International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean

Preliminary Striped Marlin Stock Assessment

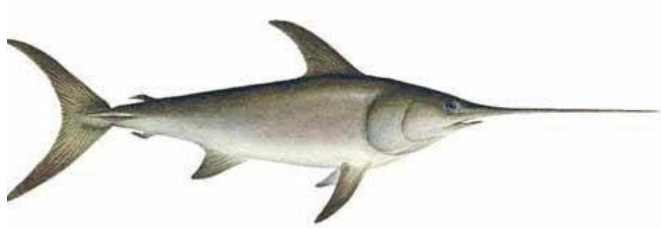
Kevin Piner[1], Hui-Hua Lee[2], Ian Taylor[3], Lyn Katahira[2], Darryl Tagami[1], and Gerard DiNardo[1]

[1]NOAA NMFS, Pacific Islands Fisheries Science Center, Honolulu, Hawaii, USA

[2]University of Hawaii, Joint Institute for Marine and Atmospheric Research, Honolulu, Hawaii, USA

[3]NOAA NMFS, Northwest Fisheries Science Center, Seattle, Washington, USA

Email: Kevin.Piner@noaa.gov



Working document submitted to the ISC Billfish Working Group Workshop, 6-16 December 2011, Honolulu, Hawaii, USA. Document not to be cited without author's written permission.

Preliminary Striped Marlin Stock Assessment[4]

Kevin Piner, Darryl Tagami and Gerard DiNardo
NOAA Fisheries
Pacific Islands Fisheries Science Center
Honolulu, HI, USA

Hui Hua Lee and Lyn Katahira
Joint Institute for Marine and Atmospheric Research, University of Hawaii
Pacific Islands Fisheries Science Center
Honolulu, HI, USA

Ian Taylor
NOAA Fisheries
Northwest Fisheries Science Center
Seattle, WA, USA

Abstract

Life-history parameters and updated time series of western and central Pacific Ocean (WCPO) striped marlin data were developed at previous meetings of the Billfish Working Group (BILLWG), International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC), for inclusion in a new striped marlin stock assessment. The new information was used in a length-based age structured Stock Synthesis model. Major model structure included single area, annual time step with observed data fit quarterly. At least one fishery was assumed asymptotic. Changes to data structure (as presented at the BILLWG meeting) included re-binning the smallest size bin to 120cm to remove the influence of misfit to the size comp of age 0 fish, and division of the Japanese Other (Harpoon) fishery into two seasonal fisheries due to a strong seasonal patterns in fish size. CPUE series were segregated into two groups based on internal model consistency, with the preferred CPUE series consisting of the Japanese longline (all areas) and the Hawaiian longline fleets. Although data series are available from 1952, the preferred model starts in 1975 when more complete data are available. Starting in 1975 also allows for the model to start without strong equilibrium assumptions. Results indicated a stock size in recent years near the lowest biomass levels observed and fishing intensity (SPR) allowing <15% of potential spawning output per recruit.

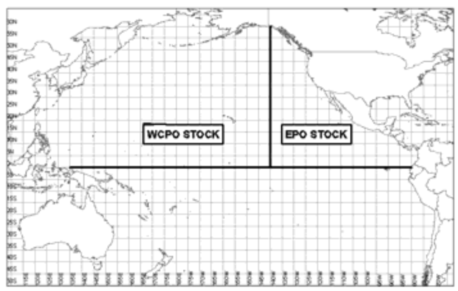
---

[4] PIFSC Working Paper WP-12-001
Issued 20 January 2012
Prepared for use in the ISC Billfish Working Group Workshop, 6-16 December 2011, Honolulu, Hawaii, USA.

Introduction

Prior ISC stock assessments of striped marlin in the North Pacific indicated a declining stock experiencing recent fishing mortality on spawners in excess of 0.7^{-yr}. Despite evidence of high fishing pressure, it was noted that there was considerable uncertainty regarding the basic biology of the stock. In particular, the stock structure, spawner-recruit resilience (h) and the growth of the species in the western North Pacific were highlighted as important areas of uncertainty.

Since the last stock assessment, considerable work on the biology of the species has been completed. Based on genetics analyses the stock boundaries were changed to reflect a Western and Central North Pacific Ocean (WCPO) stock and a separate Eastern Pacific Ocean (EPO) stock (ISC 2010).



New research has improved our understanding of growth (Sun et al. 2011a) along with new size at maturity for the same area (Sun et al. 2011b). Data for the major fisheries (DWLL Japan) were recompiled in the primary fisheries by different geographical boundaries (Area 1: 0-10°N latitude by 100°E-160°W longitude; Area 2: 10-50°N latitude by 100°E-160°E longitude; Area 3 : 10-50°N latitude by 160°E-160°W longitude) along with different time periods and updated until 2009-2010.

The objectives for this paper were to take the data series for the WCPO stock along with the new understanding of life history and explore this information inside a length-based age structured stock assessment model. We attempted to use objective means to develop model and data configurations that reduced conflict between data series. Results from this work should provide guidance for BILLWG considerations in modeling North Pacific striped marlin stock status.

Materials and methods

Overview of Methods

The methods overview will be divided into 3 sections. Section 1 describes data, Section 2 describes the preliminary analyses that were done on early model configurations to help develop base model configurations, and Section 3 describes alternative model configurations.

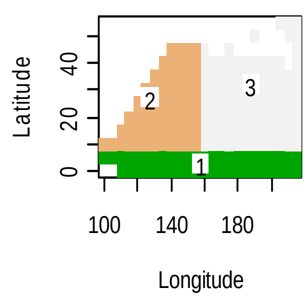
Section 1. Data

Life history

Life history information for this assessment has been taken from Sun et al. (2011a and 2011b). The combined sexes length at age relationship was based on otoliths from a maximum of age 6 fish and back-calculated lengths at age for younger ages (Figure 1). We assumed that CV on age 0.3yr fish =0.14 and age 15 yr=0.08. The assumption of the larger uncertainty in the length at age of young fish was consistent with the ageing study. The large uncertainty in the length at age of young fish is also stems from the extra variance of disparate timing of recruitment, spatial variability in growth and sexual dimorphism (although we note that the best scientific evidence does not show sexual differences in growth). Weight at length is also taken from Sun et al. (2011a). Maturity at length is based on Sun et al. (2011b) but is refit using the parameterization used in the SS3 model (Figure 2). Natural mortality (Figure 3) and steepness (h=0.87; σ_r=0.6) are the BILLWG consensus values (Piner and Lee 2011; Brodziak 2011). Observations provided in lower jaw fork length (LJFL) were transformed to eye fork length (EFL) based on Sun et al. (2011a).

Likelihood components

Likelihood components are the data that comprise the observation subcomponent of the assessment model. The three primary likelihood components are: 1) catch by fishery, 2) proportion at length of the catch by fishery, and 3) catch per unit effort (Table 1). Because of the broad spatial extent of the Japanese Distant Water Longline (DWLL) fleet, those data were divided into 3 areas resulting in 3 fisheries with all 3 types of data (catch, size composition and CPUE):



spatial distribution of JPN DWLL fisheries

Catch: Catch was inputted into the model quarterly (calendar year) from 1952 to 2010 for 19 individual fisheries in either numbers (‘000s) or weight (mt, Figure 4). We assumed catch was well known and thus it was fitted with a standard error (SE) 0.05 assuming a lognormal error distribution. We defined a poor fit to catch as models that did not remove >99% of the total catch from any fishery. Catches in 2010 were also included, using observed values when available and otherwise assuming they were equivalent to the 2009 catch.

Size Compositions: Observations of the proportion at size of the catch were included from 11 total fisheries (Figure 5). The fit to the proportion at length data assumed a multinomial error distribution with the variance described by the effective sample size (effN). Approximations of the effN were taken from an analysis of the Hawaii longline fleet (Courtney unpublished) which found ~10 fish per trip. Thus for all longline fleets effN was assumed to be number of fish measured/10. A maximum quarterly sample size was assumed to be 50. Fishery average effN are given in Table 2.

CPUE: 16 CPUE series were included and fit with a lognormal error assumption (Figure 6). The SE in log space was assumed to be the same as the CV (typically SD/estimate) described in each CPUE paper.CV was assumed to be equal to 0.2 when missing. Series with average CV<0.2 were scaled to CV=0.2 through the addition of a constant. Series with average CV>0.2 were inputted as given. Missing CVs were assumed =0.2 except for the Taiwanese longline CPUE from the early period, which was given a larger CV in accordance with the BILLWG recommendation (Table 3).

Section 2. Preliminary Analyses

This section describes preliminary analyses that went into building the base model. It includes the methods used to determine assumptions of the model. We start from analyzing appropriate larger scale structure (things unlikely to change with other changes to model) and focus on smaller scale structure at the end.

To start, a basic model was developed that described how the data was built and used the life history specifications listed above. The model was a single area model (selectivity patterns accounting for spatial patterns) that included fitting to quarterly data (Jan-Mar quarter 1 etc.). A separate selectivity pattern was fit for each fishery described above with observed composition information. Fisheries without proportion at length observations were assumed to have the same selectivity as a similar fishery as described by the BILLWG (Table 1). Time blocks to selectivity patterns (time varying) were allowed for Japanese DWLL fisheries corresponding to the time periods when CPUE catchability was assumed to change as a result of fishery practices. Ages 0-15 were included in the modeling.

<u>Assessment model</u>

This paper used Stock Synthesis (SS) as the population dynamics model. SS is a stock assessment model that estimates the population dynamics of a stock through use of a variety of fishery dependent and fishery independent information. Although its use has historically been for groundfishes, more recently it has gained popularity for stock assessments of tunas and other migratory species in the Pacific Ocean. The structure of the model allows for Bayesian estimation processes and full integration across parameter space using the Markov Chain Monte Carlo (MCMC) algorithm.

SS3 is composed of 3 subcomponents, 1) population subcomponent that recreates an estimate of the numbers/biomass at age of the population using estimates of M, growth, fecundity etc., 2) an observational sub-component that consists of the observed (measured) quantities such as CPUE or proportion at length/age, and 3) a statistical sub-component that quantifies the fit of the observations to the recreated population using likelihoods. For a complete description see (Methot 2005, 2010). This analysis uses version 3.20b.

Splitting of Japanese “Other fishery”

Preliminary data analysis showed that fishery 12 (Japanese Other fishery) contained a strong seasonal pattern in observed sizes of striped marlin (Figure 7). Larger fish are taken in the first two seasons of the year. Although some seasonality can be observed in most fisheries, this pattern was considered the most problematic, and initial modeling attempts could not adequately address the issue. We decided to split the data (catch and size) from this fishery into two seasonally-specific fisheries to reduce the influence of the misfit. This was important as this fishery (seasons 1 and 2) included observations of the largest fish and would likely be our assumed asymptotic fishery (See section below on Selectivity assumptions). We note that season 2 included both larger and smaller mode fish, but preliminary model runs showed more selectivity pattern stability if season 2 was included with season 1. All further exploration described below included the Japanese Other fishery broken into two separate fisheries: early (seasons 1-2) and late (seasons 3-4).

Determining spawning season and recruitment timing

Spawning was described in Sun et al. (2011b) as taking place from late spring throughout summer (May-August) based on histological examination. We assume that spawning biomass estimates used in the spawner-recruit model occur in season 2 (beginning of spawning cycle). Recruitment timing was assumed in the model to occur in season 3 (July-Sept) on the basis of best model fit of preliminary models (Table 4). Importantly season 3 recruitment timing showed improved fit to fisheries 1, 2 and 17 (Japan DWLL early, Japan DWLL area1 and HWLL) which take age 0 fish.

Selectivity assumptions

First, we assume that at least one fishery with observed size compositions has an asymptotic selectivity pattern and that all selectivity patterns are length-based. All other fisheries would then be allowed to be as domed, as best fits the data. This assumption means that at least one of our observational tools samples the entire population after a specific size (eliminating the possibility of model generating cryptic biomasses). This is a strong assumption and will be influential, thus the choice of the asymptotic fishery was evaluated with extensive testing. However we feel this is necessary, as models that do not have at least one asymptotic fishery often scale to unreasonable biomass levels with very good fit. We note that age-based selectivity is also invoked that allows age 0-15 fully selected for JPN DWLL early, area1, HWLL and the WCPO other fisheries. All other fisheries were considered to select only ages 1-15.

The testing for the most consistent fishery data with an asymptotic selectivity assumption consisted of sequentially assuming each fishery was the asymptotic fishery and allowing all others to be domed shaped. Because of model instability in some runs, we later changed the procedure to force two fisheries to be asymptotic (e.g. run 1 fishery 1 and 2 asymptotic, run 2 fishery 1 and 3 asymptotic etc.) for all combinations. The fishery with the best fit with assumed asymptotic selectivity pattern (in combination with other fisheries assumed asymptotic) was chosen as the asymptotic fishery. Best fit was evaluated by average total likelihood and rank (Table 5). We performed these tests across different growth assumptions and equilibrium catch assumptions. Our results indicated that the Japanese Driftnet and Japanese Other early fisheries were consistently best fitting with the asymptotic assumption. Although the driftnet result was surprising, the JPN Other Early fishery also had strong support in our testing and visually it contained the largest fish (Figure 5). Thus, we concluded that both the Japanese Driftnet and Japanese Other Early fisheries removals would be assumed to be described by an asymptotic selectivity pattern. All model runs describe from this period forward contained this result. All other fisheries are allowed to be as domed as best fit the data. We note that a few other fisheries, notably Taiwan LL, Other late and WCPO Other also had support for asymptotic selectivity patterns.

Equilibrium catch

Equilibrium catch (assumed catch prior to start of model) was set at ~75% of 1952-1953 catch. We chose only the earliest years as there was a noticeable increasing trend in catch through the 1950s. For that reason we also chose only a significant fraction of the early observed catch. We segregated catch into two fisheries: 100,000 fish into Japan DWLL early and 280mt into Japan Other early. We ran models with both 100% (high) of early catch and 50% (low). The results of the high/low catch analysis indicated that higher equilibrium catches started the model at lower levels in 1952 and lower equilibrium catches at higher levels, but by the mid 1970’s (when better data were available), the models produced similar dynamics (Figure *8*). We took this result to indicate that equilibrium catch settings primarily affected dynamics prior to the start of informative data in the 1970s and not final results.

Recruitment period

For parsimony, we chose to estimate recruitment deviations from the assumed spawner recruit curve only for years with information on recruitment. For all other years recruitment would be taken from the expectation of the S/R curve (no deviations). The years with information on recruitment was based on a preliminary model run with all recruitment deviations estimated (1952-2010). The CV of the recruitment estimates was plotted and it was assumed that information was available to inform recruitment magnitude when the CV's stabilized (Figure 9). In our case it appeared that information was available to inform recruitment starting in 1967 through 2008. Thus we now estimated recruitment during that period and used the S/R expectations for all other years. A more complex modeling process that changes the bias adjustment to account for lack of information could be used allowing for estimation of all recruitment deviations. Although this mostly affects the estimation of uncertainty, it is an area for more model development.

Lower bin bound (55 vs 120cm)

The majority of the size data supplied were binned from 55-255cm (by 5cm bins). Preliminary analysis indicated that data from 55-120cm (age 0) were both noisy and inconsistent with our model structure. We gained significant improvement in fit to proportion at size data by increasing the accumulator first bin to 120cm (accumulating all smaller fish, thus no loss in data). This improvement in fit reduced the gradients in the negative log-likelihoods from profiling across R0 for domed shaped fisheries, which we interpreted as removing misfit influence on the model results. For these reasons, we chose to include bin structure of the observations starting at 120cm- 230cm (by 5cm bins).

This decision was influenced by separate analysis that tried to explain the Hawaii longline composition information using recruitment and growth. A model was run that fit to only the composition information from the Hawaii longline fleet (best information on very young fish) and all CPUE and growth was estimated. This model explored if variation in recruitment and an alternative growth model could better explain the size data from primarily the age 0 group without the confounding influence of the other size composition information. In other words, if we could not adequately fit the age 0 size composition from that one fleet we have no hope in a model that included all fleet composition information. Results (Figure 10 (a)-(d)) indicated that we cannot adequately explain the variability in size at age 0 in the Hawaii longline samples with a time invariant but estimated growth form and variability in recruitment. This is likely due to both time varying growth and variations in spawning timing (or recruitment timing). Although changes in both growth and spawning period are possible in the model, it was deemed too complicated for implementation at this time. Changes in fishery practices may also be responsible for the misfit, but the additional process of variable year/season-specific selectivity patterns were beyond the scope of this assessment. This result reinforced our decision to bin the age 0 fish into a single bin to reduce the contribution of the misfit (due to inadequate model process) to the model results.

CPUE series to include

First we assumed that the likelihood components of both catch and proportion at length should be included in all models. However, we have 16 CPUE series which are assumed to represent the change in the segment of the population described by the selectivity pattern estimated for that fishery. It is apparent that some of the CPUE are contradictory information and thus should not be used together in a stock assessment model. We attempted to objectively segregate the CPUE series into two separate data sets based on a down-weighting analysis.

In the analysis, we sequentially down-weighted each likelihood component (excluding catch) in separate model runs. We assumed that CPUE derived from the same fishery (e.g., all 3 area 1 CPUE) described the population trajectory, and thus were included/excluded together. This analysis was performed for different assumptions (e.g. growth, equilibrium catch etc) and summarized across these assumptions. CPUE series were determined to go together if down-weighting those series led to loss of fit in the other series. Our results indicated that the Japanese DWLL area1, area 2 and area 3 and HWLL were consistent. The other series including Japan DWLL early, Japan Coastal LL and Japan Driftnet, along with Taiwanese LL early and late, were considered the alternative data set. Thus two new data sets were developed.

Data set 0 included all CPUE
Data set 1 included Japan DWLL areas 1, 2 and 3 and Hawaii LL.
Data set 2 included Japan Coastal LL, Japan Driftnet and Taiwanese LL.

Starting year

Given that the model estimation of biomass dynamics prior to the mid 1970's is influenced by our assumptions of equilibrium catch, we chose to produce a model that only fit to data from 1975-2010. In this model we estimate the initial age structure and freely estimate the fishing mortality consistent with that age structure (avoiding the equilibrium assumptions).

Based on the preliminary analyses, we explored 5 models based on data (CPUE), starting conditions and spawner recruit assumptions. Each potential model was evaluated based on the consistency of results and goodness of fit to data.

1. Model 1. Use data set 0 (all CPUE), start in 1952, fit to equilibrium catch (100K fish and 280mt) and S/R relation.
2. Model 2. Use data set 1 (subset CPUE), start in 1952, fit to equilibrium catch (100K fish and 280mt) and S/R relation.

3. Model 3. Use data set 2 (subset CPUE), start in 1952, fit to equilibrium catch (100K fish and 280mt) and S/R relation.
4. Model 4. Use data set 1 (subset CPUE), start in 1975, freely estimate equilibrium F and initial age structure and S/R relation. (Note that this model reflects the addition of a more flexible time varying selectivity pattern parameterization which was implemented during the BILLWG meeting rather than in the original draft of this working paper).
5. Model 5. Use data set 1 (subset CPUE), start in 1975, freely estimate equilibrium F and initial age structure with recruitment deviations treated as free parameters.

