)		
Length at a2 (L2)	244.6 (Female)		
	261.3 (Male)		
Growth rate (K)	0.147 (Female)		
	0.117 (Male)		
CV of L1 (CV=f(LAA))	0.25 (Female); 0.25 (Male);		ISC(2014)
CV of L2	0.1 (Female); 0.1 (Male);		
Weight-at-length	$W=5.388 \times 10^{-6} L^{3.102}$ (Female);		Yamamoto et al. (2016)
	$W=3.293 \times 10^{-6} L^{3.225}$ (Male)		
Length-at-50% Maturity	156.6 (Female)		Fujinami et al. (2016b)
Slope of maturity ogive	- 0.16 (Female)		
Fecundity (Litter size; (4)eggs=a+b*L)	Proportional to body length	Model structure	
Slope of fecundity (b)	0.46	Fixed parameter	
Intercept of fecundity (a)	-45.54		
Spawning season	1	Model structure (No season)	Fujinami et al. (2016b)
Spawner-recruit relationship	Low Fecundity SR-relationships (LFSR)	Model structure	Kai and Fujinami (2017)

Spawner-recruit steepness (*h*)	0.67 (reference)	Model structure	Kai and Fujinami (2017)
LFSR (Sfrac)	0.378 0.391 0.467		
LFSR (Beta)	1 2 3		
Log of Recruitment at virgin biomass log(R0)	11.1358 (Initial value)	Estimated	ISC (2014)
Recruitment variability (σ_R)	0.3	Fixed parameter	
Initial age structure	5 yrs (1985-1989)	Estimated	
Main recruitment deviations	1990-2013	Fixed	
Bias adjustment	1990-2013		
F ballpark for tuning early phases	0.2		
F ballpark year	2013		
F-Method	3 (hybrid)	Model structure	
Initial-F	0.315485 (Initial value) only Kinkai shallow (F4)	Estimated	

Table 2. Descriptions of the fisheries included in the models for the North Pacific blue shark stock assessment.

Fishery number	Reference Code	Fishing Countries	Gear Types	Units	Source
F1	MEX	Mexico	Mexican Pacific longline	Biomass	Sosa-Nishizak and Castillo-Geniz (2016)
F2	CAN	Canada	Troll, gillnet, seine fishery, foreign and joint-venture fisheries	Biomass	King and Surry (2016)
F3	CHINA	China		Biomass	
F4	JPN_KK_SH	Japan	Offshore shallow-set longline	Biomass	Kai (2016); Hiraoka et al.(2013)
F5	JPN_KK_DP	Japan	Offshore deep-set longline	Biomass	Kai (2016); Hiraoka et al.(2013)
F6	JPN_ENY_SHL	Japan	Distant water shallow-set longline	Biomass	Kai (2016); Hiraoka et al.(2013)
F7	PN_ENY_DP	Japan	Distant water deep-set longline	Biomass	Kai (2016); Hiraoka et al.(2013)
F8	JPN_LG_MESH	Japan	High-sea large-mesh driftnet	Biomass	Yokawa et al. (2012)
F9	JPN_CST_Oth	Japan	Coastal longline	Biomass	Kai and Yano (2016); Kimoto et al. (2011)
F10	JPN_SM_MESH	Japan	Coastal driftnet	Biomass	Kai and Yano (2016)
F11	IATTC	RFMO	Offshore longline, coastal longline, gillnet, harpoon, and others	Biomass	Alexandre Da Silva pers. comm., Nov 29, 2016
F12	KOREA	Korea	Tuna longline, observer data	Biomass	Kim et al. (2016)
F13	NON_ISC	Various flags	Longline	Biomass	
F14	USA_GILL	USA (American Samoa)	Gill net	Biomass	
F15	USA_SPORT	USA (American Samoa)	Recreational fishing	Biomass	
F16	USA_Longline	USA (Hawaii)	Longline	Biomass	Carvalho, F. (2016)

F17	TAIW_LG	Taiwan	Large-scale longline	Biomass	Liu et al. (2016)
F18	TAIW_SM	Taiwan	Small-scale longline	Biomass	Liu et al. (2016)

Table 3. Fishery-specific selectivity assumptions used in the structures used in the North Pacific blue shark stock assessment. The selectivity curves for fisheries lacking length composition data were assumed to be the same as (i.e., mirror gear) a closely related fisheries or to a fisheries operating in the same area.

Fishery reference	**Reference Code**	**Selectivity assumption**	**Mirror gear**
F1	MEX	Double-normal	Estimated
F2	CAN	Double-normal	F1
F3	CHINA	Double-normal	Estimated
F4	JPN_KK_SH	Double-normal	Estimated
F5	JPN_KK_DP	Double-normal	Estimated
F6	JPN_ENY_SHL	Double-normal	F4
F7	PN_ENY_DP	Double-normal	F5
F8	JPN_LG_MESH	Double-normal	Estimated
F9	JPN_CST_Oth	Double-normal	F7
F10	JPN_SM_MESH	Double-normal	Estimated
F11	IATTC	Double-normal	F1
F12	KOREA	Double-normal	F3
F13	NON_ISC	Double-normal	F3
F14	USA_GILL	Double-normal	Estimated
F15	USA_SPORT	Double-normal	F14
F16	USA_Longline	Double-normal	Estimated
F17	TAIW_LG	Double-normal	Estimated

F18	TAIW_SM	Double-normal	F17
S1	HW_DP	Double-normal	F16
S2	HW_SH	Double-normal	F16
S3	TAIW_LG	Double-normal	F17
S4	TAIW_SM	Double-normal	F18
S5	JPN_EARLY	Double-normal	F4
S6	JPN_LATE	Double-normal	F5
S7	JPN_RTV	Double-normal	F16
S8	SPC_OBS	Double-normal	F13
S9	SPC_OBS_TROPIC	Double-normal	F13
S10	Mex_LG	Double-normal	Estimated

Table 4. Input CV, root-mean-square-errors (RMSE), and standard deviations of the normalized residuals (SDNR) for the relative abundance indices used in the North Pacific blue shark stock assessment. S1_HW_DP, S3_TAIW_LG, S9_SPC, S10_MEX were not included in the total likelihood.

Model run: s_{Frac} = 0.467 and Beta=1

Reference code	n	Input CV	RMSE	SDNR	χ^2
S1_HW_DP	16	0.28	0.46	1.69	1.29
S3_TAIW_LG	12	0.10	0.67	6.91	1.33
S5_JPN_EARLY	18	0.10	0.08	0.95	1.27
S6_JPN_LATE	22	0.10	0.09	0.98	1.24
S9_SPC	17	0.14	0.39	2.70	1.28
S10_MEX	10	0.12	0.24	2.18	1.37

Model run: s_{Frac} = 0.391 and Beta=2

Reference code	n	Input CV	RMSE	SDNR	χ^2
S1_HW_DP	16	0.28	0.46	1.69	1.29
S3_TAIW_LG	12	0.10	0.67	7.01	1.33
S5_JPN_EARLY	18	0.10	0.08	0.89	1.27
S6_JPN_LATE	22	0.10	0.09	0.99	1.24
S9_SPC	17	0.14	0.39	2.85	1.28
S10_MEX	10	0.12	0.24	2.15	1.37

Model run: s_{Frac} = 0.378 and Beta=3

Reference code	n	Input CV	RMSE	SDNR	χ^2
S1_HW_DP	16	0.28	0.46	1.69	1.29
S3_TAIW_LG	12	0.10	0.67	6.99	1.33
S5_JPN_EARLY	18	0.10	0.08	0.90	1.27
S6_JPN_LATE	22	0.10	0.09	0.99	1.24
S9_SPC	17	0.14	0.39	2.86	1.28
S10_MEX	10	0.12	0.24	2.12	1.37

Table 5. Relative negative log-likelihoods of abundance index data components for the model run using a combination of Beta = 2 and Sfrac = 0.391 in the LFSR over a range of fixed levels of virgin recruitment in log-scale (log(R0)). Likelihoods are relative to the minimum negative log-likelihood for each respective data component. Colors indicate relative likelihood (red: high negative log-likelihood; green: low negative log-likelihood). S1_HW-DP, S3_TAIW_LG, S9_SPC, and S10_MEX were not included in the total likelihood.

R0	S1_HW_DP	S3_TAIW_LG	S5_JPN_EARLY	S6_JPN_LATE	S9_SPC	S10_MEX
10.8	3.192665	20	135.979	46.06176	6.0495	0.653301
10.9	2.156205	11.0447	89.681	27.5769	5.3626	1.478125
11	1.343675	7.7786	44.269	14.1025	4.183	2.505805
11.1	1.346335	4.4135	20.935	7.2711	0.867	1.435678
11.2	1.217195	1.1	4.026	1.164	0.7812	1.289124
11.3	1.186075	0.72	0	0	0.4507	0.807215
11.4	0.942685	0.0021	6.804	3.6564	0.2557	0.511986
11.5	0.664265	0	14.969	6.5903	0.1344	0.316444
11.6	0.41626	1.0491	27.604	11.5687	0.0586	0.173947
11.7	0.195376	5.312	37.161	22.7597	0.0161	0.072241
11.8	0	7.6273	45	34.0496	0	0

Table 6. Relative negative log-likelihoods of length composition data components for the model run using a combination of Beta = 2 and Sfrac = 0.391 in the LFSR over a range of fixed levels of virgin recruitment in log-scale (log(R0)). Likelihoods are relative to the minimum negative log-likelihood for each respective data component. Colors indicate relative likelihood (red: high negative log-likelihood; green: low negative log-likelihood).

R0	MEX	CHINA	JPN KK_SH	JPN KK_DP	JPN ENY_DP	JPN LG_MESH	USA GILL	USA Longline	TAIW_LG
10.8	0.592059	3.708873	25.21644	2.605106	19.83565	0.281332	8.612634	3.708873	0.25496
10.9	0.399855	2.04817	16.63077	2.309306	11.87549	0.636527	4.756198	2.04817	0.17219
11	0.249176	1.442492	8.209405	1.801332	6.072983	1.079079	3.349712	1.442492	0.107303
11.1	0.249669	0.818456	3.882263	0.373358	3.131166	0.618248	1.900593	0.818456	0.107516
11.2	0.225721	0.203988	0.746596	0.336409	0.501255	0.555138	0.473695	0.203988	0.097203
11.3	0.21995	0.133519	0	0.194086	0	0.347612	0.310055	0.133519	0.094717
11.4	0.174815	0.000389	1.261759	0.110113	1.574562	0.220477	0.000904	0.000389	0.075281
11.5	0.123184	0	2.775906	0.057877	2.837992	0.136271	0	0	0.053047
11.6	0.077193	0.194549	5.118987	0.025235	4.981849	0.074907	0.451776	0.194549	0.033242
11.7	0.036231	0.985077	6.891272	0.006933	9.801048	0.031109	2.287516	0.985077	0.015602
11.8	0	1.414434	8.344965	0	14.66284	0	3.284557	1.414434	0

Table 7. Estimates of key management quantities for the North Pacific blue shark stock assessment, under three different low fecundity stock recruitment relationships.