Section 3. Preferred Model Configuration

Our preferred model is model 4, with the following characteristics.

Start year: 1975

Data sets: data set 1 including all catch, all proportion at size, and CPUE data set 1.

Initial conditions: estimate initial age structure and fishing mortality

Asymptotic selectivity assumption: Japanese driftnet and Other early fishery;
3 time blocks for selectivity pattern for the JPN DWLL areas 2 and 3;
Only JPN DWLL area 1, HWLL and WCPO Other fisheries are allowed to take age 0 fish, all other fisheries are constrained to taking age 1+.

Results

In this section we describe model results for our preferred model (model 4) with comparison to the other models in some derived quantities:

Estimated selectivity patterns are given in Figure 11.

Observed and estimated CPUE by fishery are given in Figure 12.

Observed and predicted proportion at length are given in Figures 13 and 14.

Likelihood profiles across the parameter R0 (unfished recruitment the primary scaling parameter) is given are Figure 15 (total likelihood), Table 6 (size composition components) and Table 7 (CPUE components). For comparison, a likelihood profile for Model 1 (start in 1952 with all CPUE included) is also given. Likelihood values from the best fit to model 4 are given in Table 8.

Convergence to a global minimum was examined by randomly perturbing the starting values of all parameters by 10% and refitting the model (Figure 16). There is no evidence of a better fit.

Estimated spawning stock biomass and age 1+ biomass are given in Figure 17 and Figure 18. A comparison to the other candidate models is depicted. All models indicated a stock with biomass at levels below their long-term average.

Estimated spawning recruitment is given Figure 19. A comparison to the other candidate models is depicted. All models indicated a decline in recruitment over the last decade.

Estimated SPR is given Figure 20. A comparison to the other candidate models is depicted. During the period of informative data, exploitation has typically allowed 10-20% of the spawning potential. The most recent years have been closer to 12%.

Table 9 gives yearly estimated values of spawning biomass, recruitment and SPR and their asymptotic standard deviation estimates.

Discussion

We believe that the reduced model (starting in 1975 and subset of CPUE indices) is preferable to the more full models because of the limitation of model conflict apparent in the R0 profiles and because data after 1975 are, in our opinion, much more reliable (e.g. HPB available for standardization of CPUE, size data available etc.). In our preferred model, gradients of likelihood are minimized for the size composition data and indices are influential in the results. Consequently, the fit to the indices (data set 1) are acceptable except for the middle period in the JPN DWLL area 2, which is a conflicting trend to all other areas. The authors also believe that estimation of the starting conditions is preferable to specifying those same conditions. A quick check of the validity of this approach is the model's estimate of 8500 tons prior to the starting of the dynamics in 1975 is within the range (~6000-10,000 mt) reported in the decade prior to 1975.

It appears that the model generally follows the Japanese Area 3 DWLL indices, especially in the period 1975-1999. While we do not know if this is correct, but we note that the majority of the LL catch since 1975 has come from this area. Had we chosen a priori which indices to believe, it is likely we would have chosen area 3 as our primary indices. This is due to the magnitude of the catch, location of biomass and that Japanese longline fisheries are often the most trusted data source for CPUE information. This decision would have been consistent with the prior assessment philosophy as well.

We note that time varying selectivity was not allowed for Japan DWLL area 1 size information, which was inconsistent with how the CPUE were treated as 3 independent time series. It was also inconsistent with the treatment of the other JPN DWLL areas (3 time blocks of selectivity). However, preliminary runs showed selectivity instability in area 1. In models that include data back to 1952 we shared selectivity parameters from JPN DWLL area 1 and the JPN DWLL early fisheries for the same reason. Sample sizes in Area 1 are small and the data is both bimodal and noisy. It was possible to use spline functions to better approximate the bimodal distribution of sizes seen in this fishery, however it was our opinion that the bimodality was due to sampling issues and was not likely representative of the catch. Therefore we continued to use a functional form of selectivity that we felt better represented the performance of the gears. The likelihood profile across R0 for this fishery did not indicate a strong gradient of misfit, indicating any misfit for the for Japan DWLL area 1 size had little influence on model results.

The estimated selectivity patterns for most longline fisheries are decidedly dome-shaped. Whether this result reflects gear operations (such as depth fished, bait, etc.) or is related to the spatial distribution of the fleet relative to the size structure of the population is not clear. Work to address on a finer spatial scale the location of fish size caught and fishing effort is warranted. A third possibility is that this reflects a bias is the size sampling process, but this is thought to be less likely. Uncertainty in the life history parameters (growth and mortality) is also influential in the degree of dome shape.

Fit to the size composition information is generally good, especially for fisheries with the most flexible selectivity patterns and large sample sizes (JPN DWLL area 2 and 3, JPN CLL). The greatest level of model misfit appears to be from the two fisheries with the strong assumptions of asymptotic selectivity (least flexible selectivity assumption). The model generally favors more large fish than were observed. Although the likelihood profiles across R0 indicate good general support for the relative scale, it is the asymptotic selectivity pattern assumptions that prevent scaling to unrealistic biomass levels where recruitment alone can explain changes to data. We also note here that the driftnet and other fisheries effN calculations did not include the division by 10 associated with the LL fisheries and thus true sample size may be overestimated.

In the authors' opinion, the largest uncertainty in growth (Sun et al. 2011a) is how large we should expect a typical fish to get. The growth curve we used is based on observed fish size at age 6 and back calculated size-at-age for ages <6. Sun et al. point out that we really don't have strong evidence for how large fish are likely to be after age 6. This is an area that needs more research. They also note that the maturity-at-length relation does not represent that a size class is fully matured until fish are larger than routinely seen in our data. Alternative maturity ogives should be explored.

We also note that iterative rescaling of variance has not been done for any models. Typically this is something we advocate. However our model expectations are not far from the inputted variance levels and in this exploratory work, we wished to retain the comparability of likelihoods across models. We would advocate an iterative re-scaling of the inputted variances to conform to the model's expectations on any final model developed by the working group. This iterative rescaling would effectively down-weight the CPUE from the JPN distant water longline area 2 middle period via inflated SE. If the iterative rescaling is not done we would advocate the removal of the JPN DWLL area 2 middle index to prevent its misfit from causing problems.

Finally we note that it was not the intention of this paper to characterize stock status relative to reference points, but rather to explore model and data structure to facilitate BILLWG work. However, results of our analysis suggest that the stock is being fished at a relatively high level and recruitment is below long term averages.

Literature Cited

Brodziak, J.K.T., (2011) Probable values of steepness for north Pacific. ISC/11/BILLWG/2-XX

ISC. (2010) Report of the Billfish Working Group Workshop. Hakodate, Hokkaido, Japan

Methot R.D. (2005) Technical Description of the Stock Synthesis II Assessment Program Version 1.21- March 2005

Methot R.D. (2010) User manual for SS3 V.3.11a. September 2010

Piner, K.R., H.H. Lee (2011) correction to meta-analysis of striped marlin natural mortality. ISC/11/BILLWG/2/07

Sun, C.L., W.S Hsu, N.J. Su, S.Z. Yeh, Y.J. Chang and W.C. Chiang (2011a). Age and Growth of striped marlin (Kajikia audax) in the waters off Taiwan: A revision. ISC/11/BILLWG-2/07

Sun, C.L., S.Y. Chang, Y.J. Chang, S.Z. Yeh, W.C. Chiang, and N.J. Su. (2011b) Reproductive biology of female striped marlin, Kajikia audax, in the waters off Taiwan (preliminary). ISC/11/BILLWG-1/XX

Tables

Table 1. Likelihood components considered for inclusion in the population dynamics model. Fishery number is the identification numbering inside the stock assessment model. Mirrored is a term that defines which selectivity pattern is assumed to be the same as the fishery for describing removals. Time blocks indicate the number of discreet periods where selectivity was allowed to change.

Fleet number	Fishery	Catch (units)	Composition information (years)	CPUE (number of series)	Selectivity shape (est. or assumed)	Mirrored fleet (time blocks)
1	Japan DWLL early	1952-1974 (#'s)	1970-1974	1	Domed (est)	
2	Japan DWLL area 1	1975-2010 (#'s)	1975-2006	3	mirrored	
3	Japan DWLL area 2	1975-2010 (#'s)	1975-2009	3	Domed (est)	(3)
4	Japan DWLL area 3	1975-2010 (#'s)	1975-2009	3	Domed (est)	(3)
5	Japan Coastal LL	1952-2010 (mt)	1986-2009	1	Domed (est)	
6	Japan Driftnet	1952-2010 (mt)	1980-1983,91,00,05,08,09	2	Asymptotic (est)	
7	Japan OLL	1952-2010 (mt)		0	mirrored	5
8	Japan Squid	1952-2010 (mt)		0	mirrored	6
9	Japan Bait	1952-2010 (mt)		0	mirrored	5
10	Japan Net	1952-2010 (mt)		0	mirrored	5
11	Japan Trap	1952-2010 (mt)		0	mirrored	5
12	Japan Other	1952-2010 (mt)	1972-2000	0	Asymptotic (est)	
13	Taiwan Longline	1952-2010 (mt)	2006-2009	2	Domed (est)	
14	Taiwan OSLL	1952-2010 (mt)		0	mirrored	13
15	Taiwan Coastal	1952-2010 (mt)		0	mirrored	13
16	Hawaii LL	1952-2010 (mt)	1994-2010	1	Domed (est)	
17	WCPO others	1952-2010 (mt)	1993-2009	0	(Domed est)	
18	Korea LL	1952-2010 (mt)	2005	0	mirrored	3

Table 2. Inputted and estimated sample sizes of the size composition information. Fitted values are taken from model 1 (with all indices included) and model 4 (subset of indices and years). N/A indicates not used.

			model 1	model 4	
Fleet		N	mean_effN		mean(inputN)
1	JPN_DWLLearly	20	45.4538	N/A	49.255
2	JPN_DWLL1	69	13.8579	13.3282	9.79275
3	JPN_DWLL2	131	31.4654	32.8447	39.8206
4	JPN_DWLL3	135	41.3027	47.1968	42.16
5	JPN_CLL	91	47.5169	47.7605	39.0429
6	JPN_DRIFT	15	41.3261	36.4574	18.32
12	JPN_OTHER_early	37	32.9526	32.6114	44.7027
13	JPN_OTHER_late	14	33.6786	34.0571	37.5714
14	TWN_LL	13	27.8378	34.175	10.0538
17	HW_LL	66	25.4182	26.3092	26.7864
18	WCPO_OTHER	53	25.4125	25.1784	3.30377
19	KOR_LL	1	*28.6079*	30.6262	5.1

Table 3. inputted and estimated CPUE SE. Fitted values are taken from model 1 (with all indices included) and model 4 (subset of indices and years). N /A indicates not used.

			model 1	model 4	model 1	model 4	
Fleet		N	Q	Q	r.m.s.e.	r.m.s.e	Inputted SE
20	Svey1_JPN_DWLLearly	23	0.000143		0.486404		0.204414
21	Svey2_JPN_DWLL1	12	6.94E-06	7.09E-06	0.307005	0.305093	0.201767
22	Svey3_JPN_DWLL1	13	1.03E-05	9.66E-06	0.426362	0.399991	0.201506
23	Svey4_JPN_DWLL1	10	1.90E-05	1.59E-05	0.192534	0.156083	0.203003
24	Svey5_JPN_DWLL2	10	4.16E-05	3.57E-05	0.634289	0.613963	0.258454
25	Svey6_JPN_DWLL2	12	2.97E-05	2.70E-05	0.416769	0.390271	0.195439
26	Svey7_JPN_DWLL2	13	7.03E-05	6.06E-05	0.563676	0.537163	0.200912
27	Svey8_JPN_DWLL3	12	0.000856	0.0007625	0.268026	0.261032	0.204767
28	Svey9_JPN_DWLL3	13	0.000947	0.0007104	0.228455	0.225579	0.203406
29	Svey10_JPN_DWLL3	10	0.000344	0.0002693	0.47981	0.412165	0.198173
30	Svey11_JPN_CLL	16	0.000199	*N/A*	0.301389	*N/A*	0.199211
31	Svey12_JPN_DFT	17	0.000849	*N/A*	0.409024	*N/A*	0.200941
32	Svey13_JPN_DFT	9	0.009273	*N/A*	0.215227	*N/A*	0.195863
33	Svey14_TWN_EARLY	24	0.000215	*N/A*	0.454614	*N/A*	0.4
34	Svey15_TWN_LATE	15	0.000487	*N/A*	0.226786	*N/A*	0.2
35	Svey16_HWLL	14	0.003471	0.0025896	0.525457	0.465904	0.286402

Table 4. Results of the test of seasonality of recruitment. Column headings are: season of assumed recruitment, total negative log-likelihood, and the change in negative log- likelihood from season 3 for each length composition component (fishery) (negative log likelihood minus season 3 negative log likelihood). A negative value indicates better fit (highlighted in green), and a positive value indicates worse fit (highlighted in red).

Season	Total	Fishery 1	Fishery 2	Fishery 3	Fishery 4	Fishery 5	Fishery 6	Fishery 12	Fishery 13	Fishery 14	Fishery 17	Fishery 18
season 1	5775.29	22.361	-10.828	58.98	46.63	63.354	4.3981	17.304	3.14	0.2457	172.533	8.0768
season 2	5583.85	21.087	-16.015	31.28	19.42	31.155	2.4114	4.998	0.6	0.4442	94.45	-0.1291
season 3	5408.16	0	0	0	0	0	0	0	0	0	0	0
season 4	5408.19	-3.551	-21.759	-11.73	-1.56	-23.119	0.8188	-4.431	2.487	-0.6615	48.327	-0.428

Table 5. Results of the test of for asymptotic fishery. Rank average is the ordered rank (1=best) of fit from the cross pair analysis. Likelihood average is the average of the negative log likelihoods from the cross pair analysis (lower is better fit). Values bolded indicate support for asymptotic fishery. Results are shown for two of the series of model assumptions (CV of length-at-age relation).

	CV0.14		CV 0.08	
Fishery	rank average	likelihood average	rank average	likelihood average
1	5.7	175	6.2	153.9
3	7.1	459.5	7.2	849.3
4	7.4	big	6.8	412.4
5	5.1	143.8	3.5	701.6
6	2.1	28.5	2.6	45.1
12	2.9	75.9	1.6	34.9
13	4.3	306.3	4.9	84.4
14	3.3	508.6	3.2	67.5
17	4.3	121	5.7	121.4
18	4.3	124.1	2.3	30.7

Table 6. Negative log-likelihood gradients (component likelihood-minimum likelihood within component) for size composition information for model 1 and the preferred model 4. The best fitting R0 will have a likelihood of 0 and changes in negative log-likelihood across different values of R0 can be thought of as how much information there is on scaling from that likelihood component.

model 1	DW early	JPN DWLL1	JPN DWLL2	JPN DWLL3	JPN CLL	JPN DRIFT	JPN OTHER_early	JPN OTHER_late	TWN LL	HW LL	WCPO
6.3	34	20	1	0	2	1	18	0	4	0	2
6.4	15	0	6	4	4	2	2	0	5	42	3
6.5	2	3	1	5	0	0	2	1	3	3	1
6.6	5	6	0	7	0	2	0	2	21	4	0
6.7	1	2	11	12	18	25	44	3	0	10	9
6.8	0	1	8	9	18	31	67	3	0	8	12

Size comp

Model 4		JPN DWLL1	JPN DWLL2	JPN DWLL3	JPN CLL	JPN DRIFT	JPN OTHER_early	JPN OTHER_late	TWN LL	HW LL	WCPO
6.3		0	0	0	0	0	3	0	2	0	0
6.4		0	0	1	0	0	2	0	1	1	0
6.5		0	1	1	0	1	1	1	1	2	0
6.6		0	2	2	0	1	0	1	0	3	0
6.7		2	3	1	10	21	11	3	0	7	1
6.8		3	3	0	10	28	17	3	0	5	1

Table 7. Negative log-likelihood gradients (component negative log-likelihood minus minimum likelihood within component) for CPUE series information for model 1 and the preferred model 4. The best fitting R0 will have a likelihood of 0 and changes in negative log-likelihood across different values of R0 can be thought of as how much information there is on scaling from that likelihood component.

Model1	DW early	DWLL Area1			DWLL Area2			DWLL area 3			JPN _CLL	JPN_ DFT	JPN_ DFT	TWN_ EARLY	TWN_ LATE	HWLL
6.3	111	4	11	0	6	7	29	0	1	6	0	3	3	8	0	2
6.4	55	2	9	1	4	8	22	1	0	3	3	0	3	1	0	0
6.5	23	2	11	2	6	10	26	1	0	8	1	0	2	0	1	2
6.6	13	3	12	4	9	13	28	1	0	13	2	0	1	0	3	5
6.7	4	0	0	0	1	1	4	0	4	1	14	3	0	3	1	3
6.8	0	0	0	0	0	0	0	0	7	0	16	4	0	4	3	3

CPUE

Model 4			DWLLArea1		DWLLArea2			DWLLArea3								HWLL
6.3		1	10	0	4	7	20	0	0	3						0
6.4		1	10	0	4	7	20	0	0	3						0
6.5		1	10	1	4	8	20	0	0	4						1
6.6		1	10	1	5	9	20	0	0	5						1
6.7		0	1	1	0	2	5	1	4	0						6
6.8		0	0	0	0	0	0	2	7	0						9

Table 8. Negative log-likelihood values from Model 4. Top: negative log-likelihoods by component. Bottom: negative log-likelihoods by component along with lambda values (likelihood multiplier). Note that individual component values are sometime rounded off for presentation.

Likelihood Component	Negative-log Likelihood
TOTAL	4441.73
Catch	1.6666
Equil_catch	0
Survey	8.74249
Length_comp	4431.14
Recruitment	0.177562
Forecast_Recruitment	0
Parm_priors	0
Parm_softbounds	0.00933656
Parm_devs	0
Crash_Pen	0

Fleet:	ALL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Catch_lambda:	_	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Catch_like:	1.6666	0.0255	0.046	0.05	0.06	0.99	0.06	0	0.06	0.06	0.06	0	0.065	0	0.055	0.055	0.042	0.042	0
Surv_lambda:	_	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Surv_like:	8.7425	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Length_lambda:	_	1	1	1	1	1	0	0	0	0	0	1	1	1	0	0	1	1	0
Length_like:	4431.1	305.293	1170	1120	583.7	63.4	0	0	0	0	0	607.6	136	27.16	0	0	356.5	61.47	1.252

Fleet:	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
Catch_like:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Surv_lambda:	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1
Surv_like:	-5.0644	5.2524	-13	15.54	5.085	25.42	-9.25	-12.7	3.71	0.618	12.71	-2.12	-3.97	-9.06	-6.28
Length_lambda:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Length_like:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 9. Estimated Spawning biomass, recruitment and exploitation level (1-SPR) from model 4 along with their estimated asymptotic standard deviation.

Year	SSB(mt)	StdDev	Recruitment (000's)	StdDev	1-SPR	StdDev
Virgin	19811.5	779.137	592.622	23.3063		
Initial	1348.73	614.773	592.622	23.3063		
1975	3189.22	381.824	432.297	56.5365	0.869459	0.009352
1976	3191.69	275.9	501.833	50.3	0.836524	0.011224
1977	3033.05	244.366	286.261	41.2967	0.868677	0.007653
1978	2379.31	165.492	1297.45	68.7382	0.925711	0.004202
1979	1634.87	96.3727	350.886	55.5964	0.869535	0.007277
1980	2474.56	147.796	549.941	62.0262	0.859067	0.009123
1981	2972.81	218.797	554.59	54.8342	0.852473	0.010662
1982	2807.83	244.177	225.822	39.5199	0.814413	0.012694
1983	2907.84	245.622	454.892	54.6577	0.823418	0.012243
1984	2462.37	234.137	1699.74	89.2203	0.901192	0.008888
1985	2394.68	222.392	253.553	56.6132	0.838054	0.011498
1986	3538.71	249.81	351.567	45.0333	0.842594	0.010048
1987	4004.92	314.339	838.41	64.2183	0.81884	0.013704
1988	3420.8	339.629	573.174	64.3452	0.887964	0.010021
1989	3018.11	330.425	315.814	52.3444	0.846277	0.013867
1990	2934.32	328.576	939.343	69.477	0.844831	0.014616
1991	2935.61	330.468	293.179	57.5956	0.830165	0.01499
1992	3165.84	327.846	778.016	52.9497	0.811342	0.015569
1993	3331.46	331.806	122.888	32.3994	0.843359	0.012284
1994	3177.9	316.287	570.019	39.4208	0.854517	0.011899
1995	2652.96	294.538	341.078	34.3753	0.895164	0.009795
1996	2018.01	257.723	345.098	33.3305	0.875975	0.012419
1997	1796.28	236.738	680.534	51.5301	0.869713	0.014341
1998	1803.65	235.624	348.326	43.6157	0.888911	0.012266
1999	1872.56	245.133	310.558	38.0867	0.863409	0.016733
2000	1882.84	280.941	562.703	50.9575	0.850408	0.021175
2001	1943.76	324.316	310.426	39.9698	0.827797	0.025131
2002	2196.37	369.633	631.192	53.5515	0.774165	0.030271
2003	2537.12	422.677	306.922	41.6452	0.77443	0.029671
2004	3018.06	467.497	128.711	25.9375	0.730818	0.028737
2005	3232.79	482.644	434.557	33.5022	0.748071	0.026793
2006	3047.53	469.973	115.762	25.0443	0.780834	0.023887
2007	2662.39	425.596	187.273	25.1795	0.779971	0.023549
2008	2230.38	373.98	121.311	27.2894	0.868964	0.018345
2009	1459.97	319.469	403.274	33.4563	0.868791	0.026941
2010	1143.46	304.042	368.129	41.9181	0.840335	0.034457

Figures

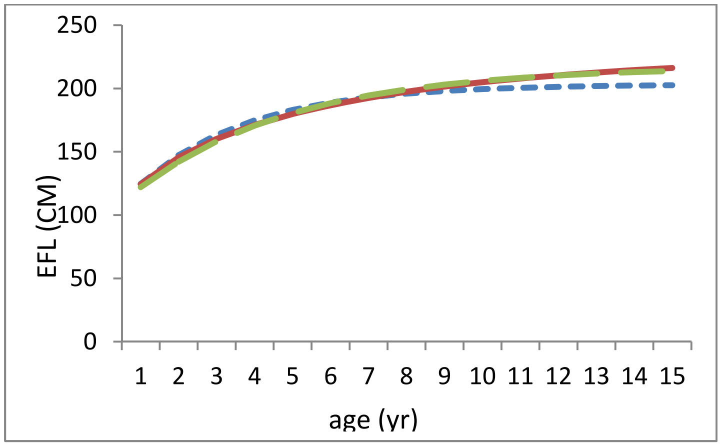


Figure 1. Plot of the Generalized Von Bertalanffy length at age (red solid line) and assessment model representation of that relationship (olive long dash). Also plotted is an alternative growth model (blue short dash) from the source paper (Sun et al. 2011a). The alternative growth form illustrates that there is uncertainty in the size of the oldest fish.

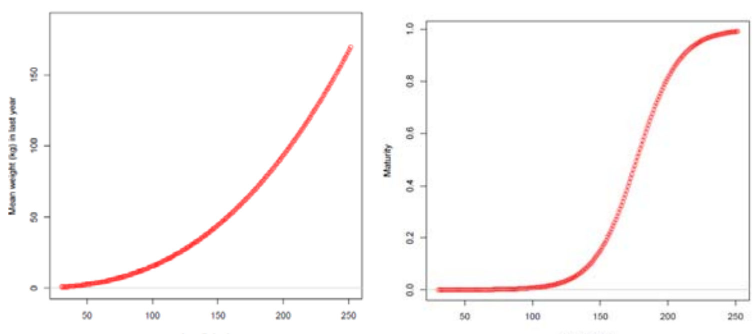


Figure 2. Weight-at-length and maturity-at-length used in the stock assessment model.

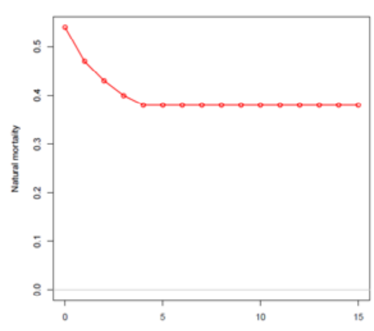


Figure 3. Natural mortality-at-age assumed in the population dynamics model.

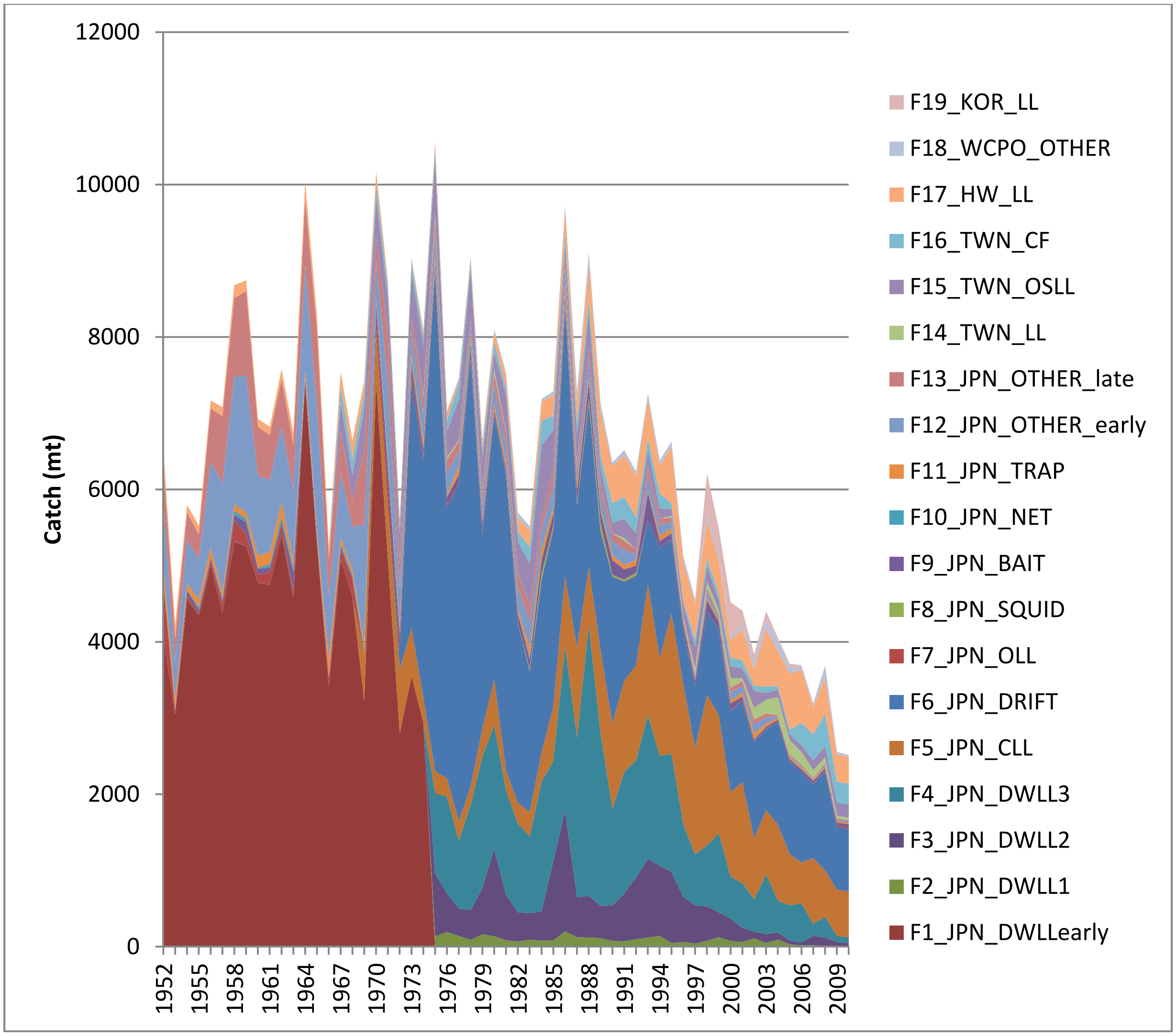


Figure 4. Plot of catch of striped marlin by year and fishery in mt. Fisheries with catch reported in numbers were converted into mt inside the stock assessment model.

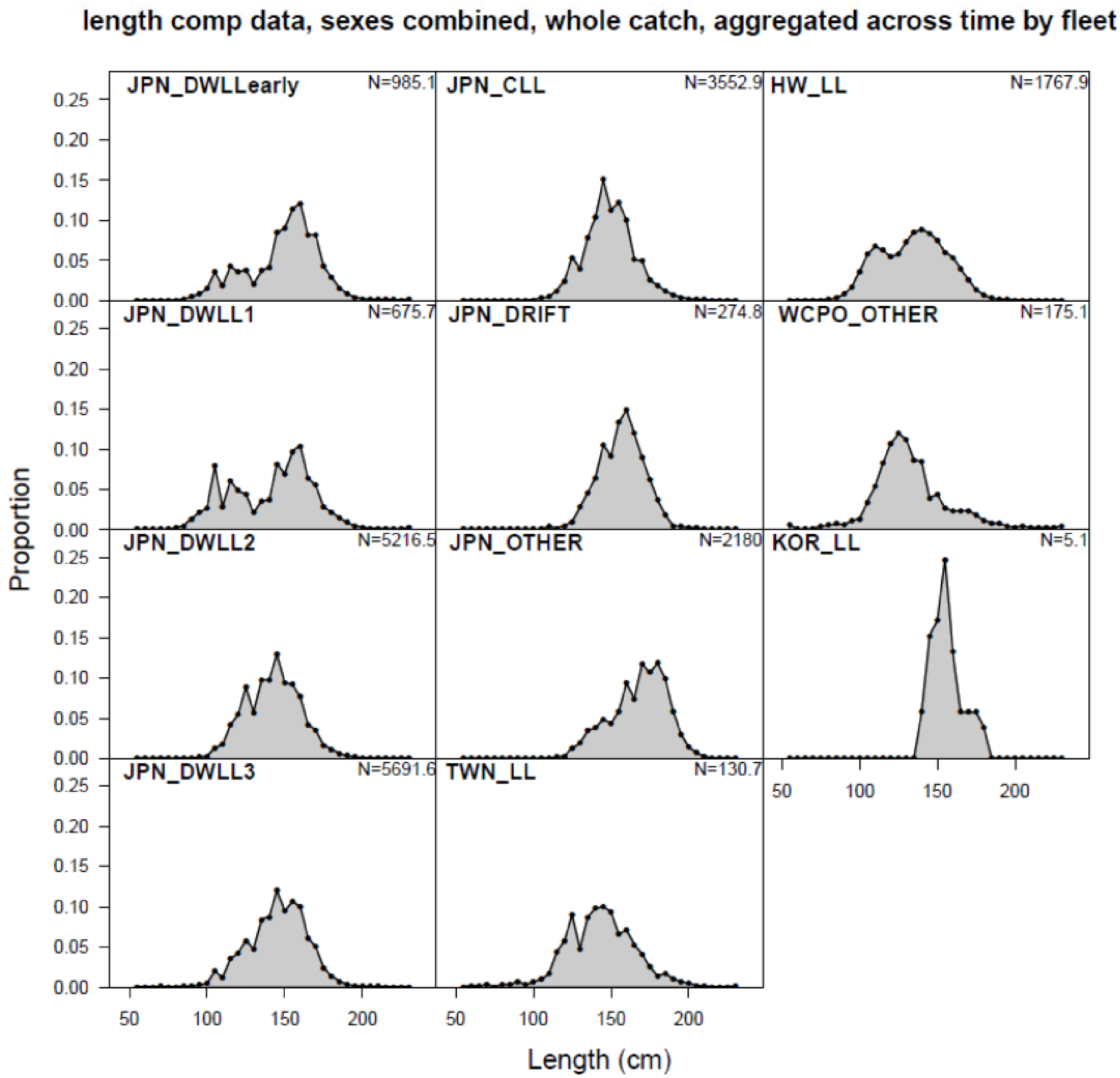


Figure 5. Observed proportion at length from all fisheries reporting size information. Samples were aggregated across year and season within fishery. Units are eye-to-fork length.

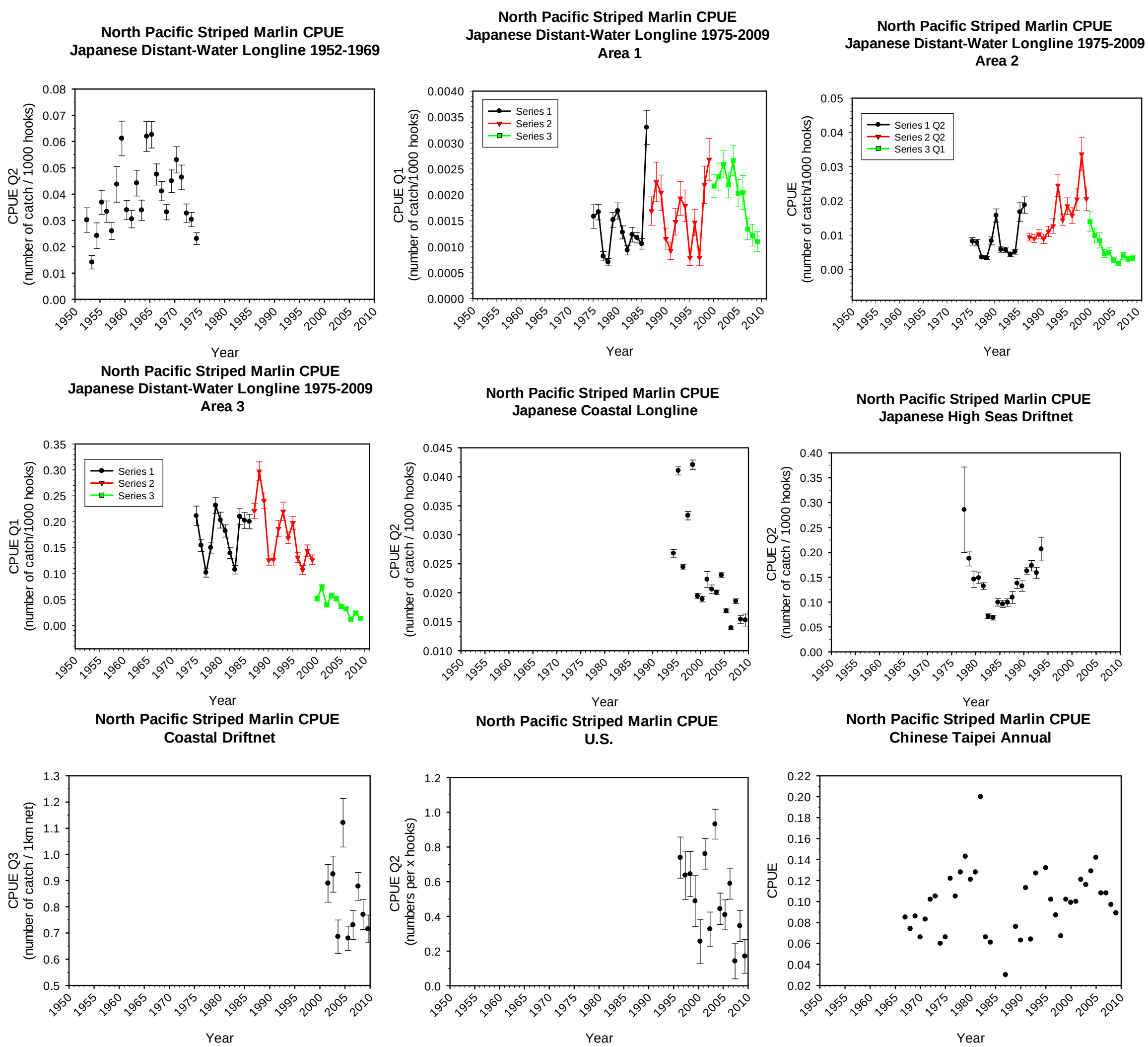


Figure 6. Plots of CPUE by fishery. For the Japanese Distant water longline fisheries there are 3 separate CPUE series (by time period) reflecting a change in fishery practices. Error bars represent 1 SD and do not include the addition of process error.

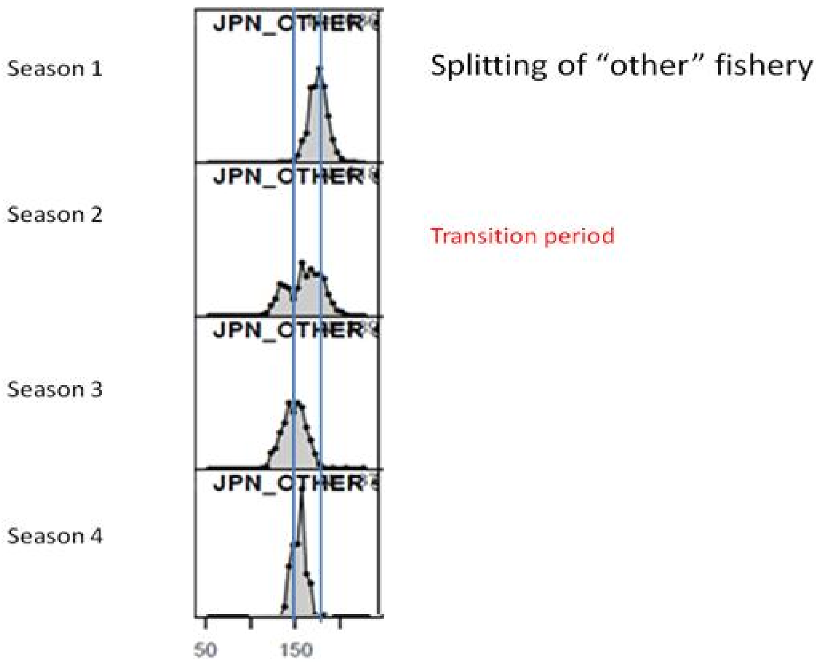


Figure 7. Seasonally summed size composition from the Japanese Other fishery showing the two size-distinct groups of fish (early and late).