	Model		
	Beta = 1 Sfrac = 0.467	**Beta = 2 Sfrac = 0.391**	**Beta = 3 Sfrac = 0.378**
SSB_{1971}	363,704	311,312	285,130
SSB_{2015}	287,792	308,286	311,444
SSB_{MSY}	196,649	179,539	174,563
F_{1971}	0.119	0.137	0.147
F_{2015}	0.109	0.090	0.098
F_{MSY}	0.443	0.355	0.329
SSB_{2015}/ SSB_{MSY}	1.463	1.717	1.784
F_{2015}/ F_{MSY}	0.246	0.253	0.297

1 Figures

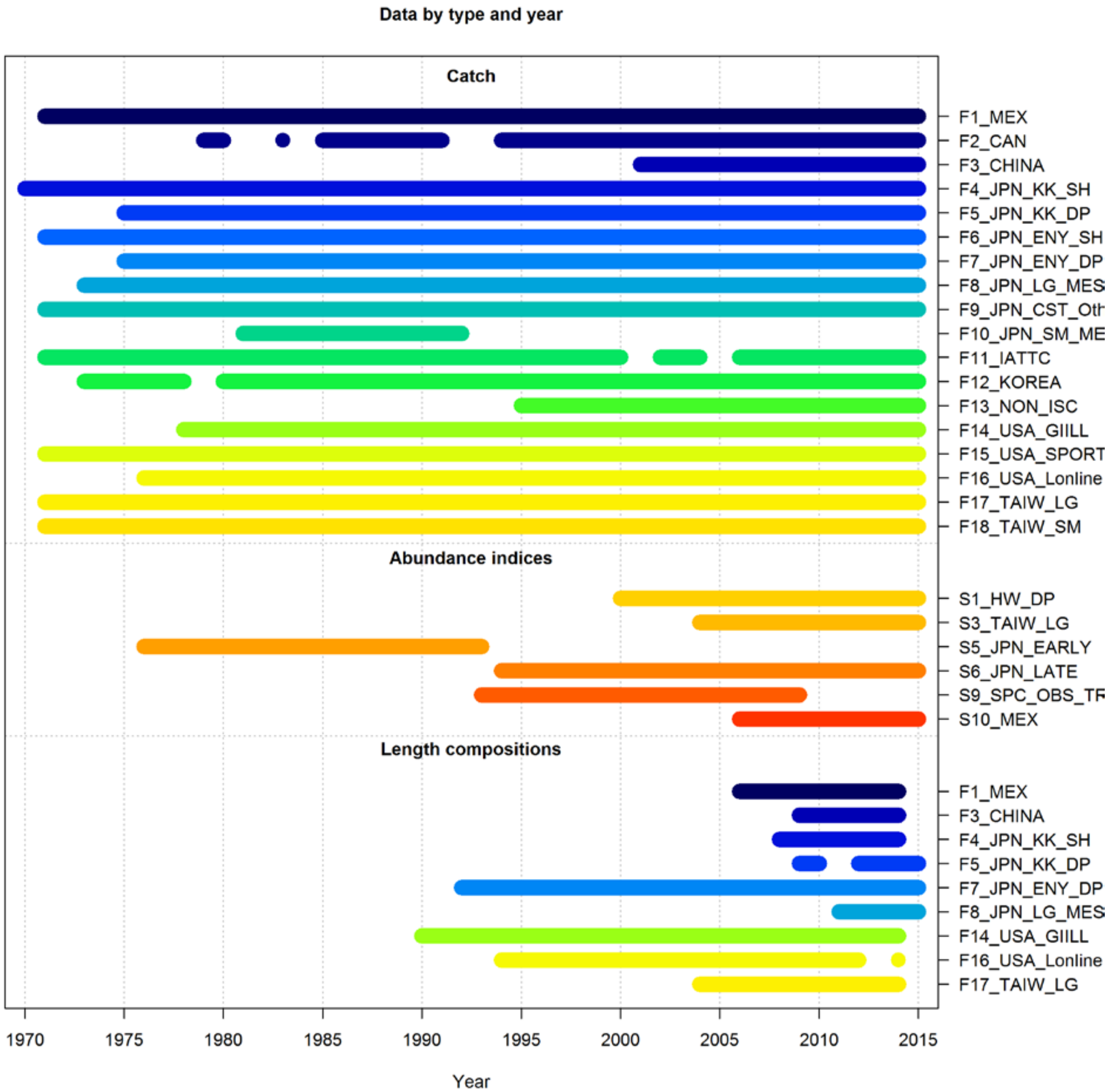


Figure 1. Temporal coverage for data included in the stock assessment of blue shark in the North Pacific Ocean.

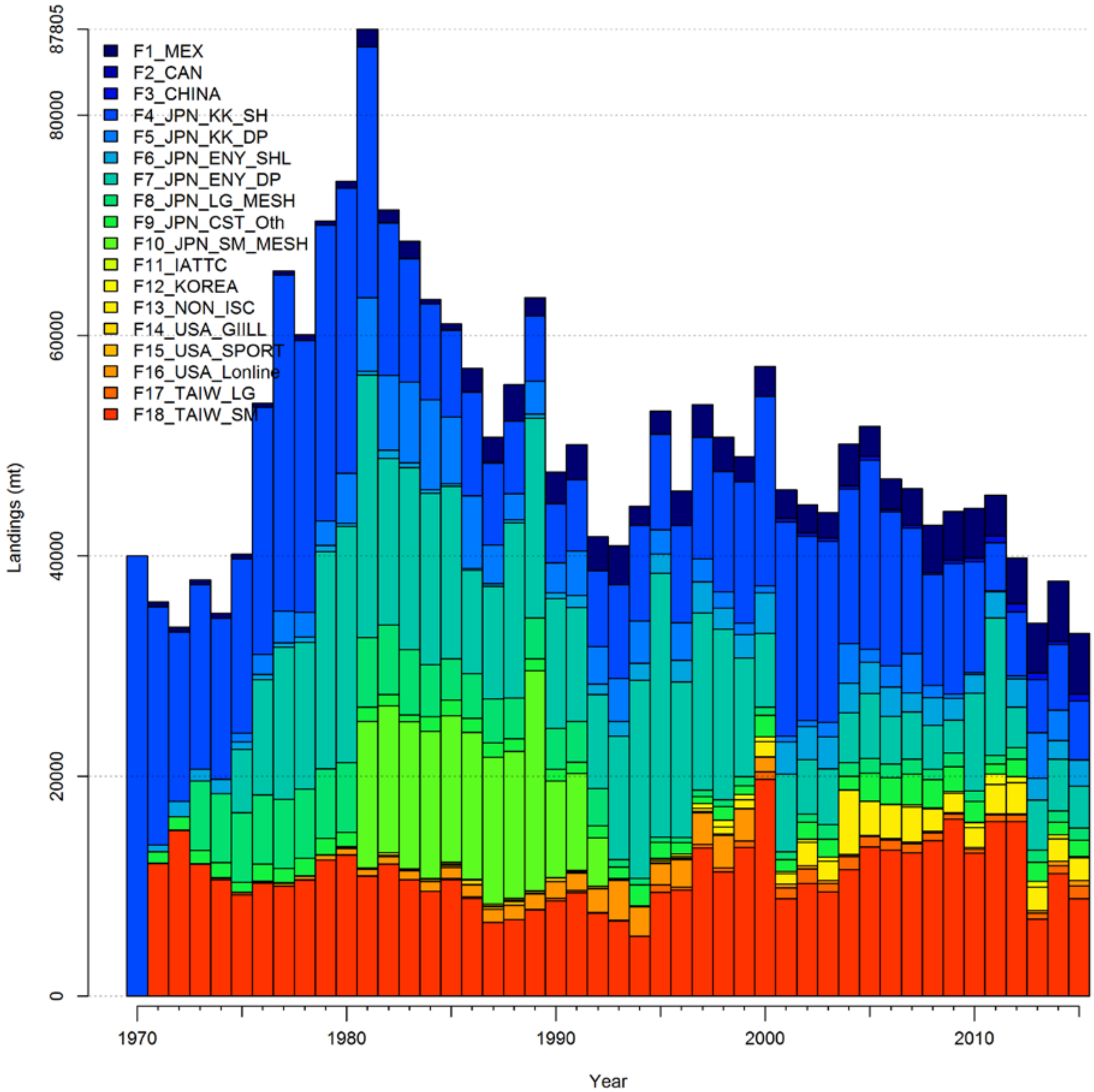


Figure 2. Assumed catches for blue shark in the North Pacific Ocean.

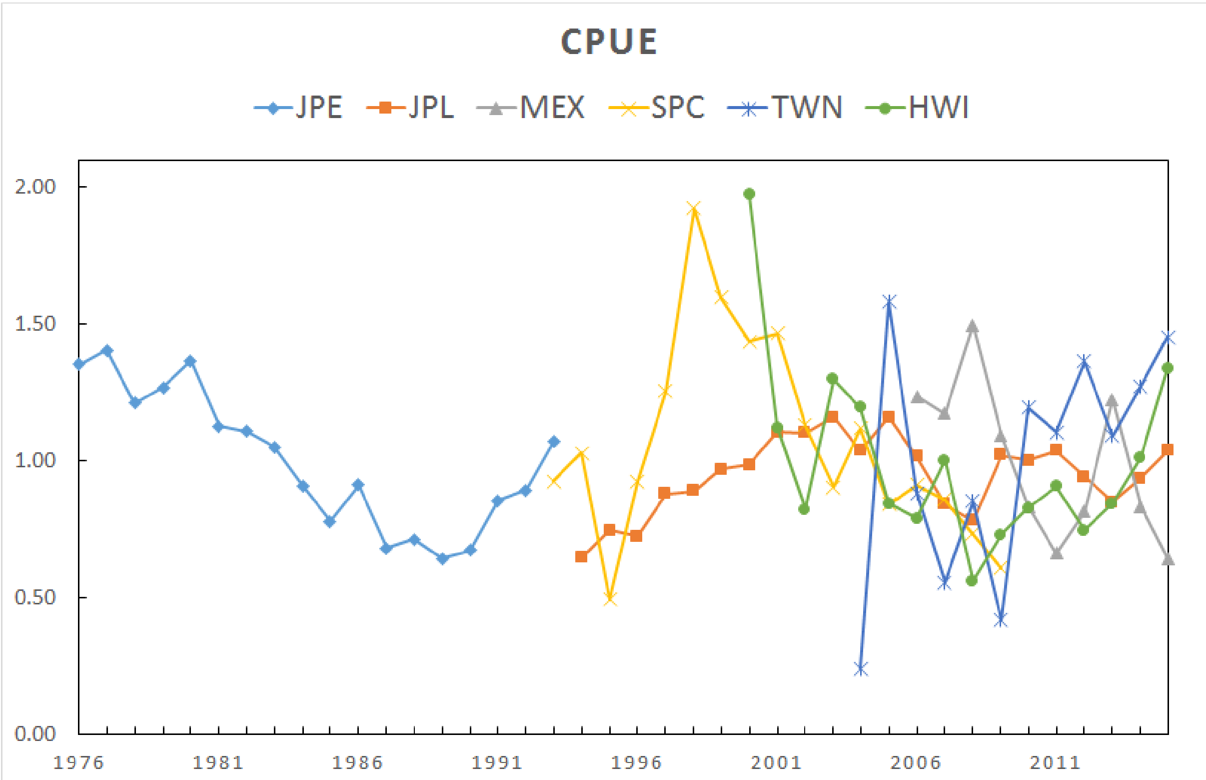


Figure 3. Yearly changes in standardized Japanese offshore shallow-set longline CPUE of North Pacific blue shark 1971-1993 (JPE) and 1994-2015 (JPL), and five standardized CPUEs (Hawaii deep-set longline: HWI, Mexico longline: MEX, SPC observed longline: SPC, Taiwan large-scale longline: TWN).

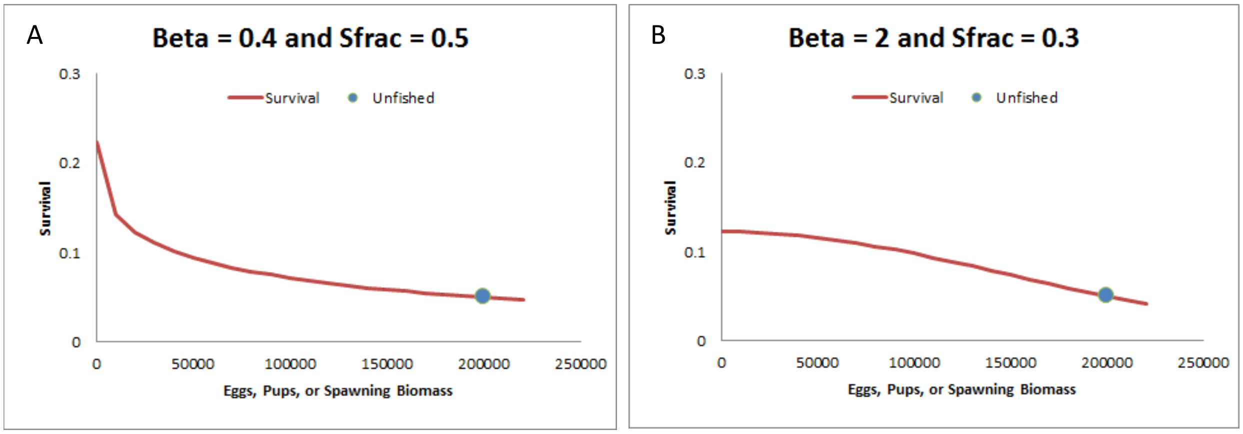


Figure 4. Examples of Pre-recruitment survival for the Low Fecundity Spawner Recruitment (LFSR) implemented in Stock Synthesis (for details see page 59 of Stock Synthesis User Manual 3.24S).

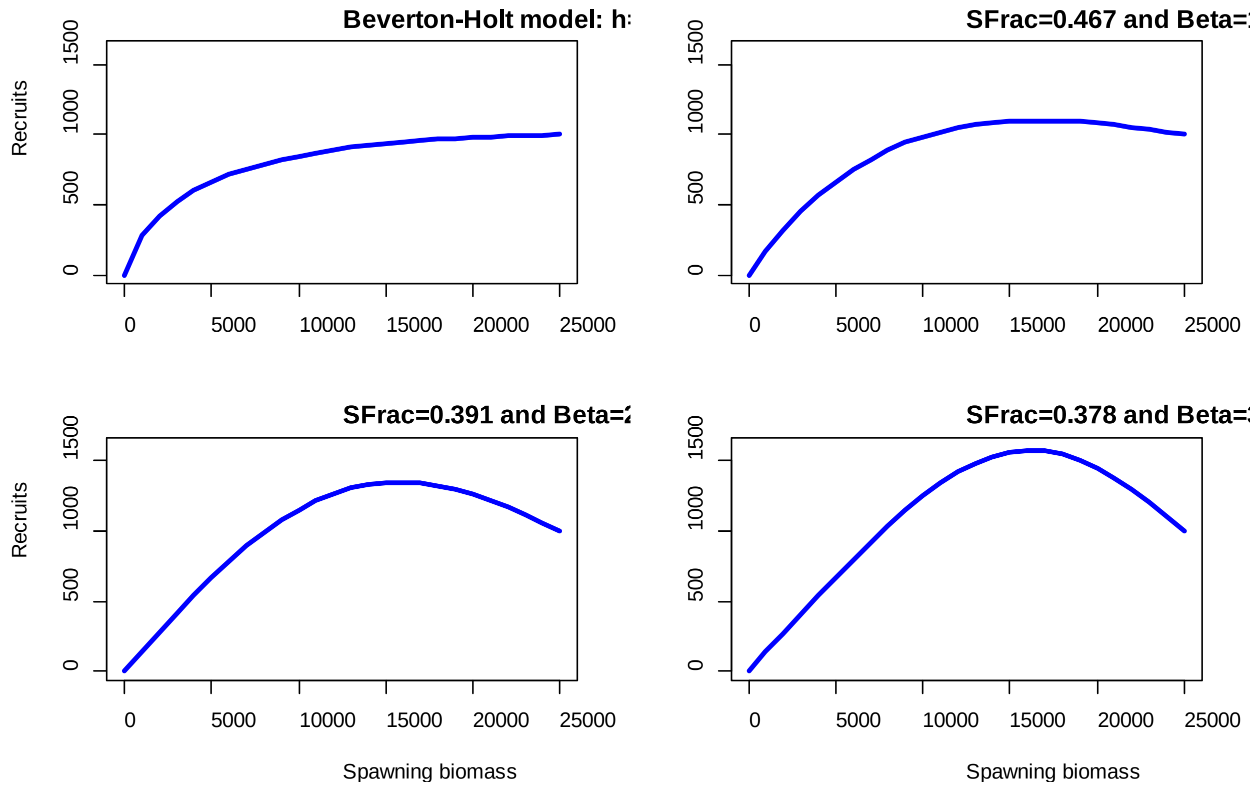


Figure 5. Illustration of the parameterizations of the stock recruitment relationships using the Beverton-Holt model for the sensitivity run (top left) and the LFSR

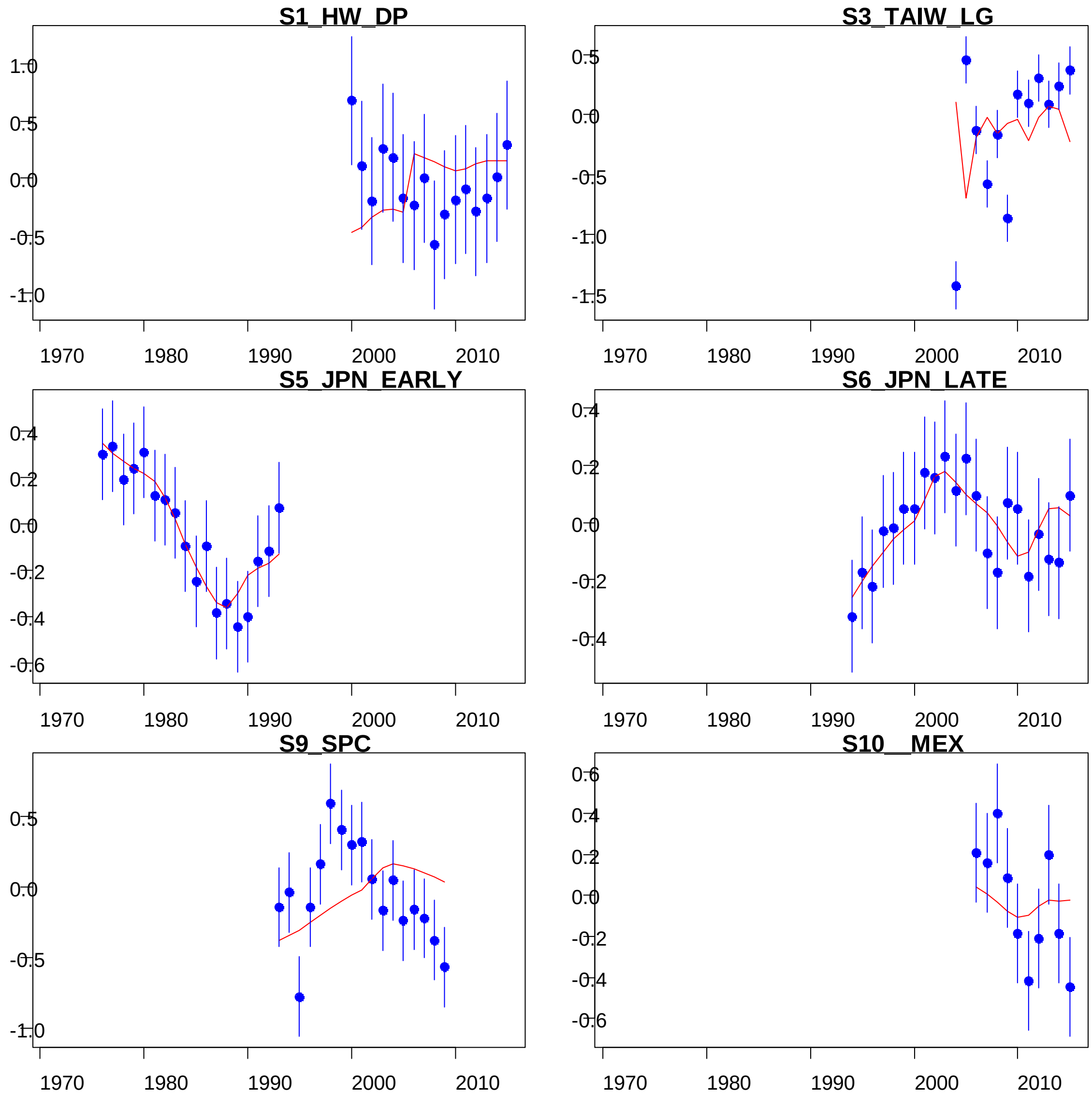


Figure 6. Model fits to the standardized catch-per-unit-effort (CPUE) (in log scale) data sets from different fisheries for the model run using a combination of Beta = 1 and Sfrac = 0.467 in the LFSR. The solid red line is the model predicted value and the solid circles are observed data values. Vertical blue lines represent the estimated confidence intervals (± 1.96 standard deviations) around the CPUE values. S1_HW_DP, S3_TAIW_LG, S9_SPC, S10_MEX were not included in the total likelihood.