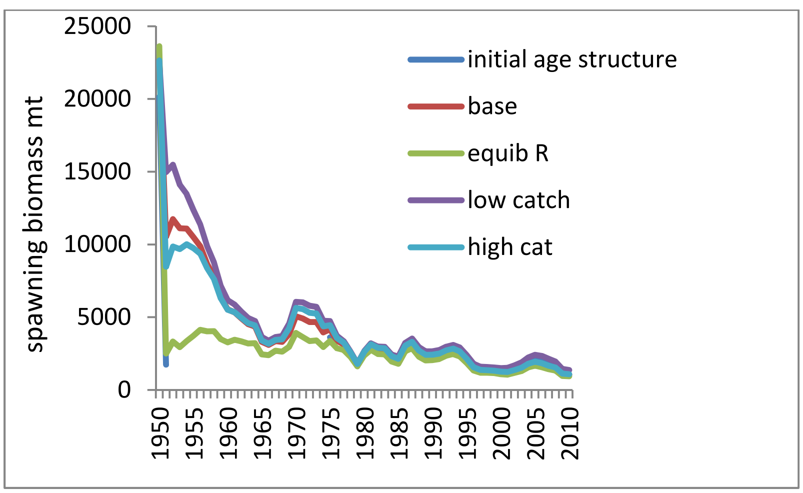


Figure 8. Spawning biomass trends by starting conditions. Initial age structure indicates a model started in 1975 without fitting to initial catch and with initial age structure estimated. Base indicates mode with 100,000 fish and 280 tons of equilibrium catch. Low and high catch are 1/3 more and less than base. Equib R indicates results from a model that decouples the equilibrium recruitment from the S/R curve.

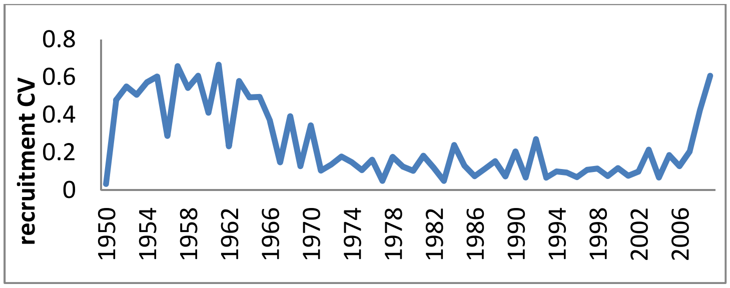


Figure 9. Estimated recruitment deviation uncertainty from 1952-2010.

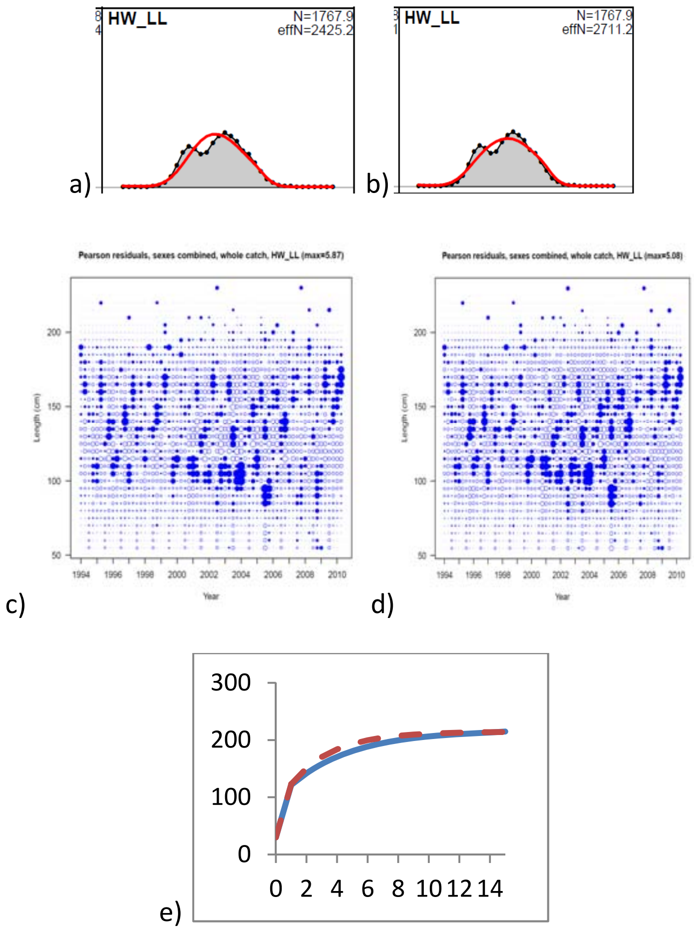


Figure 10. Observed and expected proportion at length from the Hawaii Longline fishery. Results are from a model that included all information (Full) and a model that included only the Hawaii longline proportion at size information and CPUE in fitting (Reduced). In the Reduced model growth was also estimated. Panel a) shows the combined (across years and seasons) composition fit to the size data from the Full model and Panel b) shows corresponding results from the Reduced model. Panel c) shows the Pearson residuals for the Hawaii longline composition from the Full model and Panel d) gives residuals from the Reduced model. Panel e) depicts the assumed growth in the Full model (solid) and the estimated growth from the reduced model (dashed).

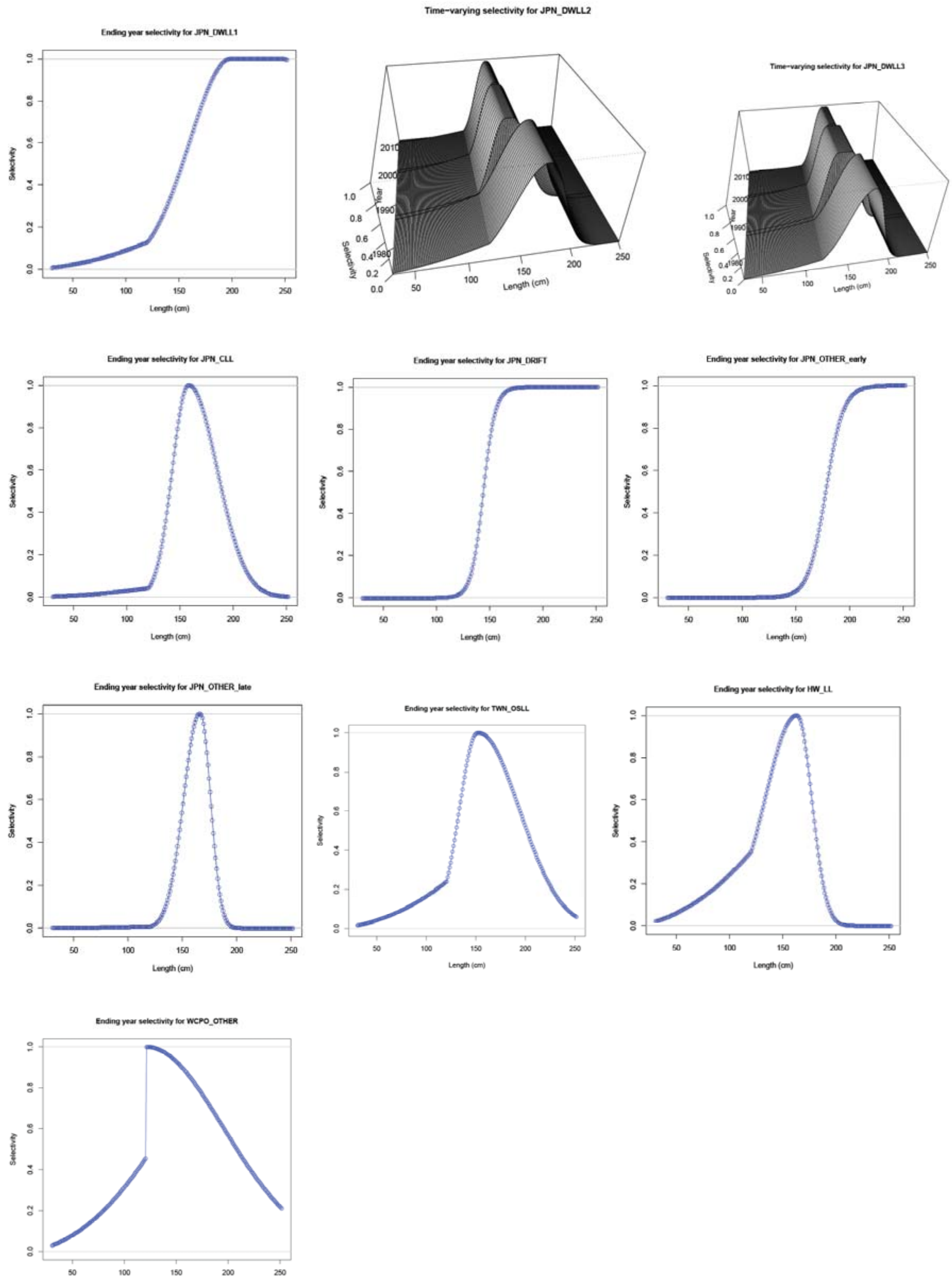


Figure 11. Estimated selectivity patterns by fishery (model 4). Fisheries with time varying selectivity patterns are displayed in 3-D plots.

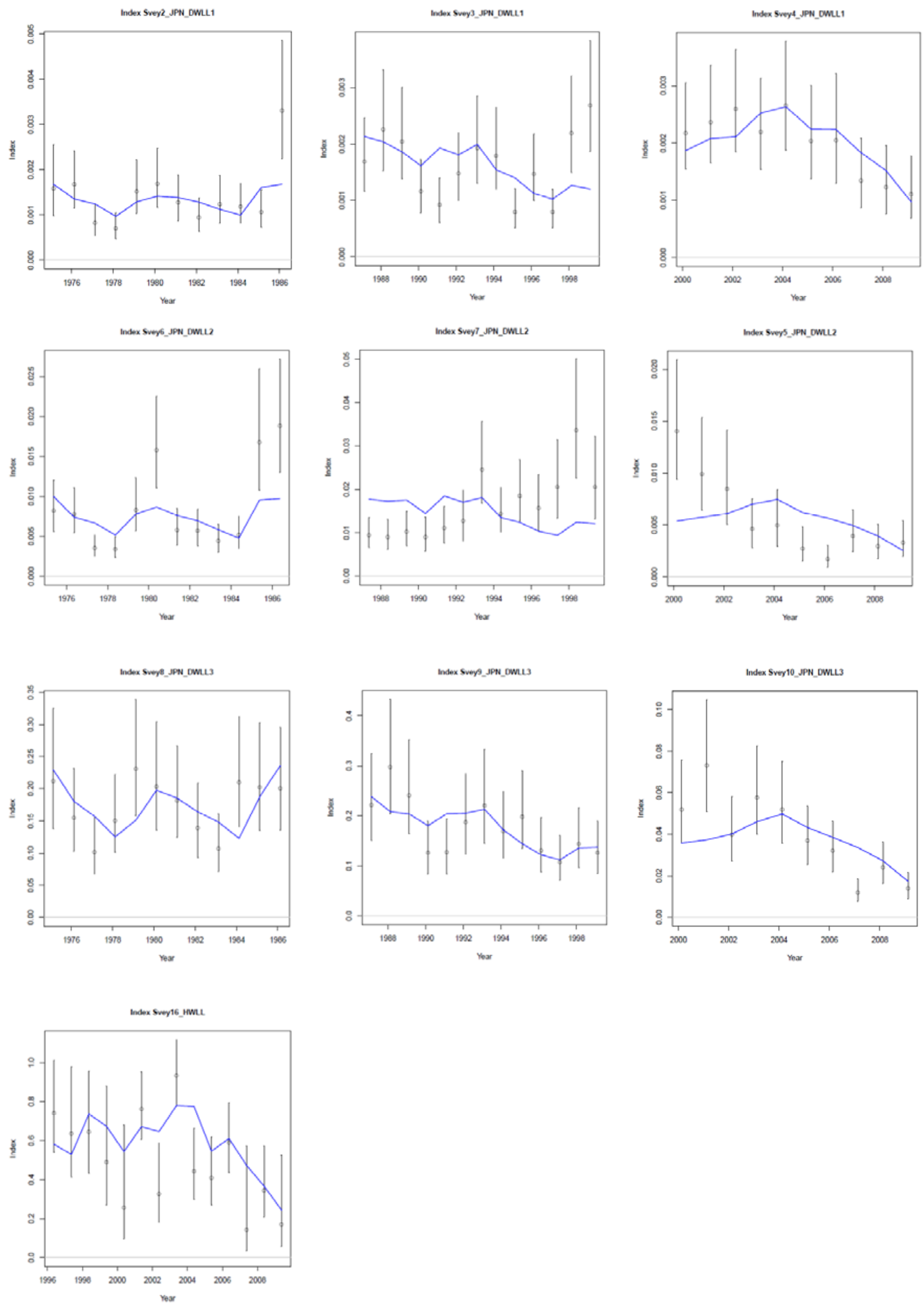


Figure 12. Observed (dots) and predicted CPUE (lines) for the Japanese DWLL fleets areas 1-3 and the Hawaiian LL fleet (model 4). The vertical lines represent the inputted CV around the observations.

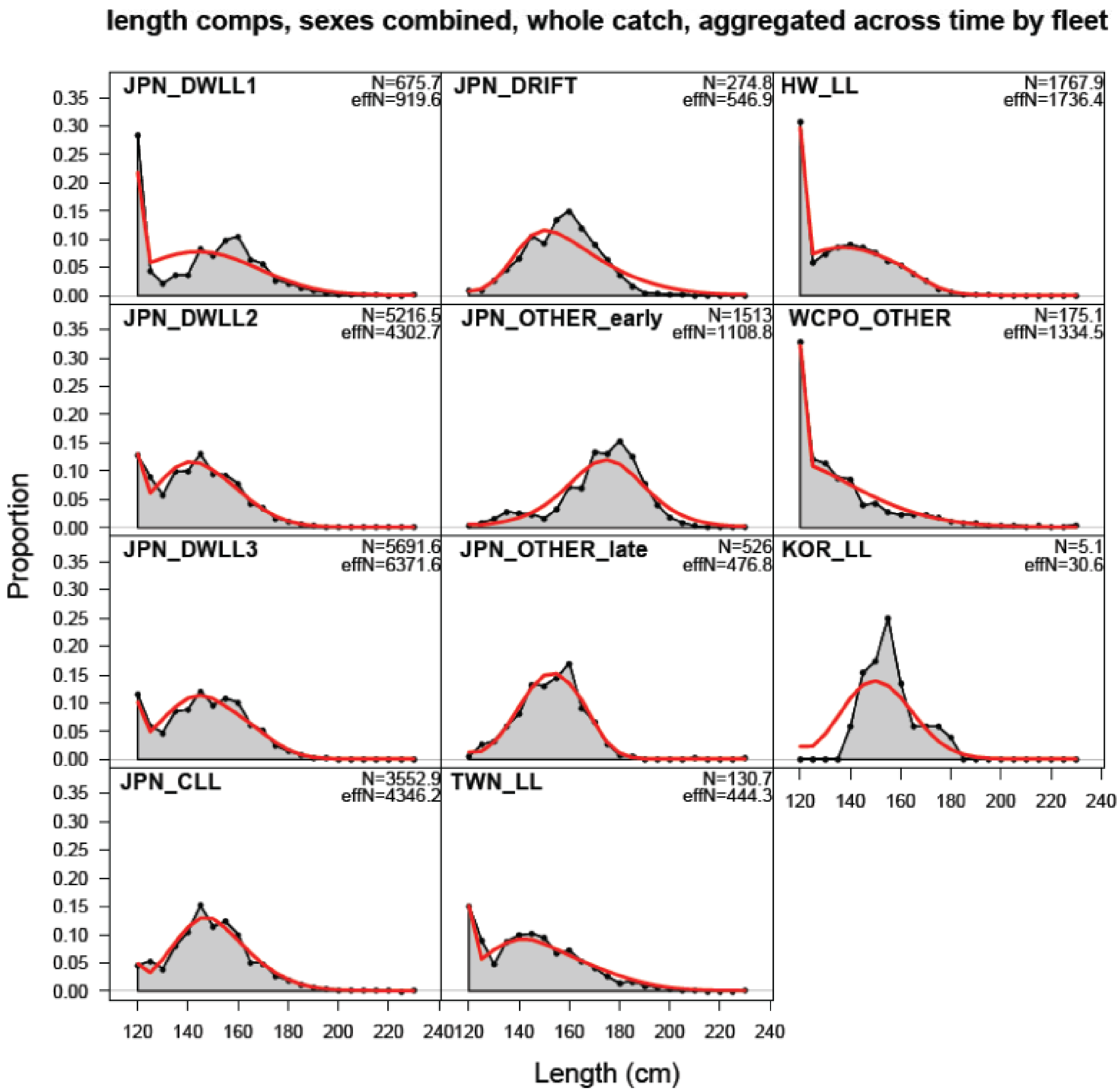
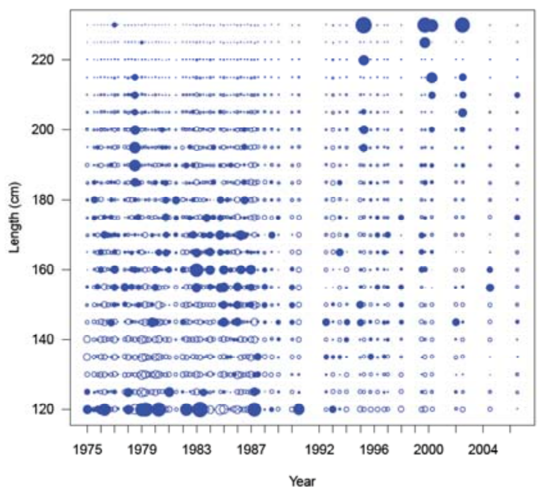
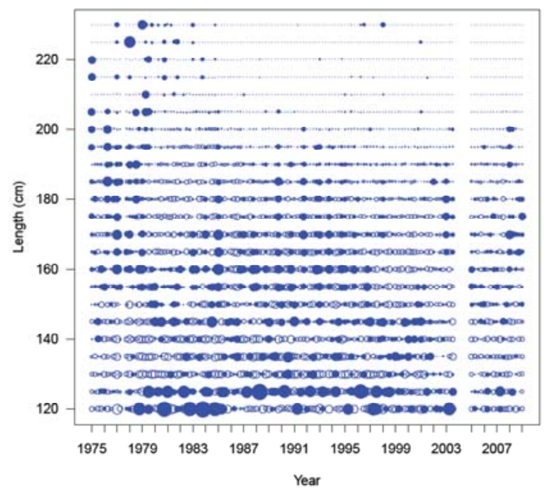


Figure 13. Plot of combined (across year and season) observed and predicted size composition by fleet (model 4).