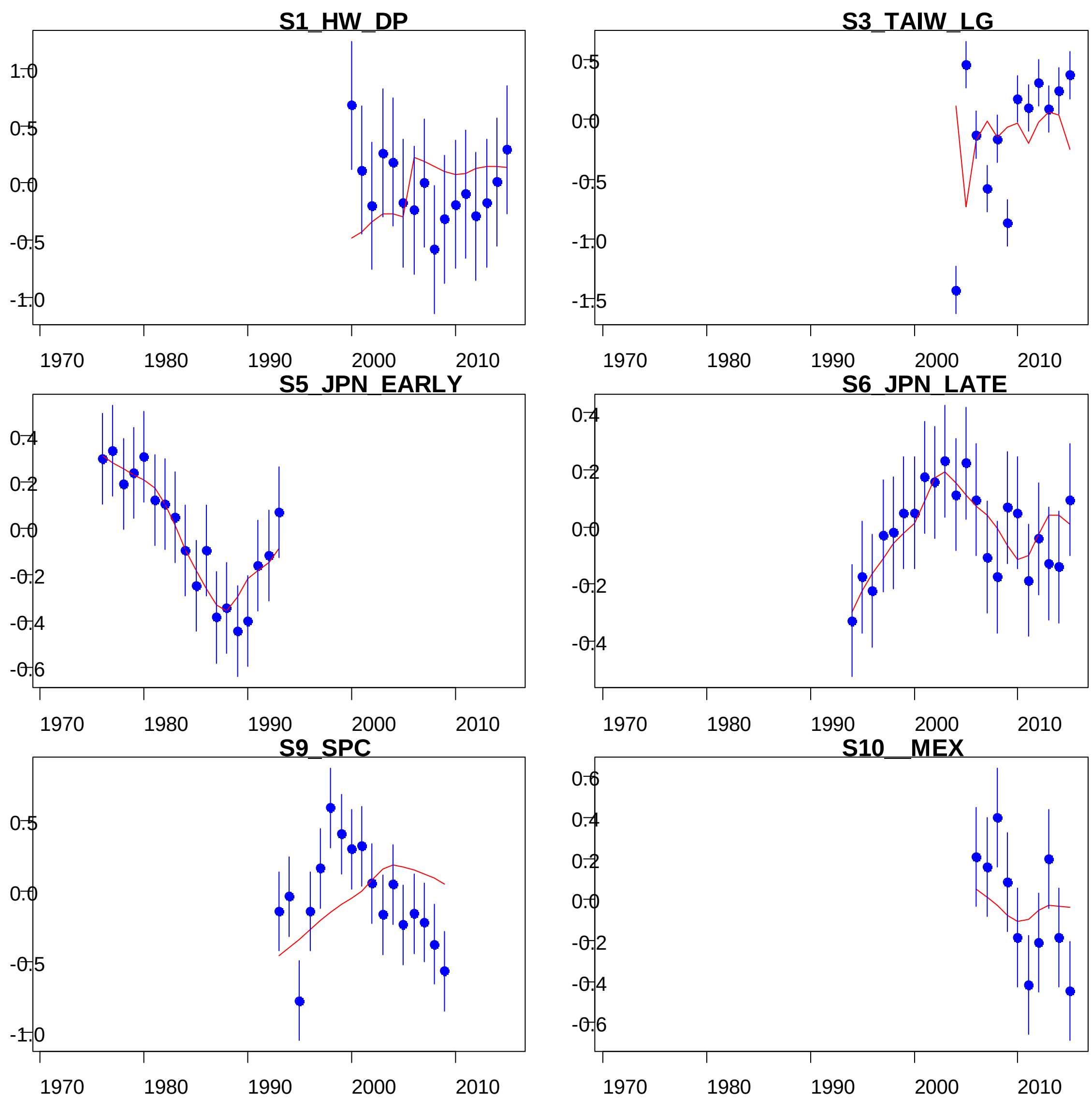


Figure 7. Model fits to the standardized catch-per-unit-effort (CPUE) (in log scale) data sets from different fisheries for the model run using a combination of Beta = 2 and Sfrac = 0.391 in the LFSR. The solid red line is the model predicted value and the solid circles are observed data values. Vertical blue lines represent the estimated confidence intervals (± 1.96 standard deviations) around the CPUE values. S1_HW_DP, S3_TAIW_LG, S9_SPC, S10_MEX were not included in the total likelihood.

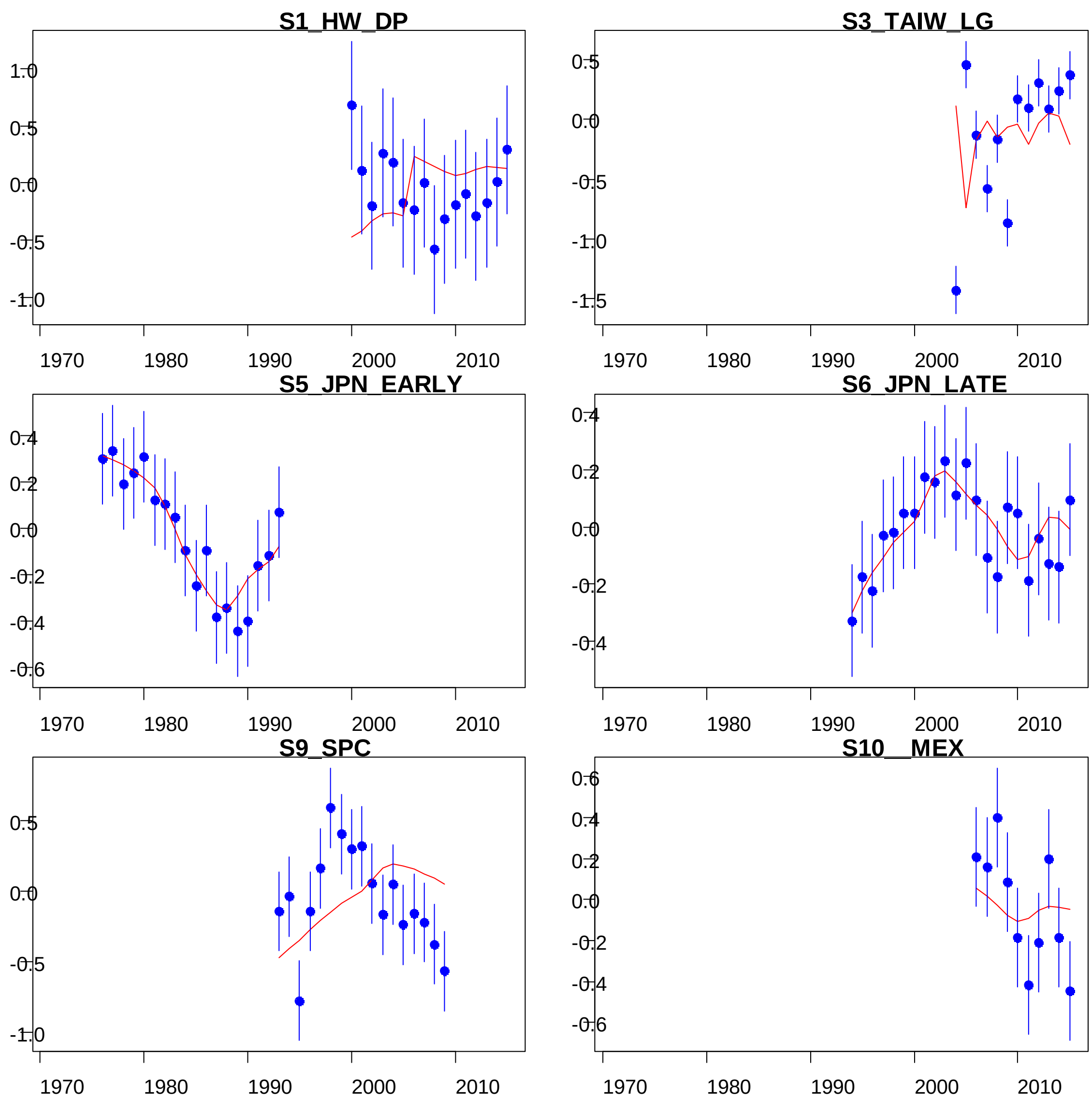


Figure 8. Model fits to the standardized catch-per-unit-effort (CPUE) (in log scale) data sets from different fisheries for the model run using a combination of Beta = 3 and Sfrac = 0.378 in the LFSR. The solid red line is the model predicted value and the solid circles are observed data values. Vertical blue lines represent the estimated confidence intervals (± 1.96 standard deviations) around the CPUE values. S1_HW_DP, S3_TAIW_LG, S9_SPC, S10_MEX were not included in the total likelihood.