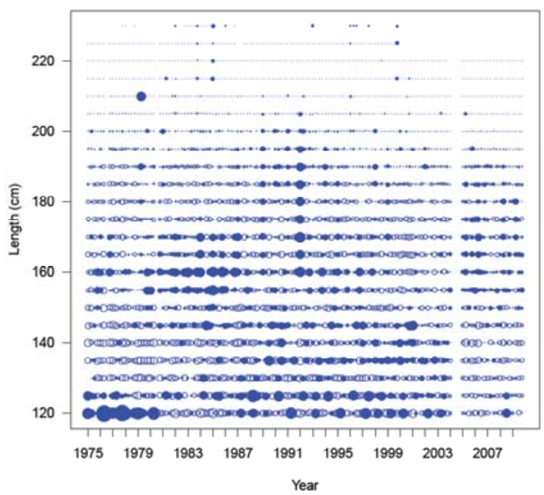
Pearson residuals, sexes combined, whole catch, JPN_DWLL1 (max=7.12)



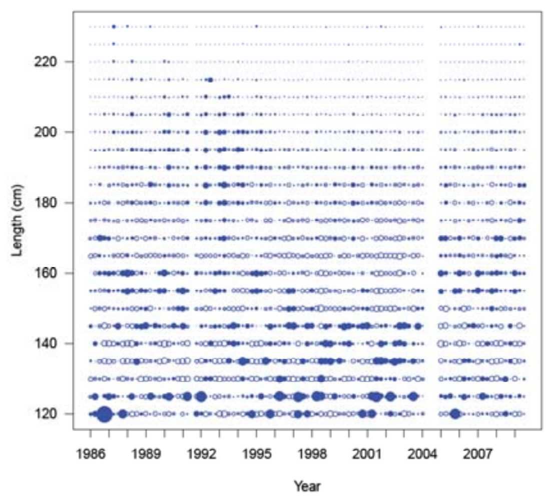
Pearson residuals, sexes combined, whole catch, JPN_DWLL2 (max=8.83)



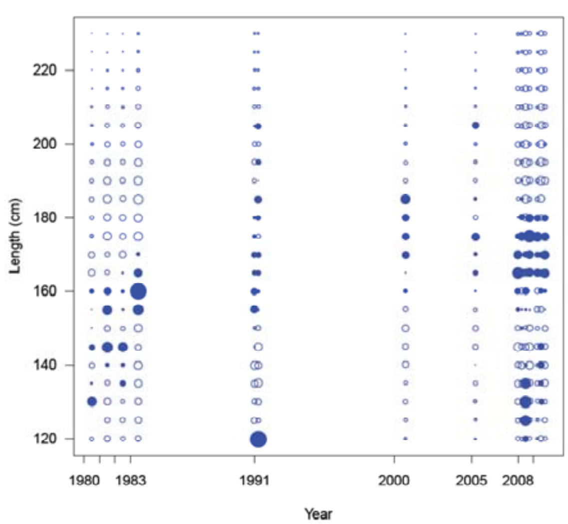
Pearson residuals, sexes combined, whole catch, JPN_DWLL3 (max=10.17)



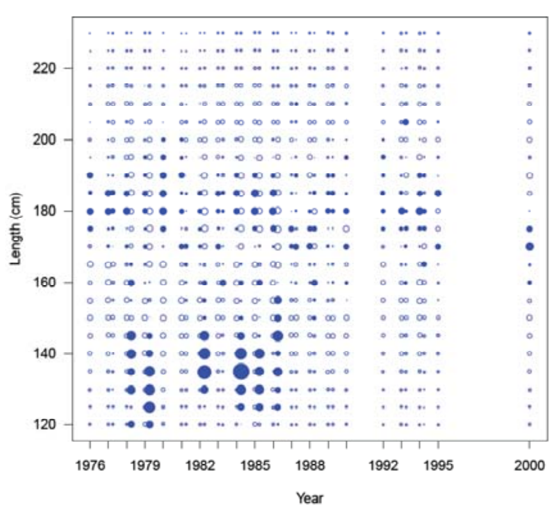
Pearson residuals, sexes combined, whole catch, JPN_CLL (max=12.8)



Pearson residuals, sexes combined, whole catch, JPN_DRIFT (max=4.13)



Pearson residuals, sexes combined, whole catch, JPN_OTHER_early (max=15.58)



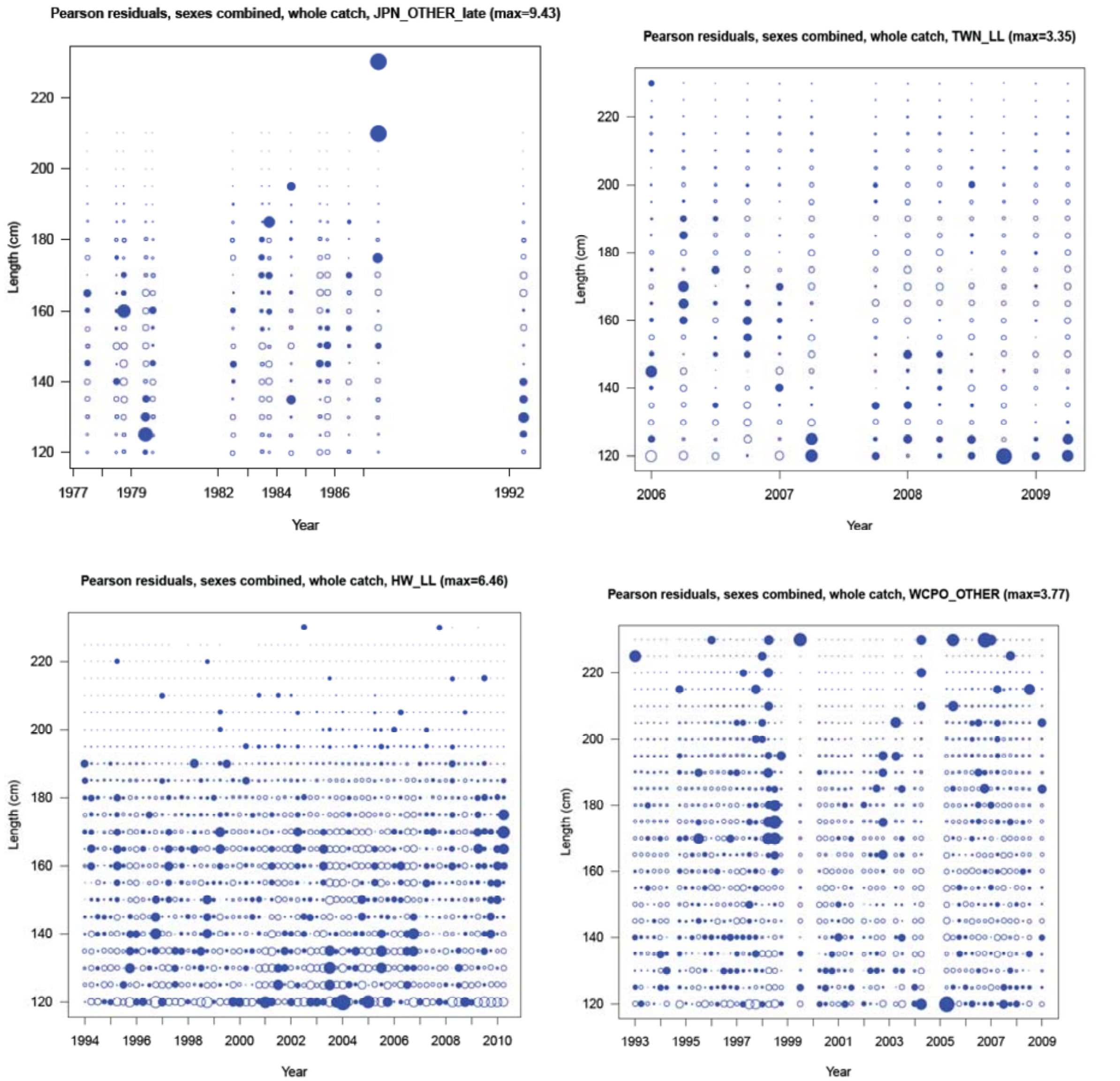


Figure 14. Pearson residuals from the fit to size composition information by fishery, season and year. Solid circles represent observations that are greater than the model predictions and open circles represent observations are less than the model predictions (model 4).

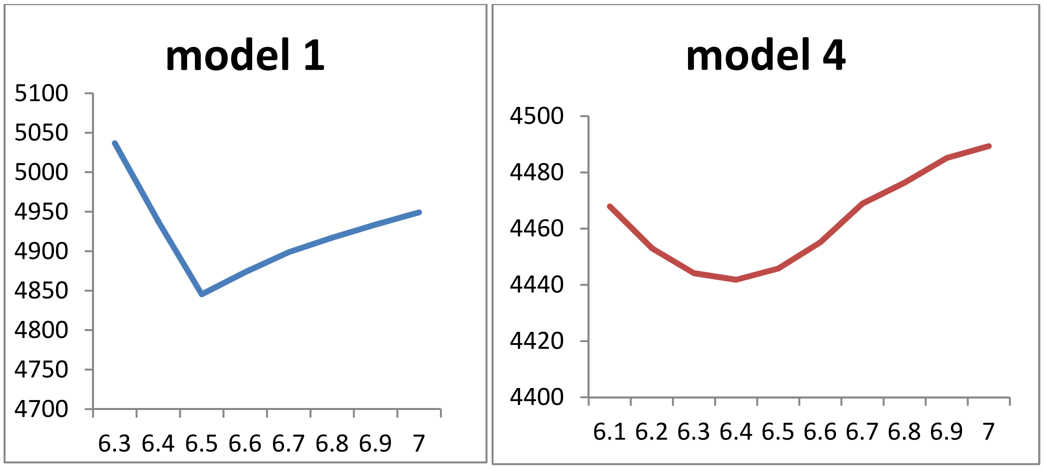


Figure 15. Total likelihood (y-axis) against fixed values of lnR0 (x-axis) from Model 1 and Model 4 .

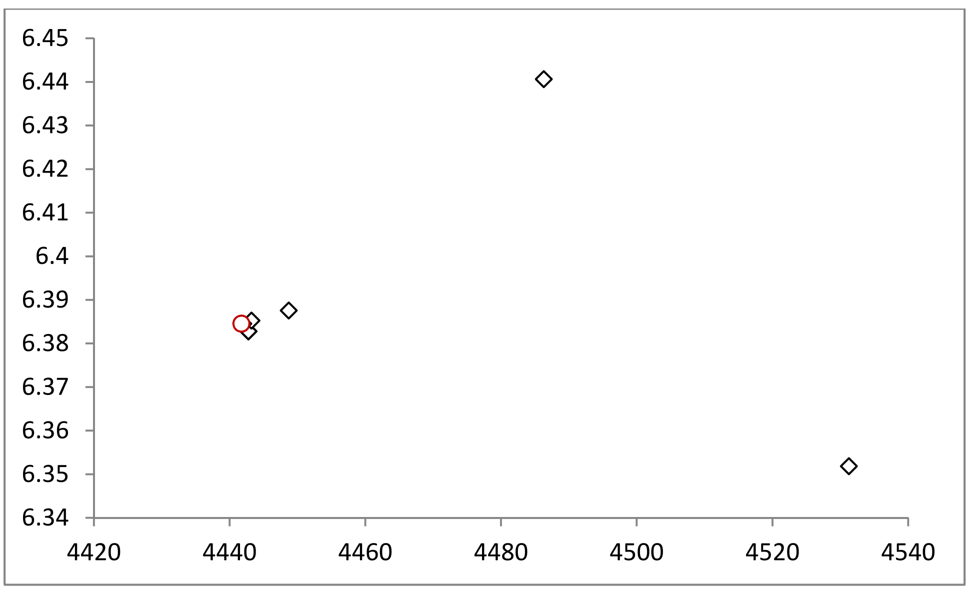


Figure 16. Plot of estimated R0 (y-axis) and total ending likelihood (x-axis) for random starting values of the model (diamonds) and the base model (circle) (model 4).

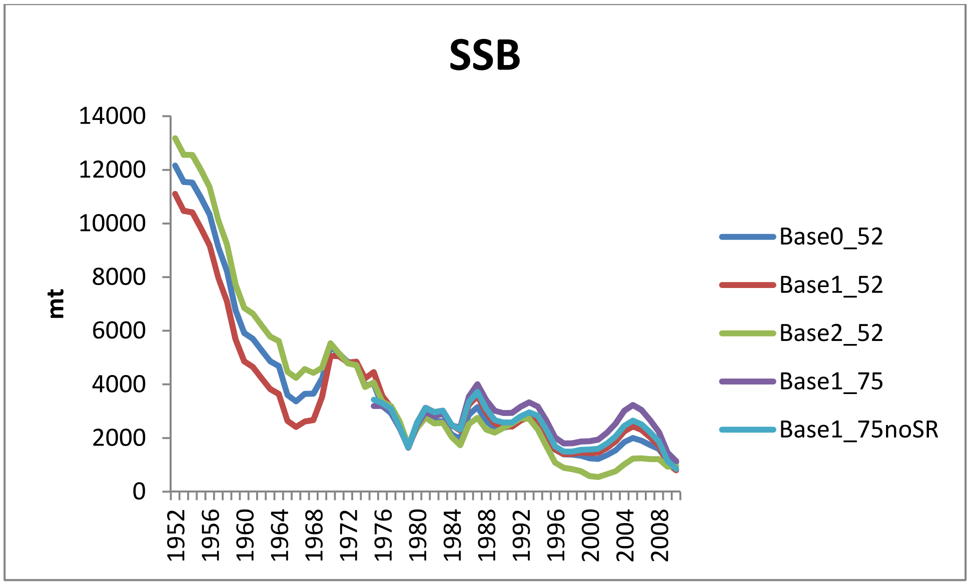


Figure 17. Plot of Spawning biomass (1952 to present) from 5 alternative models (legend shows models 1-5 in order).

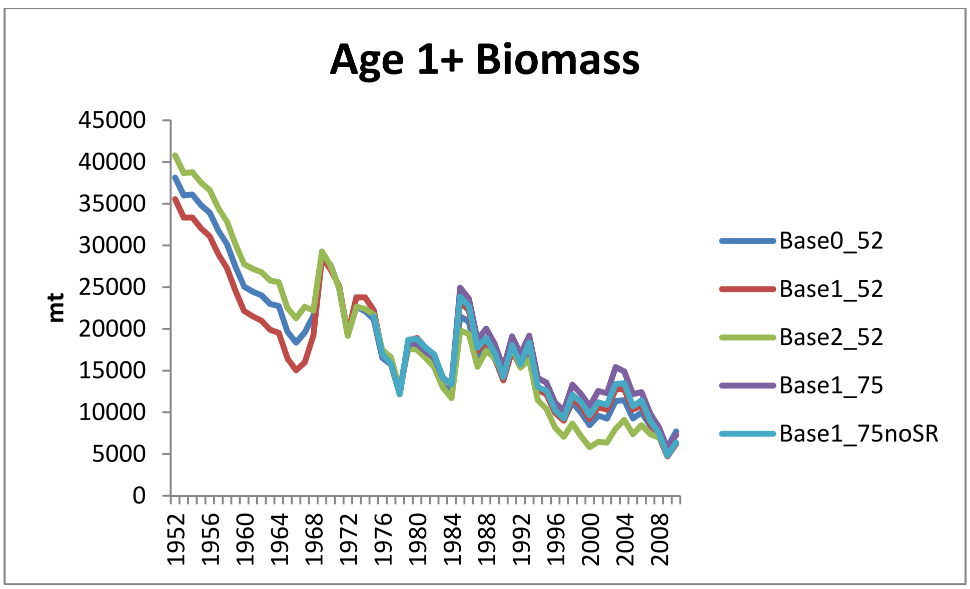


Figure 18. Plot of Age 1^{+} biomass (1952 to present) from 5 alternative models (legend shows models 1-5 in order).

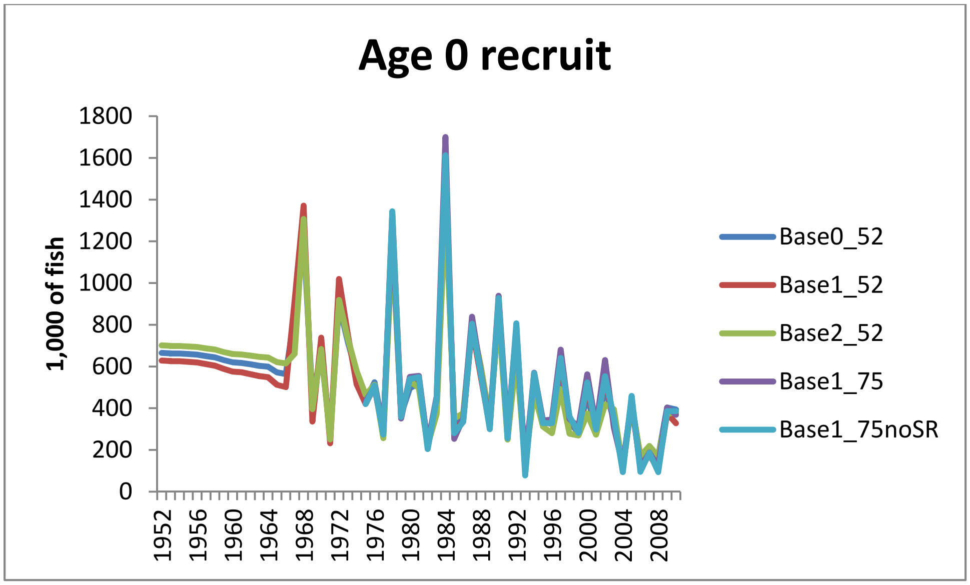


Figure 19. Plot of Recruitment (1952 to present) from 5 alternative models (legend shows models 1-5 in order).

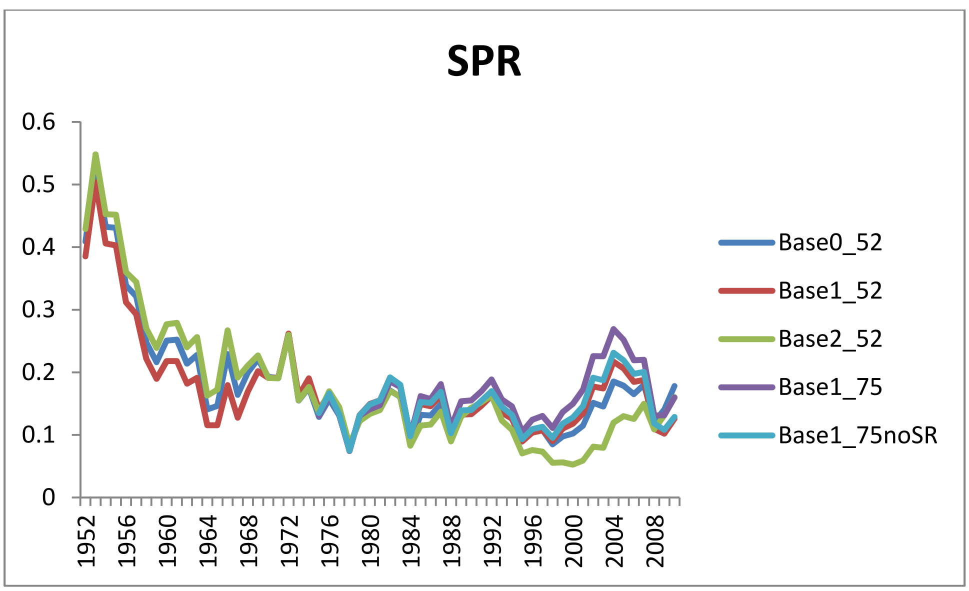
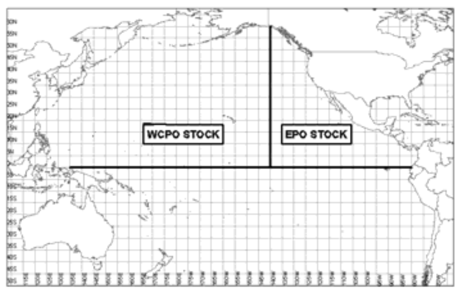


Figure 20. Plot of fishing intensity as SPR from 5 alternative models (1952-present).

Introduction

Prior ISC stock assessments of striped marlin in the North Pacific indicated a declining stock experiencing recent fishing mortality on spawners in excess of 0.7^{-yr}. Despite evidence of high fishing pressure, it was noted that there was considerable uncertainty regarding the basic biology of the stock. In particular, the stock structure, spawner-recruit resilience (h) and the growth of the species in the western North Pacific were highlighted as important areas of uncertainty.

Since the last stock assessment, considerable work on the biology of the species has been completed. Based on genetics analyses the stock boundaries were changed to reflect a Western and Central North Pacific Ocean (WCPO) stock and a separate Eastern Pacific Ocean (EPO) stock (ISC 2010).



New research has improved our understanding of growth (Sun et al. 2011a) along with new size at maturity for the same area (Sun et al. 2011b). Data for the major fisheries (DWLL Japan) were recompiled in the primary fisheries by different geographical boundaries (Area 1: 0-10°N latitude by 100°E-160°W longitude; Area 2: 10-50°N latitude by 100°E-160°E longitude; Area 3 : 10-50°N latitude by 160°E-160°W longitude) along with different time periods and updated until 2009-2010.

The objectives for this paper were to take the data series for the WCPO stock along with the new understanding of life history and explore this information inside a length-based age structured stock assessment model. We attempted to use objective means to develop model and data configurations that reduced conflict between data series. Results from this work should provide guidance for BILLWG considerations in modeling North Pacific striped marlin stock status.

Materials and methods

Overview of Methods

The methods overview will be divided into 3 sections. Section 1 describes data, Section 2 describes the preliminary analyses that were done on early model configurations to help develop base model configurations, and Section 3 describes alternative model configurations.

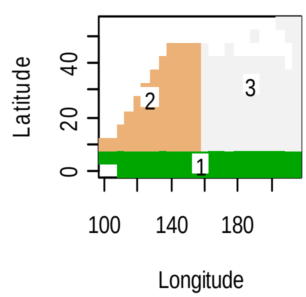
Section 1. Data

Life history

Life history information for this assessment has been taken from Sun et al. (2011a and 2011b). The combined sexes length at age relationship was based on otoliths from a maximum of age 6 fish and back-calculated lengths at age for younger ages (Figure 1). We assumed that CV on age 0.3yr fish =0.14 and age 15 yr=0.08. The assumption of the larger uncertainty in the length at age of young fish was consistent with the ageing study. The large uncertainty in the length at age of young fish is also stems from the extra variance of disparate timing of recruitment, spatial variability in growth and sexual dimorphism (although we note that the best scientific evidence does not show sexual differences in growth). Weight at length is also taken from Sun et al. (2011a). Maturity at length is based on Sun et al. (2011b) but is refit using the parameterization used in the SS3 model (Figure 2). Natural mortality (Figure 3) and steepness (h=0.87; σ_r=0.6) are the BILLWG consensus values (Piner and Lee 2011; Brodziak 2011). Observations provided in lower jaw fork length (LJFL) were transformed to eye fork length (EFL) based on Sun et al. (2011a).

Likelihood components

Likelihood components are the data that comprise the observation subcomponent of the assessment model. The three primary likelihood components are: 1) catch by fishery, 2) proportion at length of the catch by fishery, and 3) catch per unit effort (Table 1). Because of the broad spatial extent of the Japanese Distant Water Longline (DWLL) fleet, those data were divided into 3 areas resulting in 3 fisheries with all 3 types of data (catch, size composition and CPUE):



spatial distribution of JPN DWLL fisheries

Catch: Catch was inputted into the model quarterly (calendar year) from 1952 to 2010 for 19 individual fisheries in either numbers ('000s) or weight (mt, Figure 4). We assumed catch was well known and thus it was fitted with a standard error (SE) 0.05 assuming a lognormal error distribution. We defined a poor fit to catch as models that did not remove >99% of the total catch from any fishery. Catches in 2010 were also included, using observed values when available and otherwise assuming they were equivalent to the 2009 catch.

Size Compositions: Observations of the proportion at size of the catch were included from 11 total fisheries (Figure 5). The fit to the proportion at length data assumed a multinomial error distribution with the variance described by the effective sample size (effN). Approximations of the effN were taken from an analysis of the Hawaii longline fleet (Courtney unpublished) which found ~10 fish per trip. Thus for all longline fleets effN was assumed to be number of fish measured/10. A maximum quarterly sample size was assumed to be 50. Fishery average effN are given in Table 2.

CPUE: 16 CPUE series were included and fit with a lognormal error assumption (Figure 6). The SE in log space was assumed to be the same as the CV (typically SD/estimate) described in each CPUE paper.CV was assumed to be equal to 0.2 when missing. Series with average CV<0.2 were scaled to CV=0.2 through the addition of a constant. Series with average CV>0.2 were inputted as given. Missing CVs were assumed =0.2 except for the Taiwanese longline CPUE from the early period, which was given a larger CV in accordance with the BILLWG recommendation (Table 3).

Section 2. Preliminary Analyses

This section describes preliminary analyses that went into building the base model. It includes the methods used to determine assumptions of the model. We start from analyzing appropriate larger scale structure (things unlikely to change with other changes to model) and focus on smaller scale structure at the end.

To start, a basic model was developed that described how the data was built and used the life history specifications listed above. The model was a single area model (selectivity patterns accounting for spatial patterns) that included fitting to quarterly data (Jan-Mar quarter 1 etc.). A separate selectivity pattern was fit for each fishery described above with observed composition information. Fisheries without proportion at length observations were assumed to have the same selectivity as a similar fishery as described by the BILLWG (Table 1). Time blocks to selectivity patterns (time varying) were allowed for Japanese DWLL fisheries corresponding to the time periods when CPUE catchability was assumed to change as a result of fishery practices. Ages 0-15 were included in the modeling.

<u>Assessment model</u>

This paper used Stock Synthesis (SS) as the population dynamics model. SS is a stock assessment model that estimates the population dynamics of a stock through use of a variety of fishery dependent and fishery independent information. Although its use has historically been for groundfishes, more recently it has gained popularity for stock assessments of tunas and other migratory species in the Pacific Ocean. The structure of the model allows for Bayesian estimation processes and full integration across parameter space using the Markov Chain Monte Carlo (MCMC) algorithm.

SS3 is composed of 3 subcomponents, 1) population subcomponent that recreates an estimate of the numbers/biomass at age of the population using estimates of M, growth, fecundity etc., 2) an observational sub-component that consists of the observed (measured) quantities such as CPUE or proportion at length/age, and 3) a statistical sub-component that quantifies the fit of the observations to the recreated population using likelihoods. For a complete description see (Methot 2005, 2010). This analysis uses version 3.20b.

Splitting of Japanese "Other fishery"

Preliminary data analysis showed that fishery 12 (Japanese Other fishery) contained a strong seasonal pattern in observed sizes of striped marlin (Figure 7). Larger fish are taken in the first two seasons of the year. Although some seasonality can be observed in most fisheries, this pattern was considered the most problematic, and initial modeling attempts could not adequately address the issue. We decided to split the data (catch and size) from this fishery into two seasonally-specific fisheries to reduce the influence of the misfit. This was important as this fishery (seasons 1 and 2) included observations of the largest fish and would likely be our assumed asymptotic fishery (See section below on Selectivity assumptions). We note that season 2 included both larger and smaller mode fish, but preliminary model runs showed more selectivity pattern stability if season 2 was included with season 1. All further exploration described below included the Japanese Other fishery broken into two separate fisheries: early (seasons 1-2) and late (seasons 3-4).

Determining spawning season and recruitment timing

Spawning was described in Sun et al. (2011b) as taking place from late spring throughout summer (May-August) based on histological examination. We assume that spawning biomass estimates used in the spawner-recruit model occur in season 2 (beginning of spawning cycle). Recruitment timing was assumed in the model to occur in season 3 (July-Sept) on the basis of best model fit of preliminary models (Table 4). Importantly season 3 recruitment timing showed improved fit to fisheries 1, 2 and 17 (Japan DWLL early, Japan DWLL area1 and HWLL) which take age 0 fish.

Selectivity assumptions

First, we assume that at least one fishery with observed size compositions has an asymptotic selectivity pattern and that all selectivity patterns are length-based. All other fisheries would then be allowed to be as domed, as best fits the data. This assumption means that at least one of our observational tools samples the entire population after a specific size (eliminating the possibility of model generating cryptic biomasses). This is a strong assumption and will be influential, thus the choice of the asymptotic fishery was evaluated with extensive testing. However we feel this is necessary, as models that do not have at least one asymptotic fishery often scale to unreasonable biomass levels with very good fit. We note that age-based selectivity is also invoked that allows age 0-15 fully selected for JPN DWLL early, area1, HWLL and the WCPO other fisheries. All other fisheries were considered to select only ages 1-15.

The testing for the most consistent fishery data with an asymptotic selectivity assumption consisted of sequentially assuming each fishery was the asymptotic fishery and allowing all others to be domed shaped. Because of model instability in some runs, we later changed the procedure to force two fisheries to be asymptotic (e.g. run 1 fishery 1 and 2 asymptotic, run 2 fishery 1 and 3 asymptotic etc.) for all combinations. The fishery with the best fit with assumed asymptotic selectivity pattern (in combination with other fisheries assumed asymptotic) was chosen as the asymptotic fishery. Best fit was evaluated by average total likelihood and rank (Table 5). We performed these tests across different growth assumptions and equilibrium catch assumptions. Our results indicated that the Japanese Driftnet and Japanese Other early fisheries were consistently best fitting with the asymptotic assumption. Although the driftnet result was surprising, the JPN Other Early fishery also had strong support in our testing and visually it contained the largest fish (Figure 5). Thus, we concluded that both the Japanese Driftnet and Japanese Other Early fisheries removals would be assumed to be described by an asymptotic selectivity pattern. All model runs describe from this period forward contained this result. All other fisheries are allowed to be as domed as best fit the data. We note that a few other fisheries, notably Taiwan LL, Other late and WCPO Other also had support for asymptotic selectivity patterns.

Equilibrium catch

Equilibrium catch (assumed catch prior to start of model) was set at ~75% of 1952-1953 catch. We chose only the earliest years as there was a noticeable increasing trend in catch through the 1950s. For that reason we also chose only a significant fraction of the early observed catch. We segregated catch into two fisheries: 100,000 fish into Japan DWLL early and 280mt into Japan Other early. We ran models with both 100% (high) of early catch and 50% (low). The results of the high/low catch analysis indicated that higher equilibrium catches started the model at lower levels in 1952 and lower equilibrium catches at higher levels, but by the mid 1970's (when better data were available), the models produced similar dynamics (Figure *8*). We took this result to indicate that equilibrium catch settings primarily affected dynamics prior to the start of informative data in the 1970s and not final results.

Recruitment period

For parsimony, we chose to estimate recruitment deviations from the assumed spawner recruit curve only for years with information on recruitment. For all other years recruitment would be taken from the expectation of the S/R curve (no deviations). The years with information on recruitment was based on a preliminary model run with all recruitment deviations estimated (1952-2010). The CV of the recruitment estimates was plotted and it was assumed that information was available to inform recruitment magnitude when the CV's stabilized (Figure 9). In our case it appeared that information was available to inform recruitment starting in 1967 through 2008. Thus we now estimated recruitment during that period and used the S/R expectations for all other years. A more complex modeling process that changes the bias adjustment to account for lack of information could be used allowing for estimation of all recruitment deviations. Although this mostly affects the estimation of uncertainty, it is an area for more model development.

Lower bin bound (55 vs 120cm)

The majority of the size data supplied were binned from 55-255cm (by 5cm bins). Preliminary analysis indicated that data from 55-120cm (age 0) were both noisy and inconsistent with our model structure. We gained significant improvement in fit to proportion at size data by increasing the accumulator first bin to 120cm (accumulating all smaller fish, thus no loss in data). This improvement in fit reduced the gradients in the negative log-likelihoods from profiling across R0 for domed shaped fisheries, which we interpreted as removing misfit influence on the model results. For these reasons, we chose to include bin structure of the observations starting at 120cm- 230cm (by 5cm bins).

This decision was influenced by separate analysis that tried to explain the Hawaii longline composition information using recruitment and growth. A model was run that fit to only the composition information from the Hawaii longline fleet (best information on very young fish) and all CPUE and growth was estimated. This model explored if variation in recruitment and an alternative growth model could better explain the size data from primarily the age 0 group without the confounding influence of the other size composition information. In other words, if we could not adequately fit the age 0 size composition from that one fleet we have no hope in a model that included all fleet composition information. Results (Figure 10 (a)-(d)) indicated that we cannot adequately explain the variability in size at age 0 in the Hawaii longline samples with a time invariant but estimated growth form and variability in recruitment. This is likely due to both time varying growth and variations in spawning timing (or recruitment timing). Although changes in both growth and spawning period are possible in the model, it was deemed too complicated for implementation at this time. Changes in fishery practices may also be responsible for the misfit, but the additional process of variable year/season-specific selectivity patterns were beyond the scope of this assessment. This result reinforced our decision to bin the age 0 fish into a single bin to reduce the contribution of the misfit (due to inadequate model process) to the model results.

CPUE series to include

First we assumed that the likelihood components of both catch and proportion at length should be included in all models. However, we have 16 CPUE series which are assumed to represent the change in the segment of the population described by the selectivity pattern estimated for that fishery. It is apparent that some of the CPUE are contradictory information and thus should not be used together in a stock assessment model. We attempted to objectively segregate the CPUE series into two separate data sets based on a down-weighting analysis.

In the analysis, we sequentially down-weighted each likelihood component (excluding catch) in separate model runs. We assumed that CPUE derived from the same fishery (e.g., all 3 area 1 CPUE) described the population trajectory, and thus were included/excluded together. This analysis was performed for different assumptions (e.g. growth, equilibrium catch etc) and summarized across these assumptions. CPUE series were determined to go together if down-weighting those series led to loss of fit in the other series. Our results indicated that the Japanese DWLL area1, area 2 and area 3 and HWLL were consistent. The other series including Japan DWLL early, Japan Coastal LL and Japan Driftnet, along with Taiwanese LL early and late, were considered the alternative data set. Thus two new data sets were developed.

Data set 0 included all CPUE
Data set 1 included Japan DWLL areas 1, 2 and 3 and Hawaii LL.
Data set 2 included Japan Coastal LL, Japan Driftnet and Taiwanese LL.

Starting year

Given that the model estimation of biomass dynamics prior to the mid 1970's is influenced by our assumptions of equilibrium catch, we chose to produce a model that only fit to data from 1975-2010. In this model we estimate the initial age structure and freely estimate the fishing mortality consistent with that age structure (avoiding the equilibrium assumptions).

Based on the preliminary analyses, we explored 5 models based on data (CPUE), starting conditions and spawner recruit assumptions. Each potential model was evaluated based on the consistency of results and goodness of fit to data.

1. Model 1. Use data set 0 (all CPUE), start in 1952, fit to equilibrium catch (100K fish and 280mt) and S/R relation.
2. Model 2. Use data set 1 (subset CPUE), start in 1952, fit to equilibrium catch (100K fish and 280mt) and S/R relation.

3. Model 3. Use data set 2 (subset CPUE), start in 1952, fit to equilibrium catch (100K fish and 280mt) and S/R relation.
4. Model 4. Use data set 1 (subset CPUE), start in 1975, freely estimate equilibrium F and initial age structure and S/R relation. (Note that this model reflects the addition of a more flexible time varying selectivity pattern parameterization which was implemented during the BILLWG meeting rather than in the original draft of this working paper).
5. Model 5. Use data set 1 (subset CPUE), start in 1975, freely estimate equilibrium F and initial age structure with recruitment deviations treated as free parameters.

Section 3. Preferred Model Configuration

Our preferred model is model 4, with the following characteristics.

Start year: 1975

Data sets: data set 1 including all catch, all proportion at size, and CPUE data set 1.

Initial conditions: estimate initial age structure and fishing mortality

Asymptotic selectivity assumption: Japanese driftnet and Other early fishery;
3 time blocks for selectivity pattern for the JPN DWLL areas 2 and 3;
Only JPN DWLL area 1, HWLL and WCPO Other fisheries are allowed to take age 0 fish, all other fisheries are constrained to taking age 1+.

Results

In this section we describe model results for our preferred model (model 4) with comparison to the other models in some derived quantities:

Estimated selectivity patterns are given in Figure 11.

Observed and estimated CPUE by fishery are given in Figure 12.

Observed and predicted proportion at length are given in Figures 13 and 14.

Likelihood profiles across the parameter R0 (unfished recruitment the primary scaling parameter) is given are Figure 15 (total likelihood), Table 6 (size composition components) and Table 7 (CPUE components). For comparison, a likelihood profile for Model 1 (start in 1952 with all CPUE included) is also given. Likelihood values from the best fit to model 4 are given in Table 8.

Convergence to a global minimum was examined by randomly perturbing the starting values of all parameters by 10% and refitting the model (Figure 16). There is no evidence of a better fit.

Estimated spawning stock biomass and age 1+ biomass are given in Figure 17 and Figure 18. A comparison to the other candidate models is depicted. All models indicated a stock with biomass at levels below their long-term average.

Estimated spawning recruitment is given Figure 19. A comparison to the other candidate models is depicted. All models indicated a decline in recruitment over the last decade.

Estimated SPR is given Figure 20. A comparison to the other candidate models is depicted. During the period of informative data, exploitation has typically allowed 10-20% of the spawning potential. The most recent years have been closer to 12%.

Table 9 gives yearly estimated values of spawning biomass, recruitment and SPR and their asymptotic standard deviation estimates.