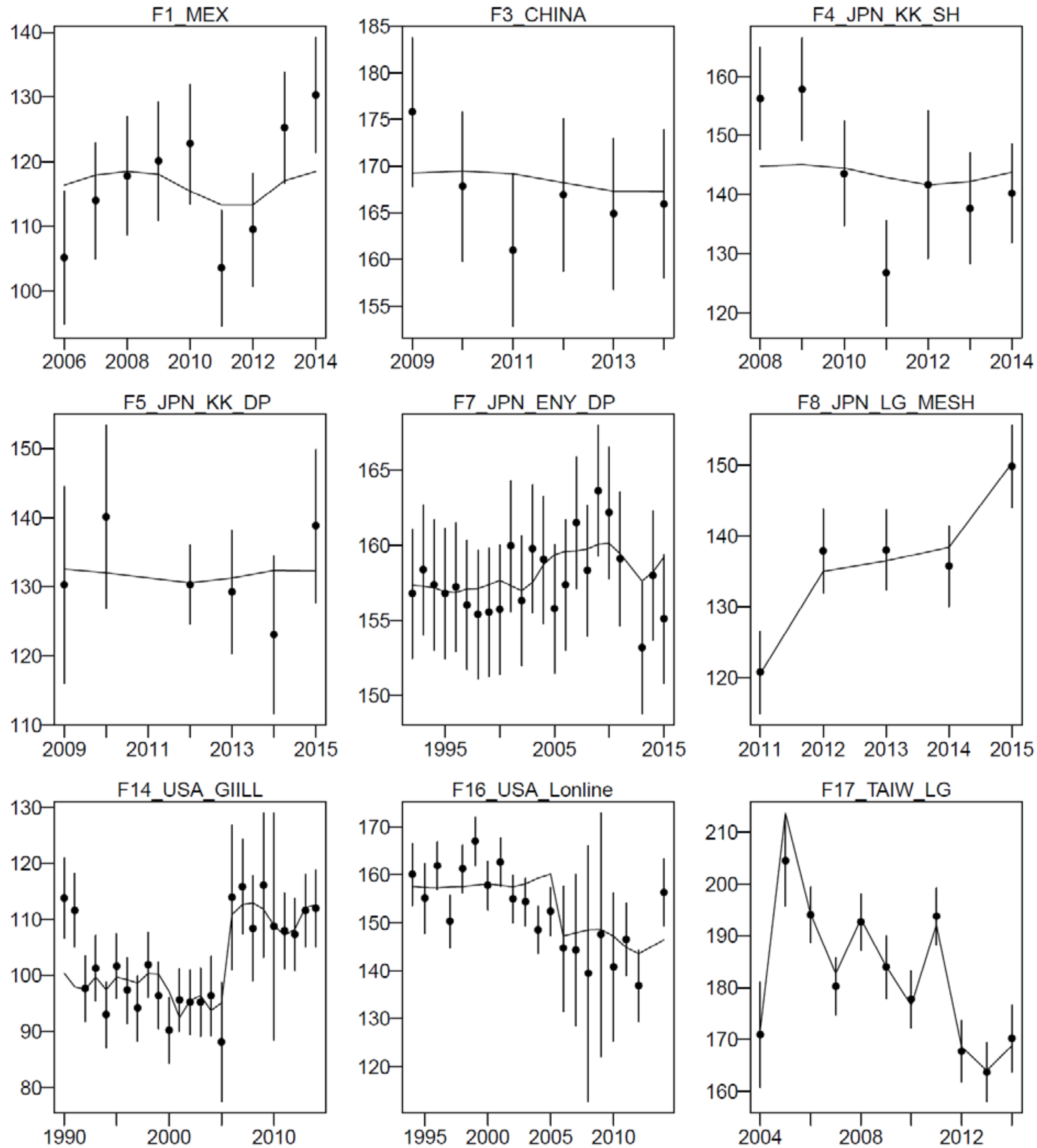


Figure 9. Model fit (black solid lines) to mean PCL (in cm) of the composition data using a combination of Beta = 1 and Sfrac = 0.467 in the LFSR. The solid circles are the observed mean length and the vertical black solid lines are 95% credible limits around mean length. All measurements were in pre-caudal length (PCL, cm).

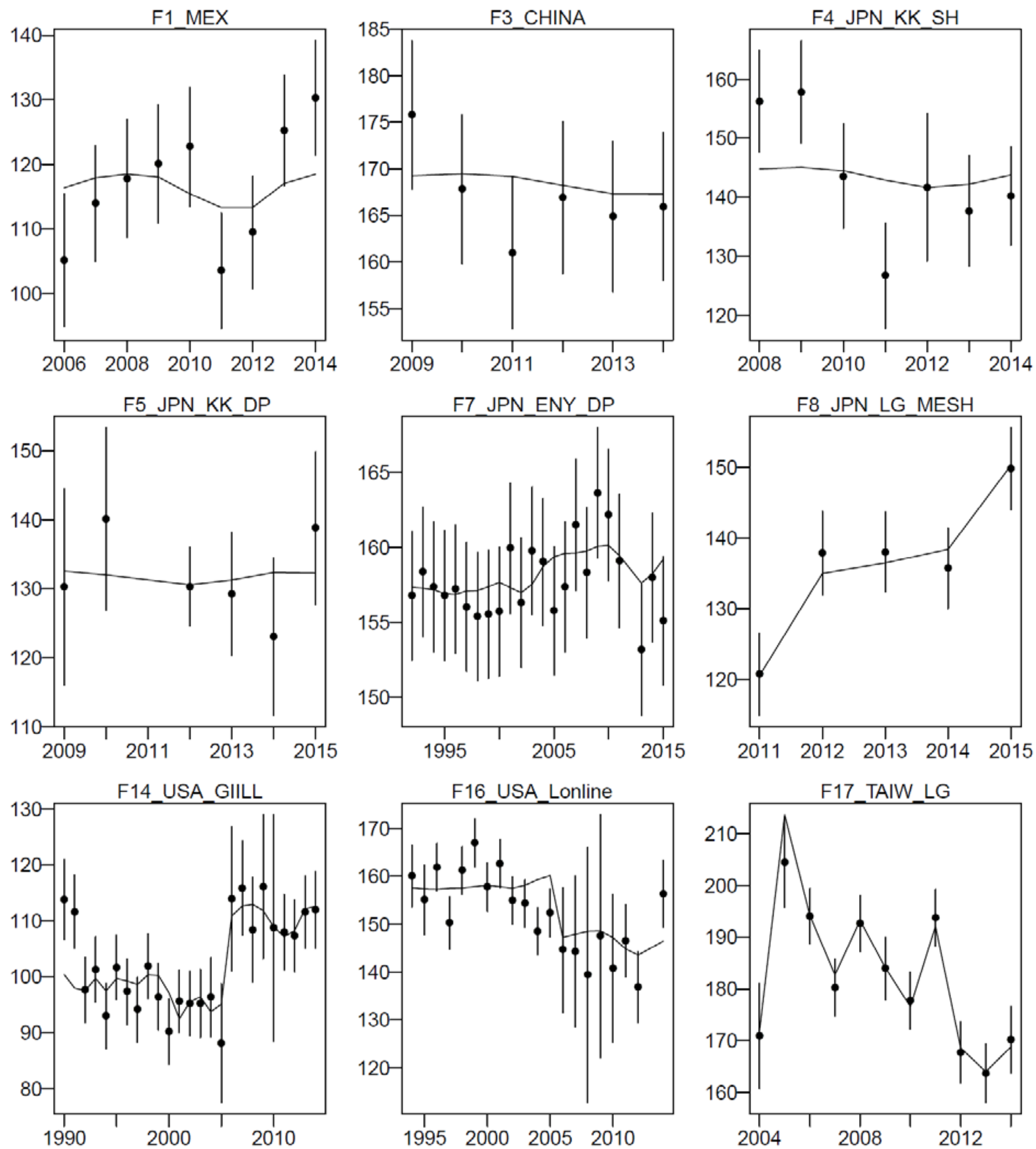


Figure 10. Model fit (black solid lines) to mean PCL (in cm) of the composition data using a combination of Beta = 2 and Sfrac = 0.391 in the LFSR. The solid circles are the observed mean length and the vertical black solid lines are 95% credible limits around mean length. All measurements were in pre-caudal length (PCL, cm).

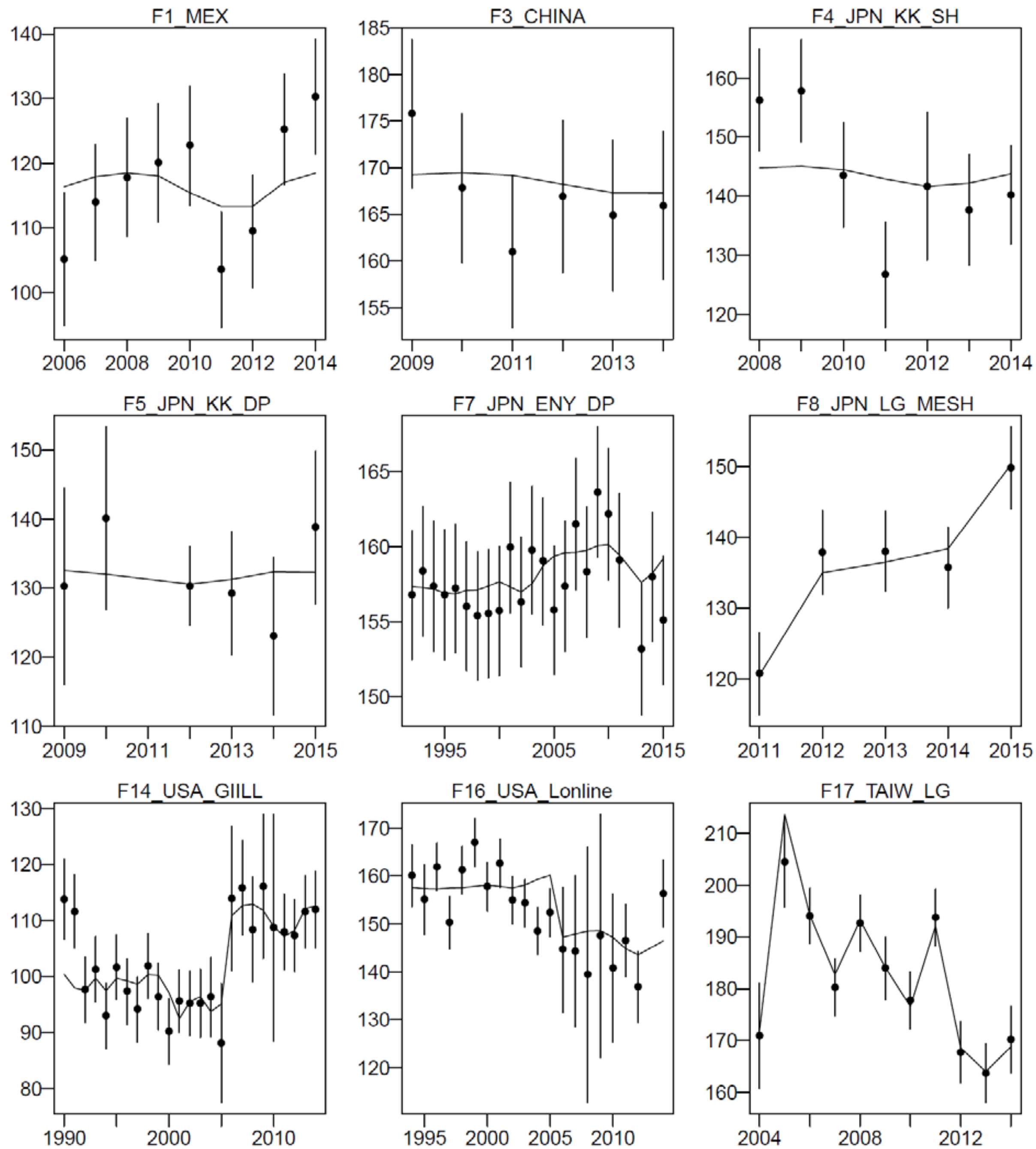


Figure 11. Model fit (black solid lines) to mean PCL (in cm) of the composition data using a combination of Beta = 3 and Sfrac = 0.378 in the LFSR. The solid circles are the observed mean length and the vertical black solid lines are 95% credible limits around mean length. All measurements were in pre-caudal length (PCL, cm).