Discussion

We believe that the reduced model (starting in 1975 and subset of CPUE indices) is preferable to the more full models because of the limitation of model conflict apparent in the R0 profiles and because data after 1975 are, in our opinion, much more reliable (e.g. HPB available for standardization of CPUE, size data available etc.). In our preferred model, gradients of likelihood are minimized for the size composition data and indices are influential in the results. Consequently, the fit to the indices (data set 1) are acceptable except for the middle period in the JPN DWLL area 2, which is a conflicting trend to all other areas. The authors also believe that estimation of the starting conditions is preferable to specifying those same conditions. A quick check of the validity of this approach is the model's estimate of 8500 tons prior to the starting of the dynamics in 1975 is within the range (~6000-10,000 mt) reported in the decade prior to 1975.

It appears that the model generally follows the Japanese Area 3 DWLL indices, especially in the period 1975-1999. While we do not know if this is correct, but we note that the majority of the LL catch since 1975 has come from this area. Had we chosen a priori which indices to believe, it is likely we would have chosen area 3 as our primary indices. This is due to the magnitude of the catch, location of biomass and that Japanese longline fisheries are often the most trusted data source for CPUE information. This decision would have been consistent with the prior assessment philosophy as well.

We note that time varying selectivity was not allowed for Japan DWLL area 1 size information, which was inconsistent with how the CPUE were treated as 3 independent time series. It was also inconsistent with the treatment of the other JPN DWLL areas (3 time blocks of selectivity). However, preliminary runs showed selectivity instability in area 1. In models that include data back to 1952 we shared selectivity parameters from JPN DWLL area 1 and the JPN DWLL early fisheries for the same reason. Sample sizes in Area 1 are small and the data is both bimodal and noisy. It was possible to use spline functions to better approximate the bimodal distribution of sizes seen in this fishery, however it was our opinion that the bimodality was due to sampling issues and was not likely representative of the catch. Therefore we continued to use a functional form of selectivity that we felt better represented the performance of the gears. The likelihood profile across R0 for this fishery did not indicate a strong gradient of misfit, indicating any misfit for the for Japan DWLL area 1 size had little influence on model results.

The estimated selectivity patterns for most longline fisheries are decidedly dome-shaped. Whether this result reflects gear operations (such as depth fished, bait, etc.) or is related to the spatial distribution of the fleet relative to the size structure of the population is not clear. Work to address on a finer spatial scale the location of fish size caught and fishing effort is warranted. A third possibility is that this reflects a bias is the size sampling process, but this is thought to be less likely. Uncertainty in the life history parameters (growth and mortality) is also influential in the degree of dome shape.

Fit to the size composition information is generally good, especially for fisheries with the most flexible selectivity patterns and large sample sizes (JPN DWLL area 2 and 3, JPN CLL). The greatest level of model misfit appears to be from the two fisheries with the strong assumptions of asymptotic selectivity (least flexible selectivity assumption). The model generally favors more large fish than were observed. Although the likelihood profiles across R0 indicate good general support for the relative scale, it is the asymptotic selectivity pattern assumptions that prevent scaling to unrealistic biomass levels where recruitment alone can explain changes to data. We also note here that the driftnet and other fisheries effN calculations did not include the division by 10 associated with the LL fisheries and thus true sample size may be overestimated.

In the authors' opinion, the largest uncertainty in growth (Sun et al. 2011a) is how large we should expect a typical fish to get. The growth curve we used is based on observed fish size at age 6 and back calculated size-at-age for ages <6. Sun et al. point out that we really don't have strong evidence for how large fish are likely to be after age 6. This is an area that needs more research. They also note that the maturity-at-length relation does not represent that a size class is fully matured until fish are larger than routinely seen in our data. Alternative maturity ogives should be explored.

We also note that iterative rescaling of variance has not been done for any models. Typically this is something we advocate. However our model expectations are not far from the inputted variance levels and in this exploratory work, we wished to retain the comparability of likelihoods across models. We would advocate an iterative re-scaling of the inputted variances to conform to the model's expectations on any final model developed by the working group. This iterative rescaling would effectively down-weight the CPUE from the JPN distant water longline area 2 middle period via inflated SE. If the iterative rescaling is not done we would advocate the removal of the JPN DWLL area 2 middle index to prevent its misfit from causing problems.

Finally we note that it was not the intention of this paper to characterize stock status relative to reference points, but rather to explore model and data structure to facilitate BILLWG work. However, results of our analysis suggest that the stock is being fished at a relatively high level and recruitment is below long term averages.

Literature Cited

Brodziak, J.K.T., (2011) Probable values of steepness for north Pacific. ISC/11/BILLWG/2-XX

ISC. (2010) Report of the Billfish Working Group Workshop. Hakodate, Hokkaido, Japan

Methot R.D. (2005) Technical Description of the Stock Synthesis II Assessment Program Version 1.21- March 2005

Methot R.D. (2010) User manual for SS3 V.3.11a. September 2010

Piner, K.R., H.H. Lee (2011) correction to meta-analysis of striped marlin natural mortality. ISC/11/BILLWG/2/07

Sun, C.L., W.S Hsu, N.J. Su, S.Z. Yeh, Y.J. Chang and W.C. Chiang (2011a). Age and Growth of striped marlin (Kajikia audax) in the waters off Taiwan: A revision. ISC/11/BILLWG-2/07

Sun, C.L., S.Y. Chang, Y.J. Chang, S.Z. Yeh, W.C. Chiang, and N.J. Su. (2011b) Reproductive biology of female striped marlin, Kajikia audax, in the waters off Taiwan (preliminary). ISC/11/BILLWG-1/XX

Tables

Table 1. Likelihood components considered for inclusion in the population dynamics model. Fishery number is the identification numbering inside the stock assessment model. Mirrored is a term that defines which selectivity pattern is assumed to be the same as the fishery for describing removals. Time blocks indicate the number of discreet periods where selectivity was allowed to change.

Fleet number	Fishery	Catch (units)	Composition information (years)	CPUE (number of series)	Selectivity shape (est. or assumed)	Mirrored fleet (time blocks)
1	Japan DWLL early	1952-1974 (#'s)	1970-1974	1	Domed (est)	
2	Japan DWLL area 1	1975-2010 (#'s)	1975-2006	3	mirrored	
3	Japan DWLL area 2	1975-2010 (#'s)	1975-2009	3	Domed (est)	(3)
4	Japan DWLL area 3	1975-2010 (#'s)	1975-2009	3	Domed (est)	(3)
5	Japan Coastal LL	1952-2010 (mt)	1986-2009	1	Domed (est)	
6	Japan Driftnet	1952-2010 (mt)	1980-1983,91,00,05,08,09	2	Asymptotic (est)	
7	Japan OLL	1952-2010 (mt)		0	mirrored	5
8	Japan Squid	1952-2010 (mt)		0	mirrored	6
9	Japan Bait	1952-2010 (mt)		0	mirrored	5
10	Japan Net	1952-2010 (mt)		0	mirrored	5
11	Japan Trap	1952-2010 (mt)		0	mirrored	5
12	Japan Other	1952-2010 (mt)	1972-2000	0	Asymptotic (est)	
13	Taiwan Longline	1952-2010 (mt)	2006-2009	2	Domed (est)	
14	Taiwan OSLL	1952-2010 (mt)		0	mirrored	13
15	Taiwan Coastal	1952-2010 (mt)		0	mirrored	13
16	Hawaii LL	1952-2010 (mt)	1994-2010	1	Domed (est)	
17	WCPO others	1952-2010 (mt)	1993-2009	0	(Domed est)	
18	Korea LL	1952-2010 (mt)	2005	0	mirrored	3

Table 2. Inputted and estimated sample sizes of the size composition information. Fitted values are taken from model 1 (with all indices included) and model 4 (subset of indices and years). N/A indicates not used.

			model 1	model 4	
Fleet		N	mean_effN		mean(inputN)
1	JPN_DWLLearly	20	45.4538	N/A	49.255
2	JPN_DWLL1	69	13.8579	13.3282	9.79275
3	JPN_DWLL2	131	31.4654	32.8447	39.8206
4	JPN_DWLL3	135	41.3027	47.1968	42.16
5	JPN_CLL	91	47.5169	47.7605	39.0429
6	JPN_DRIFT	15	41.3261	36.4574	18.32
12	JPN_OTHER_early	37	32.9526	32.6114	44.7027
13	JPN_OTHER_late	14	33.6786	34.0571	37.5714
14	TWN_LL	13	27.8378	34.175	10.0538
17	HW_LL	66	25.4182	26.3092	26.7864
18	WCPO_OTHER	53	25.4125	25.1784	3.30377
19	KOR_LL	1	*28.6079*	30.6262	5.1

Table 3. inputted and estimated CPUE SE. Fitted values are taken from model 1 (with all indices included) and model 4 (subset of indices and years). N /A indicates not used.

			model 1	model 4	model 1	model 4	
Fleet		N	Q	Q	r.m.s.e.	r.m.s.e	Inputted SE
20	Svey1_JPN_DWLLearly	23	0.000143		0.486404		0.204414
21	Svey2_JPN_DWLL1	12	6.94E-06	7.09E-06	0.307005	0.305093	0.201767
22	Svey3_JPN_DWLL1	13	1.03E-05	9.66E-06	0.426362	0.399991	0.201506
23	Svey4_JPN_DWLL1	10	1.90E-05	1.59E-05	0.192534	0.156083	0.203003
24	Svey5_JPN_DWLL2	10	4.16E-05	3.57E-05	0.634289	0.613963	0.258454
25	Svey6_JPN_DWLL2	12	2.97E-05	2.70E-05	0.416769	0.390271	0.195439
26	Svey7_JPN_DWLL2	13	7.03E-05	6.06E-05	0.563676	0.537163	0.200912
27	Svey8_JPN_DWLL3	12	0.000856	0.0007625	0.268026	0.261032	0.204767
28	Svey9_JPN_DWLL3	13	0.000947	0.0007104	0.228455	0.225579	0.203406
29	Svey10_JPN_DWLL3	10	0.000344	0.0002693	0.47981	0.412165	0.198173
30	Svey11_JPN_CLL	16	0.000199	*N/A*	0.301389	*N/A*	0.199211
31	Svey12_JPN_DFT	17	0.000849	*N/A*	0.409024	*N/A*	0.200941
32	Svey13_JPN_DFT	9	0.009273	*N/A*	0.215227	*N/A*	0.195863
33	Svey14_TWN_EARLY	24	0.000215	*N/A*	0.454614	*N/A*	0.4
34	Svey15_TWN_LATE	15	0.000487	*N/A*	0.226786	*N/A*	0.2
35	Svey16_HWLL	14	0.003471	0.0025896	0.525457	0.465904	0.286402

Table 4. Results of the test of seasonality of recruitment. Column headings are: season of assumed recruitment, total negative log-likelihood, and the change in negative log- likelihood from season 3 for each length composition component (fishery) (negative log likelihood minus season 3 negative log likelihood). A negative value indicates better fit (highlighted in green), and a positive value indicates worse fit (highlighted in red).

Season	Total	Fishery 1	Fishery 2	Fishery 3	Fishery 4	Fishery 5	Fishery 6	Fishery 12	Fishery 13	Fishery 14	Fishery 17	Fishery 18
season 1	5775.29	22.361	-10.828	58.98	46.63	63.354	4.3981	17.304	3.14	0.2457	172.533	8.0768
season 2	5583.85	21.087	-16.015	31.28	19.42	31.155	2.4114	4.998	0.6	0.4442	94.45	-0.1291
season 3	5408.16	0	0	0	0	0	0	0	0	0	0	0
season 4	5408.19	-3.551	-21.759	-11.73	-1.56	-23.119	0.8188	-4.431	2.487	-0.6615	48.327	-0.428

Table 5. Results of the test of for asymptotic fishery. Rank average is the ordered rank (1=best) of fit from the cross pair analysis. Likelihood average is the average of the negative log likelihoods from the cross pair analysis (lower is better fit). Values bolded indicate support for asymptotic fishery. Results are shown for two of the series of model assumptions (CV of length-at-age relation).

	CV0.14		CV 0.08	
Fishery	rank average	likelihood average	rank average	likelihood average
1	5.7	175	6.2	153.9
3	7.1	459.5	7.2	849.3
4	7.4	big	6.8	412.4
5	5.1	143.8	**3.5**	701.6
6	**2.1**	**28.5**	**2.6**	**45.1**
12	**2.9**	**75.9**	**1.6**	**34.9**
13	4.3	306.3	4.9	**84.4**
14	**3.3**	508.6	**3.2**	**67.5**
17	4.3	**121**	5.7	121.4
18	4.3	**124.1**	**2.3**	**30.7**

Table 6. Negative log-likelihood gradients (component likelihood-minimum likelihood within component) for size composition information for model 1 and the preferred model 4. The best fitting R0 will have a likelihood of 0 and changes in negative log-likelihood across different values of R0 can be thought of as how much information there is on scaling from that likelihood component.

model 1	DW early	JPN DWLL1	JPN DWLL2	JPN DWLL3	JPN CLL	JPN DRIFT	JPN OTHER_early	JPN OTHER_late	TWN LL	HW LL	WCPO
6.3	34	20	1	0	2	1	18	0	4	0	2
6.4	15	0	6	4	4	2	2	0	5	42	3
6.5	2	3	1	5	0	0	2	1	3	3	1
6.6	5	6	0	7	0	2	0	2	21	4	0
6.7	1	2	11	12	18	25	44	3	0	10	9
6.8	0	1	8	9	18	31	67	3	0	8	12

Size comp

Model 4	JPN DWLL1	JPN DWLL2	JPN DWLL3	JPN CLL	JPN DRIFT	JPN OTHER_early	JPN OTHER_late	TWN LL	HW LL	WCPO
6.3	0	0	0	0	0	3	0	2	0	0
6.4	0	0	1	0	0	2	0	1	1	0
6.5	0	1	1	0	1	1	1	1	2	0
6.6	0	2	2	0	1	0	1	0	3	0
6.7	2	3	1	10	21	11	3	0	7	1
6.8	3	3	0	10	28	17	3	0	5	1

Table 7. Negative log-likelihood gradients (component negative log-likelihood minus minimum likelihood within component) for CPUE series information for model 1 and the preferred model 4. The best fitting R0 will have a likelihood of 0 and changes in negative log-likelihood across different values of R0 can be thought of as how much information there is on scaling from that likelihood component.

Model1	DW early	DWLL Area1			DWLL Area2			DWLL area 3			JPN _CLL	JPN_ DFT	JPN_ DFT	TWN_ EARLY	TWN_ LATE	HWLL
6.3	111	4	11	0	6	7	29	0	1	6	0	3	3	8	0	2
6.4	55	2	9	1	4	8	22	1	0	3	3	0	3	1	0	0
6.5	23	2	11	2	6	10	26	1	0	8	1	0	2	0	1	2
6.6	13	3	12	4	9	13	28	1	0	13	2	0	1	0	3	5
6.7	4	0	0	0	1	1	4	0	4	1	14	3	0	3	1	3
6.8	0	0	0	0	0	0	0	0	7	0	16	4	0	4	3	3

CPUE

Model 4			DWLLArea1		DWLLArea2			DWLLArea3								HWLL
6.3		1	10	0	4	7	20	0	0	3						0
6.4		1	10	0	4	7	20	0	0	3						0
6.5		1	10	1	4	8	20	0	0	4						1
6.6		1	10	1	5	9	20	0	0	5						1
6.7		0	1	1	0	2	5	1	4	0						6
6.8		0	0	0	0	0	0	2	7	0						9

Table 8. Negative log-likelihood values from Model 4. Top: negative log-likelihoods by component. Bottom: negative log-likelihoods by component along with lambda values (likelihood multiplier). Note that individual component values are sometime rounded off for presentation.

Likelihood Component	Negative-log Likelihood
TOTAL	4441.73
Catch	1.6666
Equil_catch	0
Survey	8.74249
Length_comp	4431.14
Recruitment	0.177562
Forecast_Recruitment	0
Parm_priors	0
Parm_softbounds	0.00933656
Parm_devs	0
Crash_Pen	0

Fleet:	ALL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Catch_lambda:	_	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Catch_like:	1.6666	0.0255	0.046	0.05	0.06	0.99	0.06	0	0.06	0.06	0.06	0	0.065	0	0.055	0.055	0.042	0.042	0
Surv_lambda:	_	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Surv_like:	8.7425	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Length_lambda:	_	1	1	1	1	1	0	0	0	0	0	1	1	1	0	0	1	1	0
Length_like:	4431.1	305.293	1170	1120	583.7	63.4	0	0	0	0	0	607.6	136	27.16	0	0	356.5	61.47	1.252

Fleet:	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
Catch_like:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Surv_lambda:	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1
Surv_like:	- 5.0644	5.2524	-13	15.54	5.085	25.42	- 9.25	- 12.7	3.71	0.618	12.71	-2.12	-3.97	-9.06	-6.28
Length_lambda:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Length_like:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 9. Estimated Spawning biomass, recruitment and exploitation level (1-SPR) from model 4 along with their estimated asymptotic standard deviation.