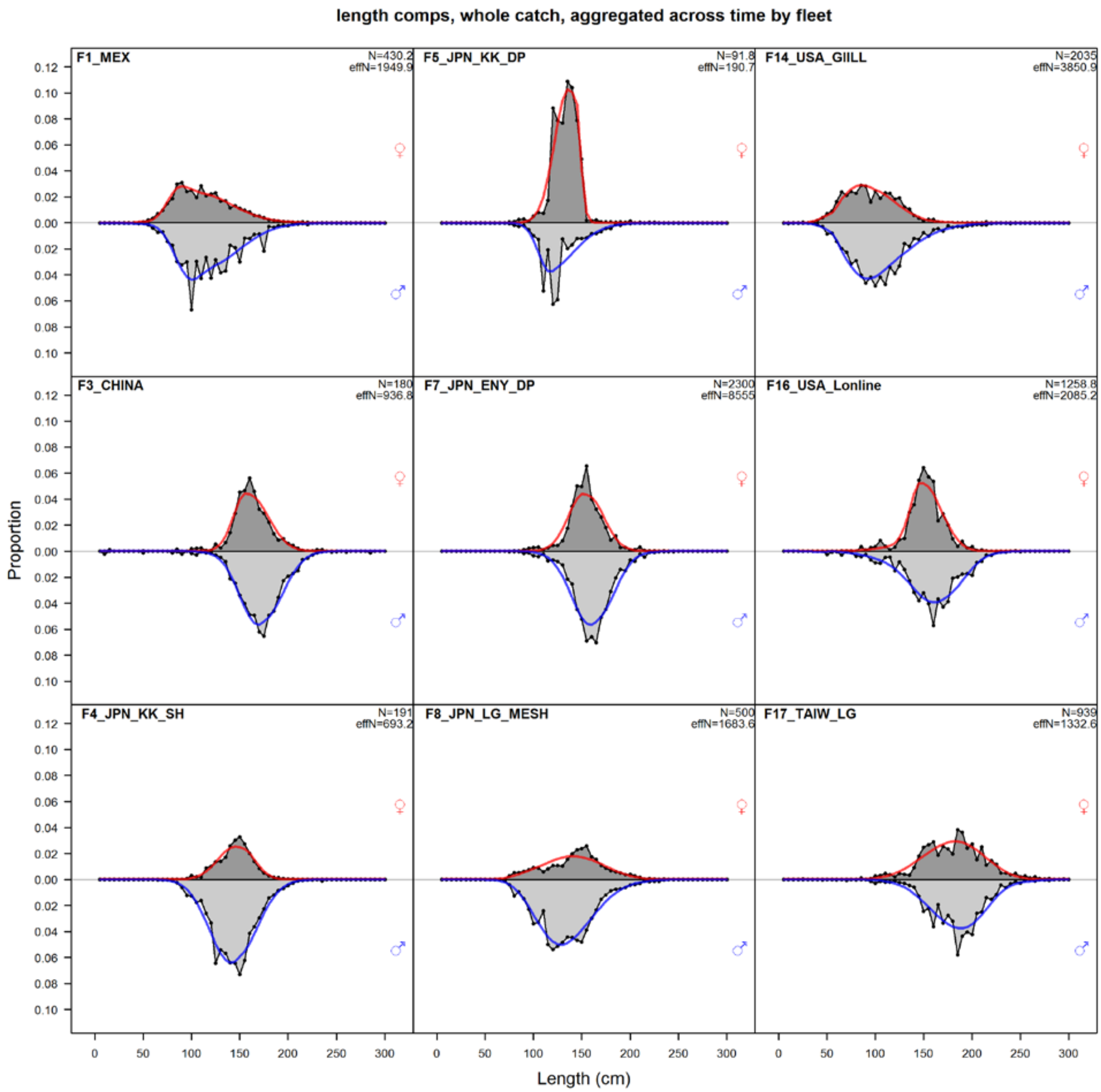


Figure 12. Sex specific comparison of observed (gray shaded area) and model predicted (blue and red solid lines) length compositions for different fisheries in the stock assessment model using Beta = 1 and Sfrac = 0.467 in the LFSR.

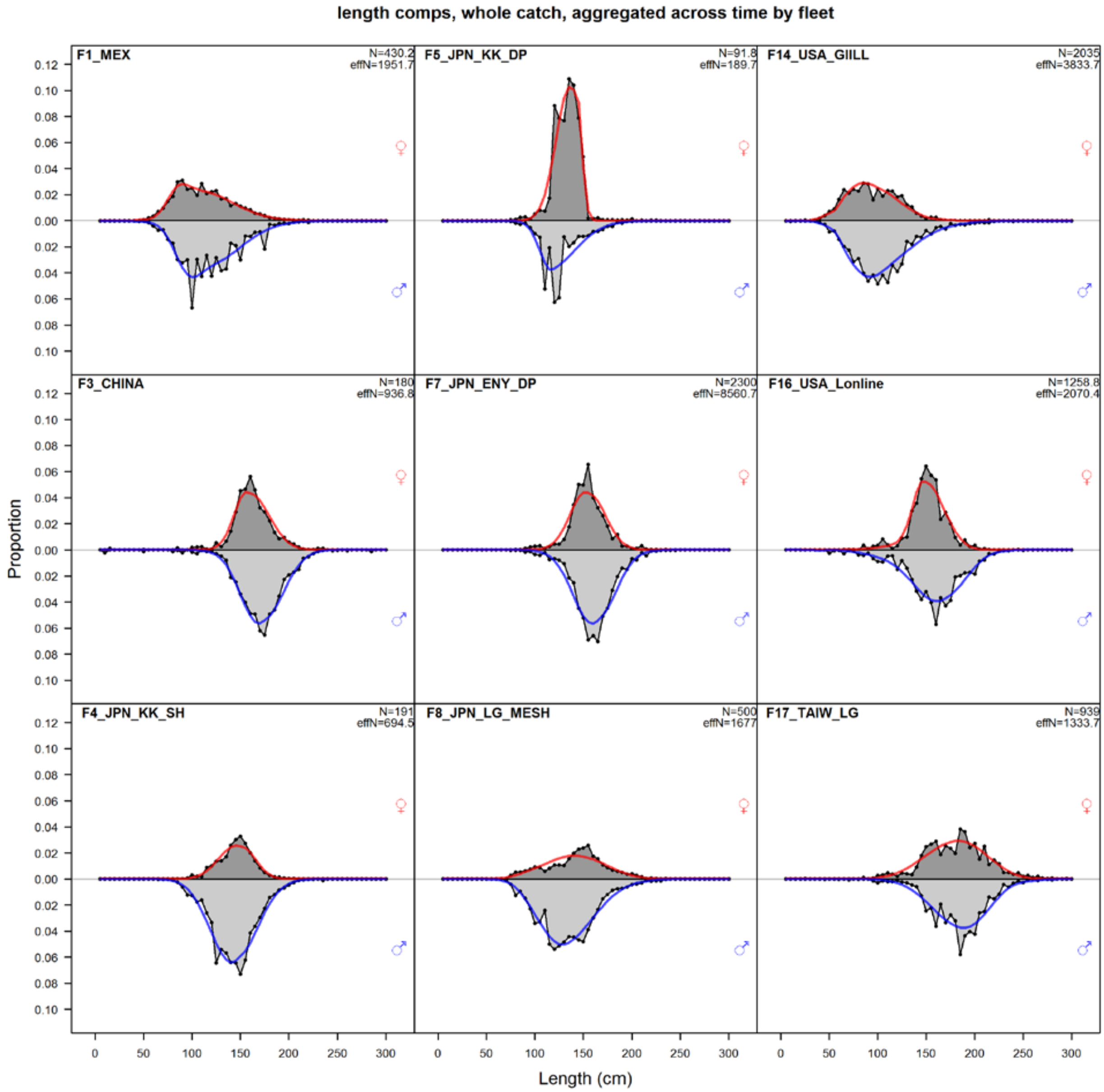


Figure 13. Sex specific comparison of observed (gray shaded area) and model predicted (blue and red solid line) length compositions for different fisheries in the stock assessment model using Beta = 2 and Sfrac = 0.391 in the LFSR.

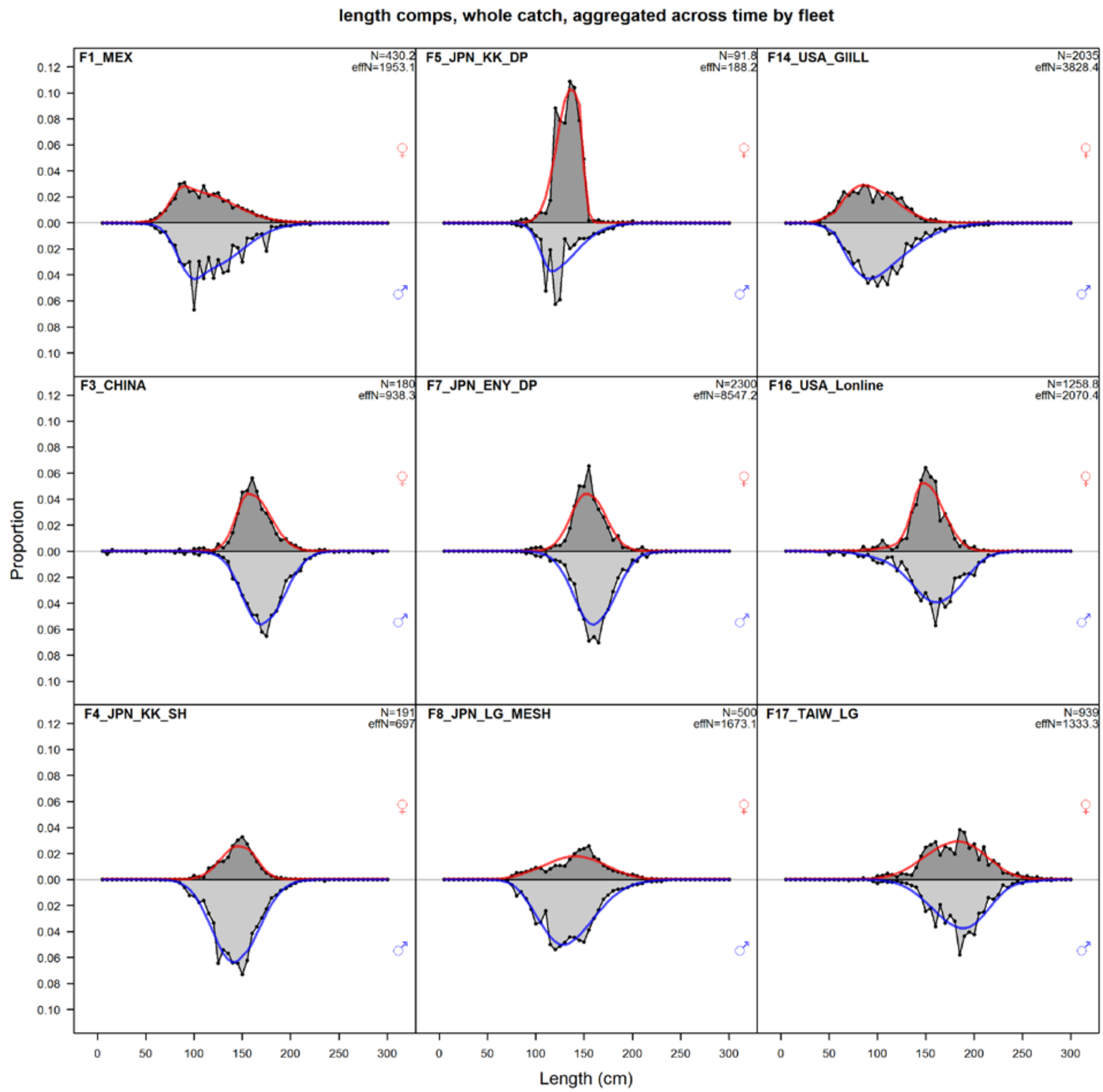


Figure 14. Sex specific comparison of observed (gray shaded area) and model predicted (blue and red solid line) length compositions for different fisheries in the stock assessment model using Beta = 3 and Sfrac = 0.378 in the LFSR.

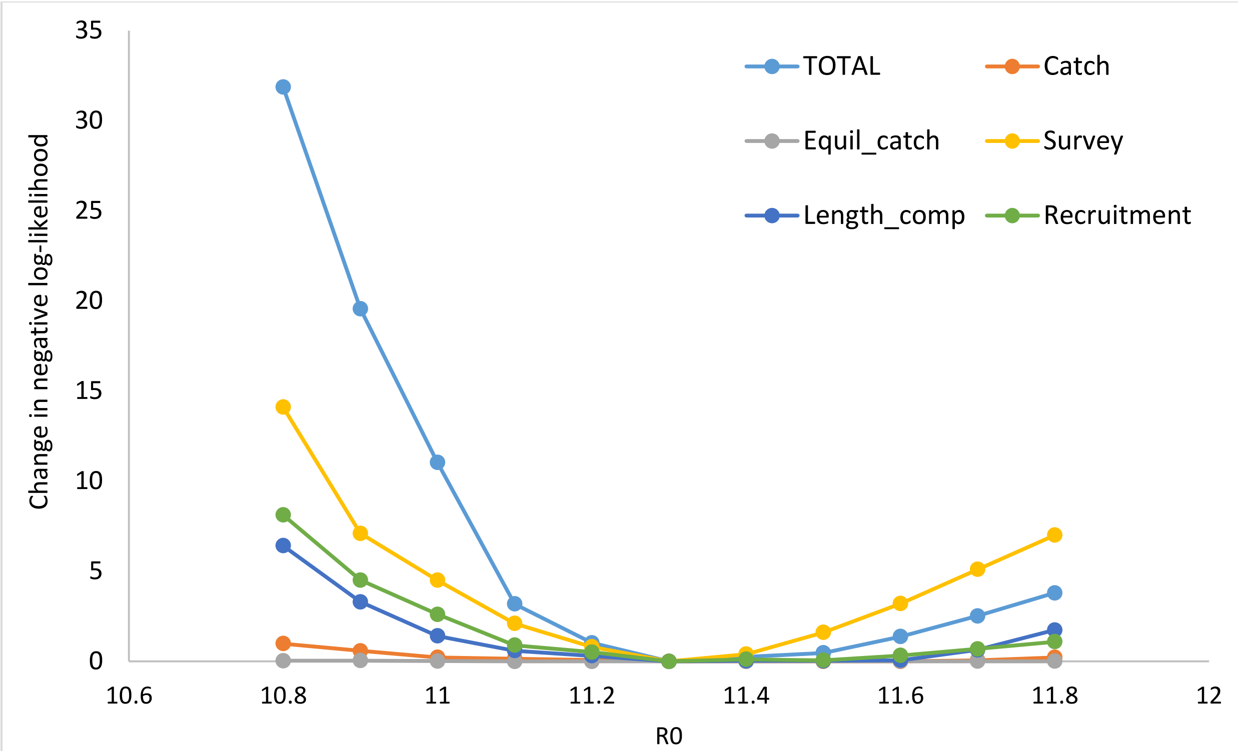


Figure 15. Profiles of the relative-negative log likelihoods by different data components for the virgin recruitment in log-scale (log(R0)) for the stock assessment model using Beta = 2 and Sfrac = 0.391 in the LFSR.

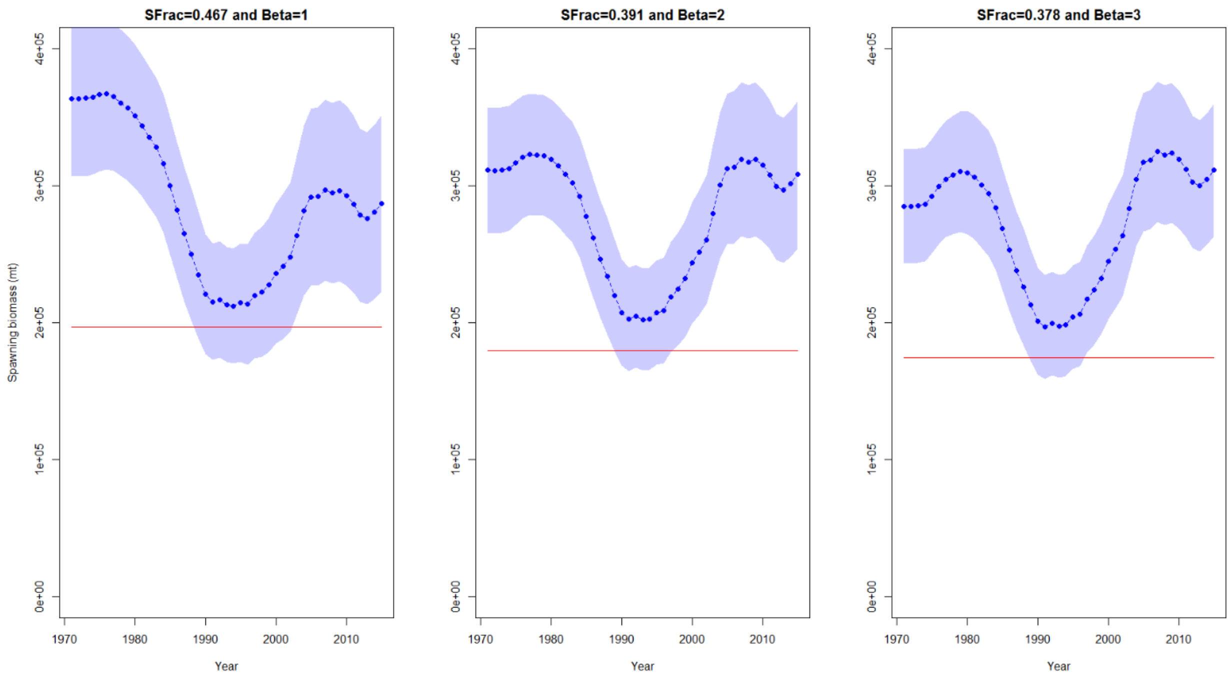


Figure 16. Comparison of time series of SSB in mt for North Pacific blue shark for different low fecundity stock recruitment relationships. Red solid lines indicate the estimates of MSY, and blue shaded area the 95%CI.

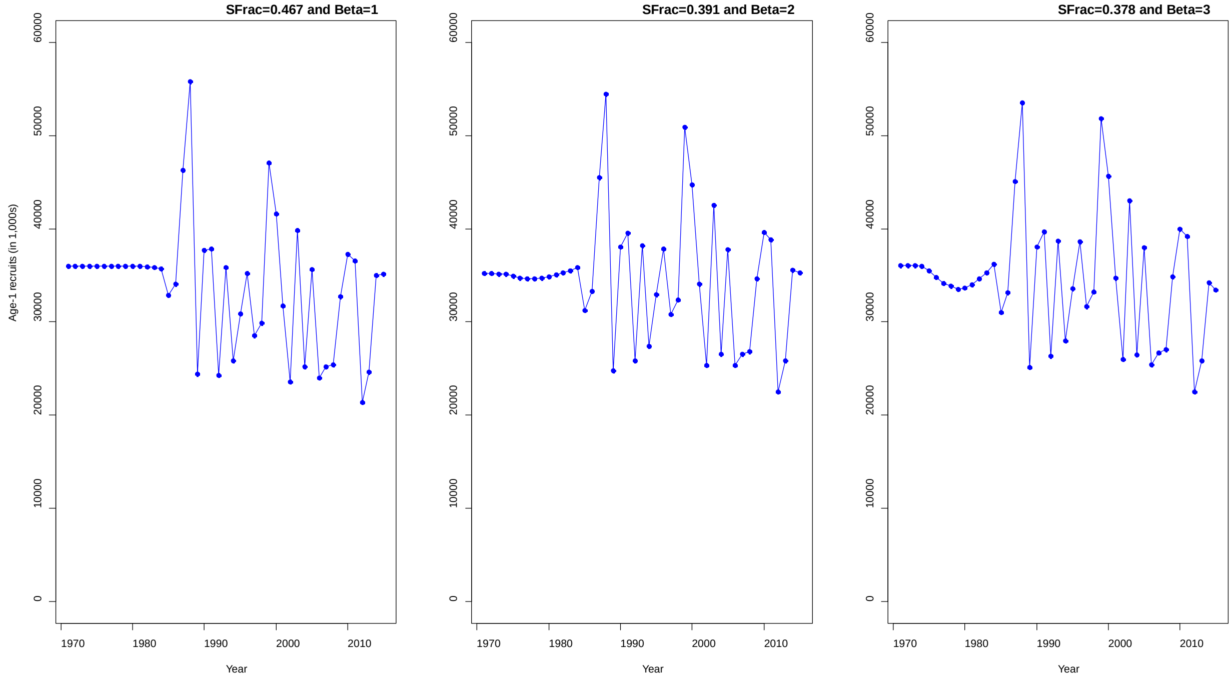


Figure 17. Comparison of time series of recruits (age-1 fish) for North Pacific blue shark for different low fecundity stock recruitment relationships.

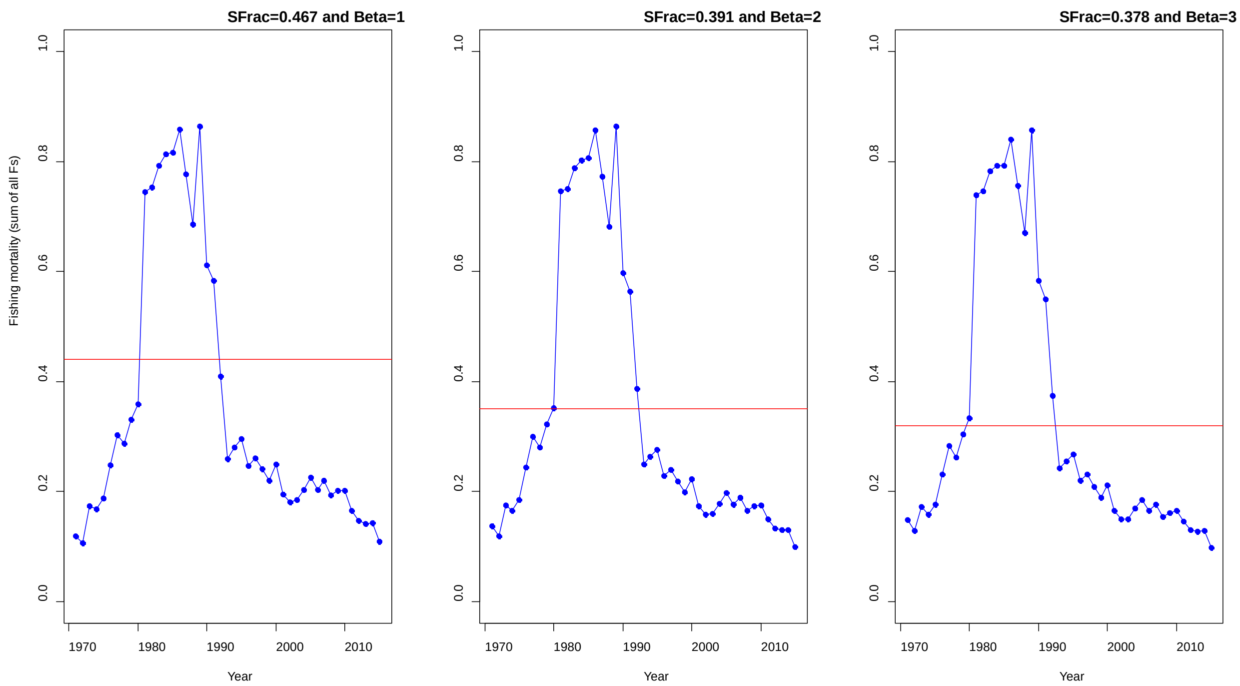


Figure 18. Comparison of time series of fishing mortality for North Pacific blue shark for different low fecundity stock recruitment relationships. Red solid lines indicate the estimates of MSY.

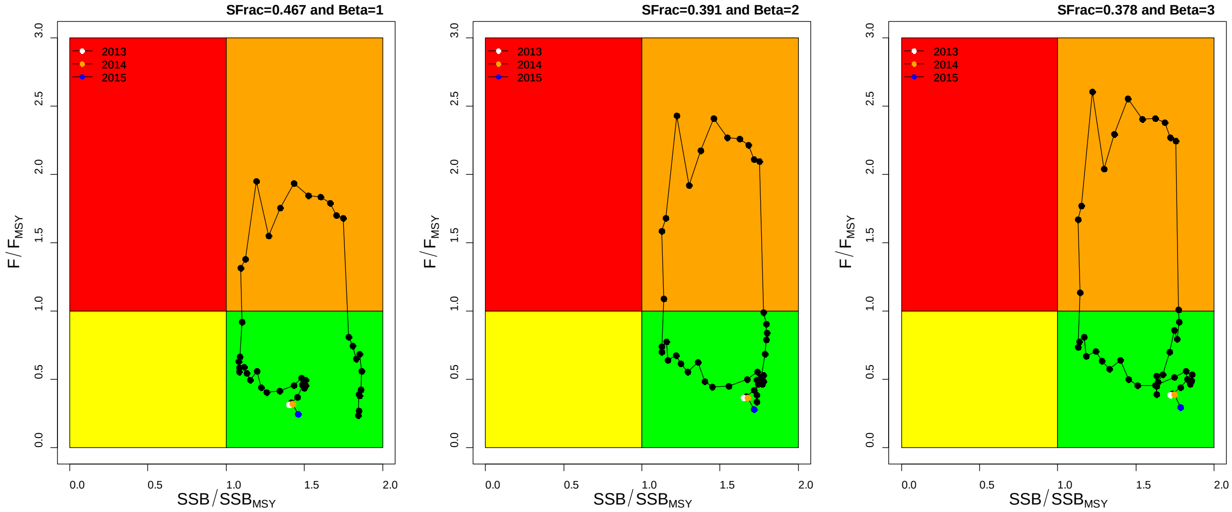


Figure 19. Kobe plot of the trends in estimates of relative fishing mortality and spawning stock biomass of North Pacific blue shark between 1971-2015 for different low fecundity stock recruitment relationships.

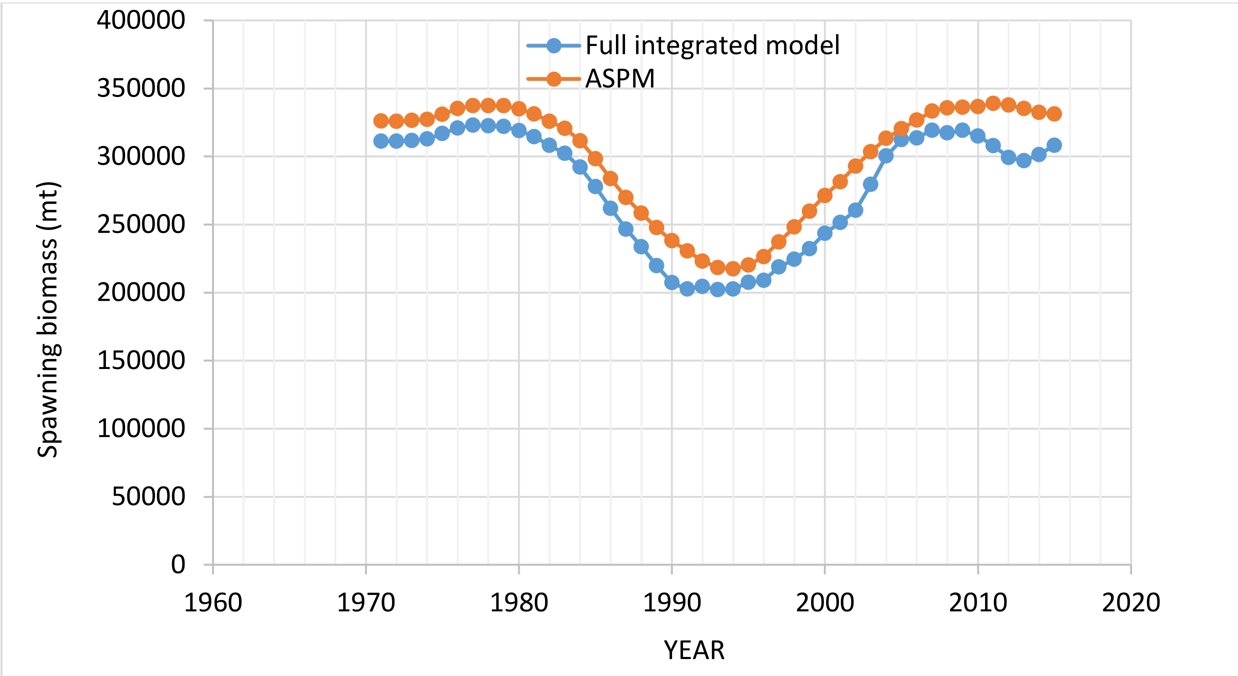


Figure 20. Age-Structured Production Model diagnostic (ASPM)

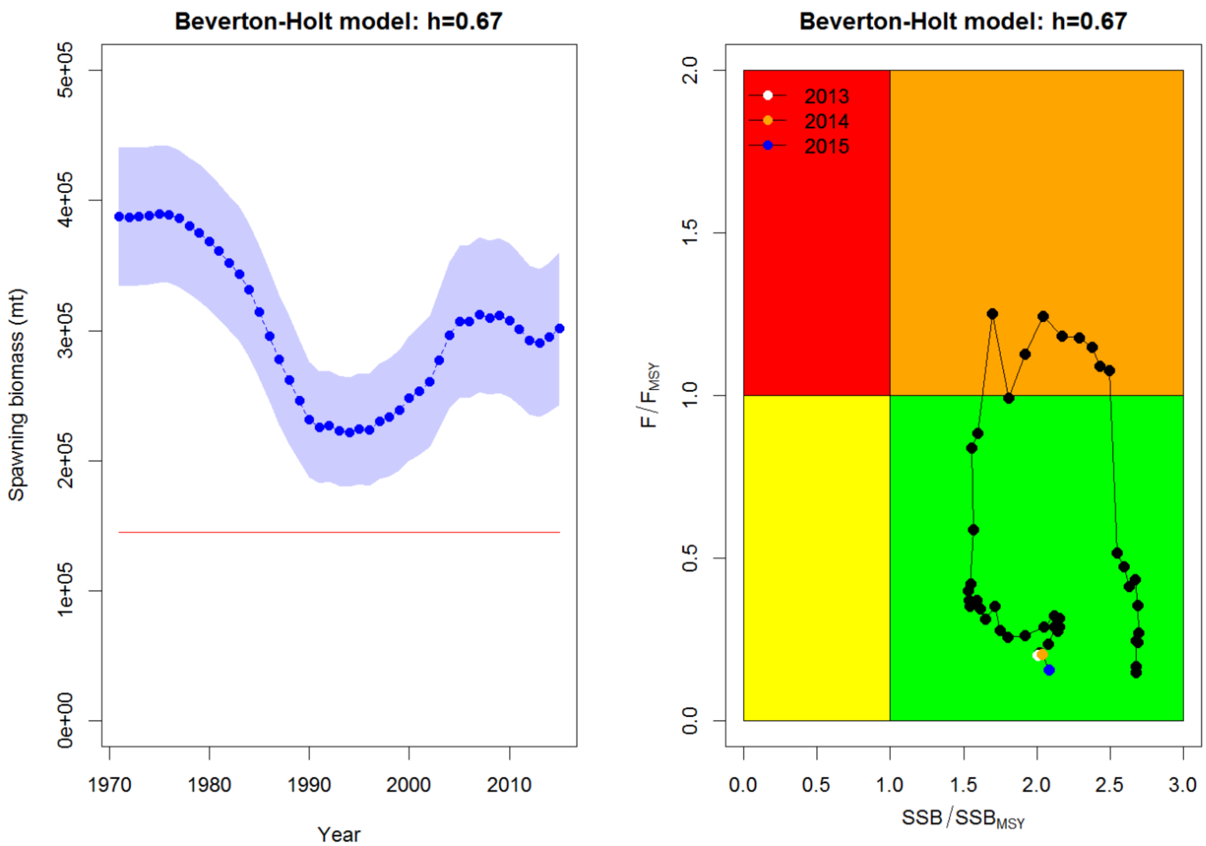


Figure 21. Results of sensitivity run with a Beverton-Holt stock recruitment relationship. Time series of spawning biomass in mt for North Pacific blue shark. Red solid lines indicate the estimates of MSY, and blue shaded area the 95%CI (left panel). Kobe plot of the trends in estimates of relative fishing mortality and spawning biomass of North Pacific blue shark between 1971-2015 (right panel).