Year	SSB(mt)	StdDev	Recruitment (000's)	StdDev	1-SPR	StdDev
Virgin	19811.5	779.137	592.622	23.3063		
Initial	1348.73	614.773	592.622	23.3063		
1975	3189.22	381.824	432.297	56.5365	0.869459	0.009352
1976	3191.69	275.9	501.833	50.3	0.836524	0.011224
1977	3033.05	244.366	286.261	41.2967	0.868677	0.007653
1978	2379.31	165.492	1297.45	68.7382	0.925711	0.004202
1979	1634.87	96.3727	350.886	55.5964	0.869535	0.007277
1980	2474.56	147.796	549.941	62.0262	0.859067	0.009123
1981	2972.81	218.797	554.59	54.8342	0.852473	0.010662
1982	2807.83	244.177	225.822	39.5199	0.814413	0.012694
1983	2907.84	245.622	454.892	54.6577	0.823418	0.012243
1984	2462.37	234.137	1699.74	89.2203	0.901192	0.008888
1985	2394.68	222.392	253.553	56.6132	0.838054	0.011498
1986	3538.71	249.81	351.567	45.0333	0.842594	0.010048
1987	4004.92	314.339	838.41	64.2183	0.81884	0.013704
1988	3420.8	339.629	573.174	64.3452	0.887964	0.010021
1989	3018.11	330.425	315.814	52.3444	0.846277	0.013867
1990	2934.32	328.576	939.343	69.477	0.844831	0.014616
1991	2935.61	330.468	293.179	57.5956	0.830165	0.01499
1992	3165.84	327.846	778.016	52.9497	0.811342	0.015569
1993	3331.46	331.806	122.888	32.3994	0.843359	0.012284
1994	3177.9	316.287	570.019	39.4208	0.854517	0.011899
1995	2652.96	294.538	341.078	34.3753	0.895164	0.009795
1996	2018.01	257.723	345.098	33.3305	0.875975	0.012419
1997	1796.28	236.738	680.534	51.5301	0.869713	0.014341
1998	1803.65	235.624	348.326	43.6157	0.888911	0.012266
1999	1872.56	245.133	310.558	38.0867	0.863409	0.016733
2000	1882.84	280.941	562.703	50.9575	0.850408	0.021175
2001	1943.76	324.316	310.426	39.9698	0.827797	0.025131
2002	2196.37	369.633	631.192	53.5515	0.774165	0.030271
2003	2537.12	422.677	306.922	41.6452	0.77443	0.029671
2004	3018.06	467.497	128.711	25.9375	0.730818	0.028737
2005	3232.79	482.644	434.557	33.5022	0.748071	0.026793
2006	3047.53	469.973	115.762	25.0443	0.780834	0.023887
2007	2662.39	425.596	187.273	25.1795	0.779971	0.023549
2008	2230.38	373.98	121.311	27.2894	0.868964	0.018345
2009	1459.97	319.469	403.274	33.4563	0.868791	0.026941
2010	1143.46	304.042	368.129	41.9181	0.840335	0.034457

Figures

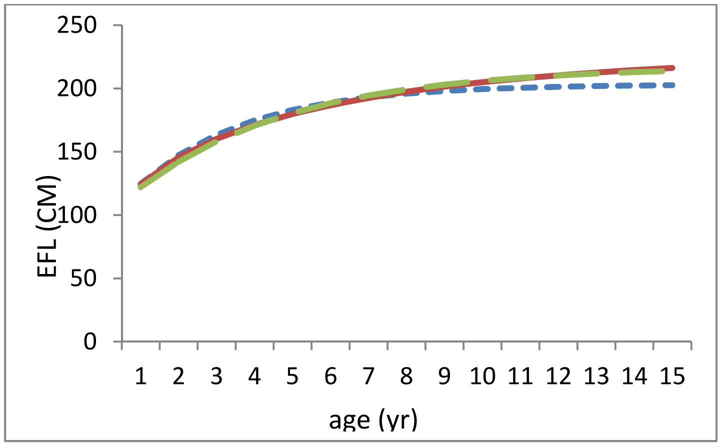


Figure 1. Plot of the Generalized Von Bertalanffy length at age (red solid line) and assessment model representation of that relationship (olive long dash). Also plotted is an alternative growth model (blue short dash) from the source paper (Sun et al. 2011a). The alternative growth form illustrates that there is uncertainty in the size of the oldest fish.

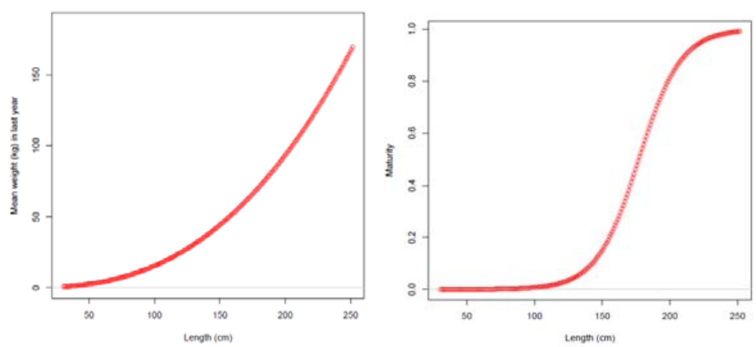


Figure 2. Weight-at-length and maturity-at-length used in the stock assessment model.

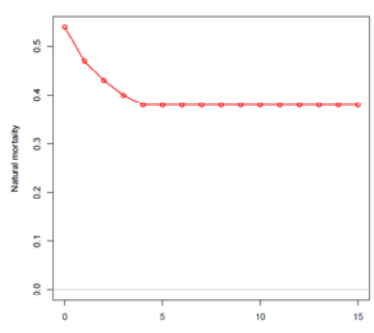


Figure 3. Natural mortality-at-age assumed in the population dynamics model.

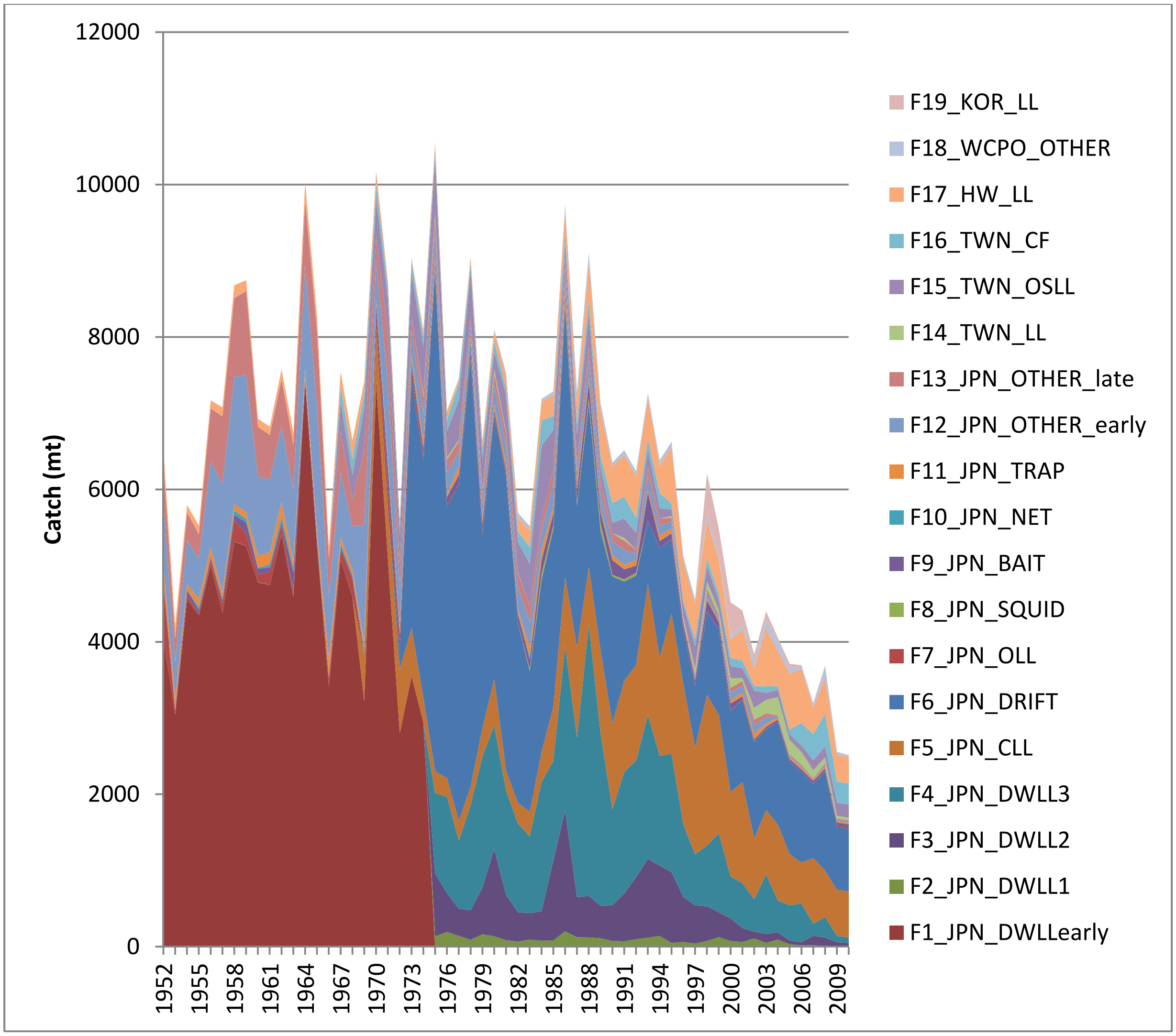


Figure 4. Plot of catch of striped marlin by year and fishery in mt. Fisheries with catch reported in numbers were converted into mt inside the stock assessment model.

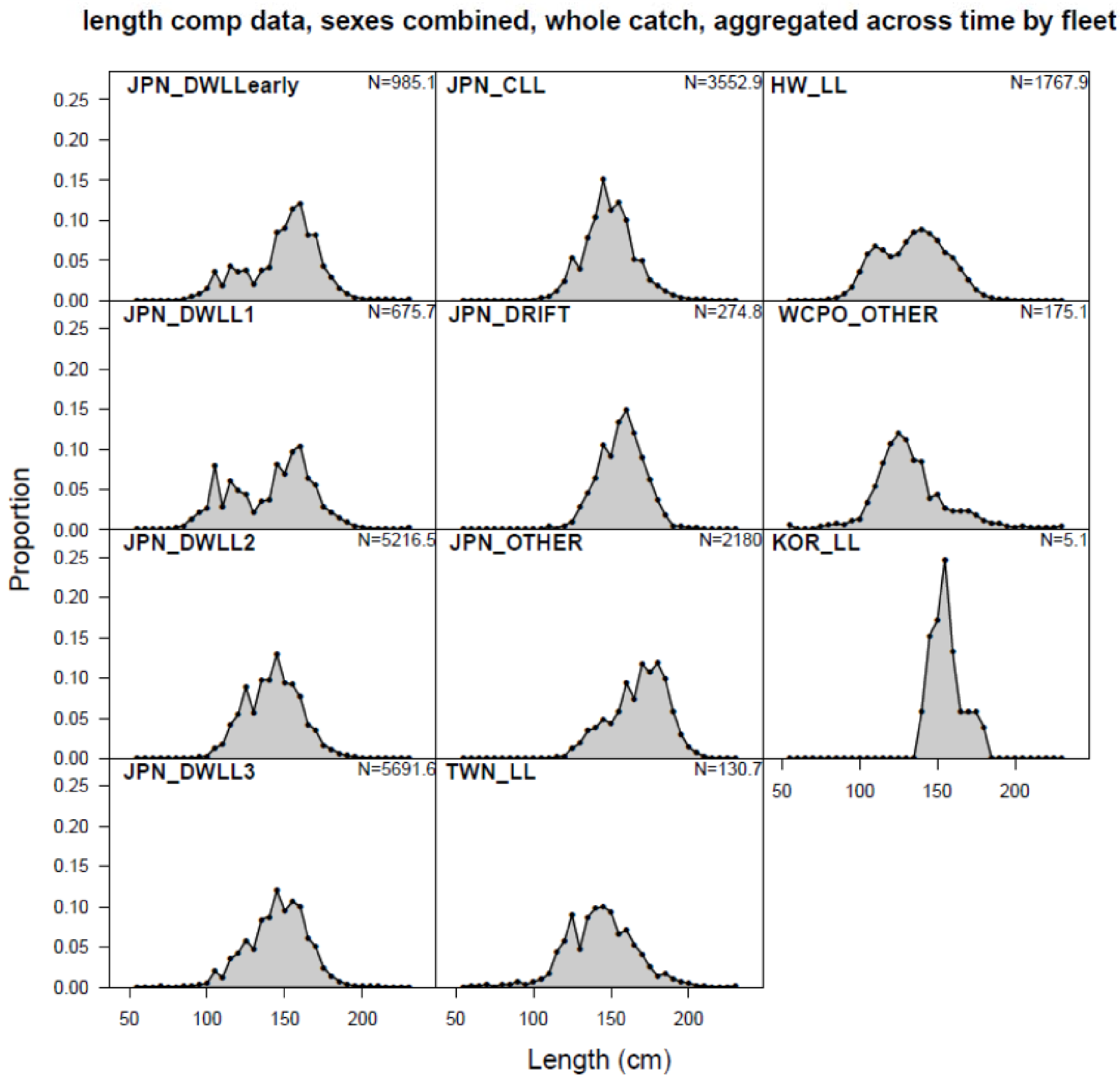


Figure 5. Observed proportion at length from all fisheries reporting size information. Samples were aggregated across year and season within fishery. Units are eye-to-fork length.

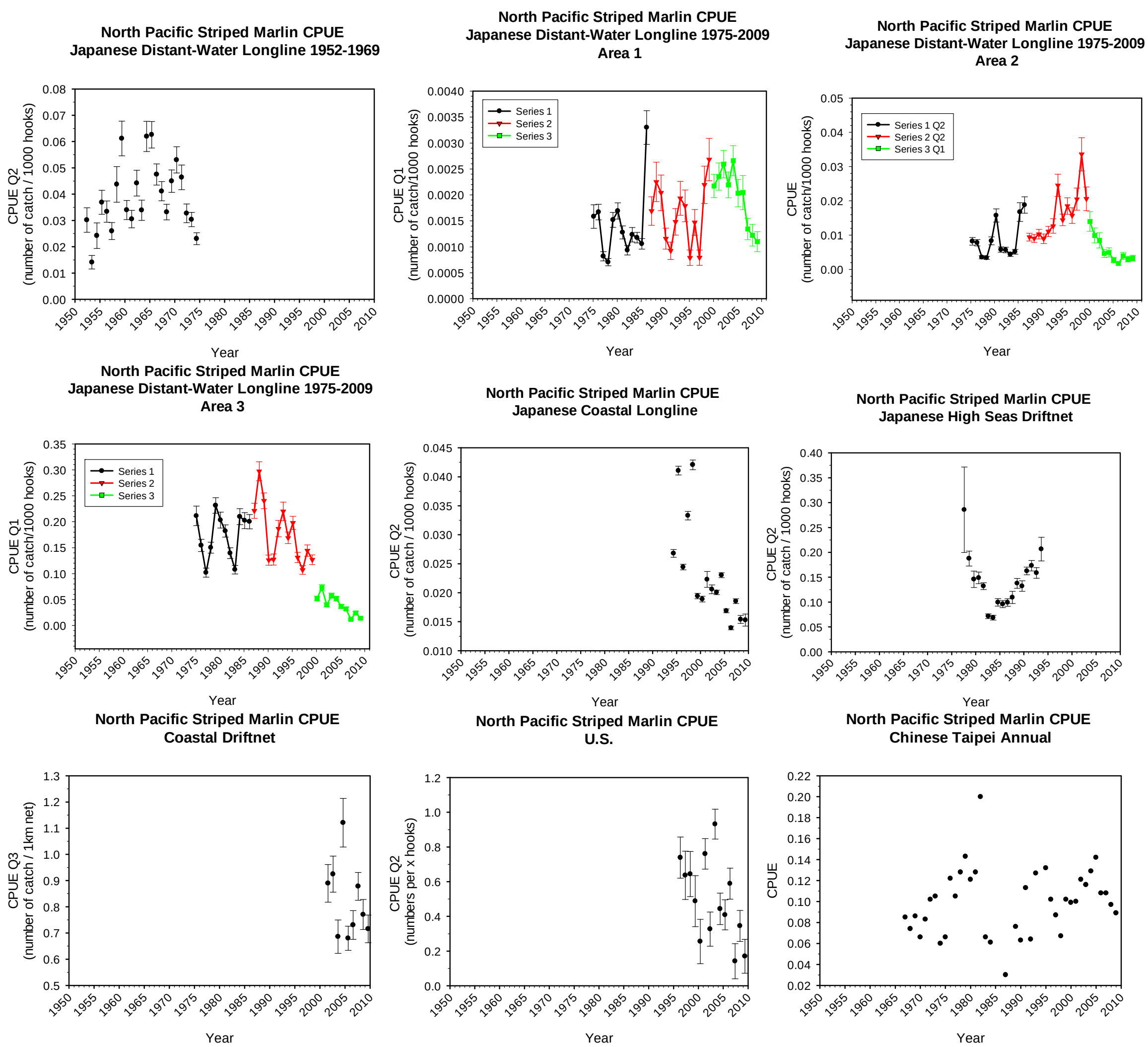


Figure 6. Plots of CPUE by fishery. For the Japanese Distant water longline fisheries there are 3 separate CPUE series (by time period) reflecting a change in fishery practices. Error bars represent 1 SD and do not include the addition of process error.

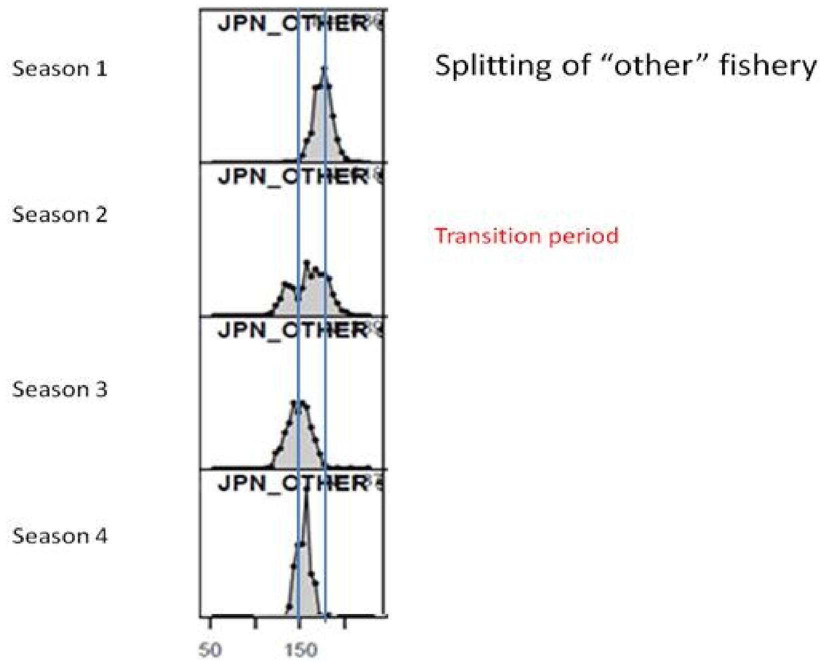


Figure 7. Seasonally summed size composition from the Japanese Other fishery showing the two size-distinct groups of fish (early and late).

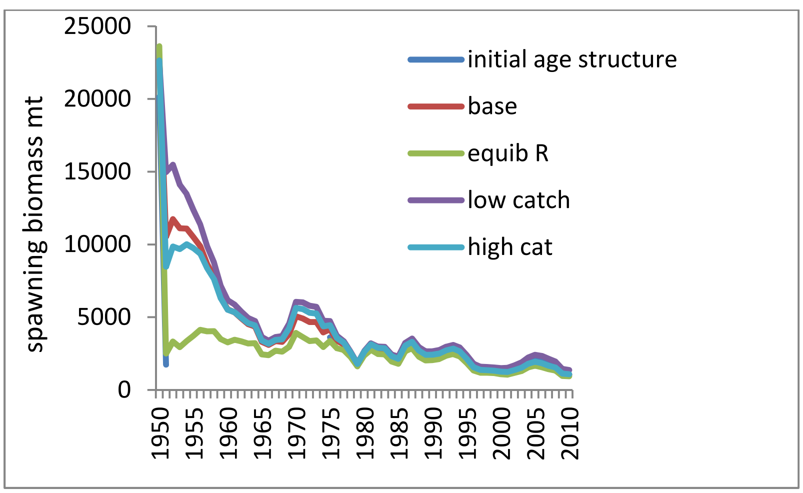


Figure 8. Spawning biomass trends by starting conditions. Initial age structure indicates a model started in 1975 without fitting to initial catch and with initial age structure estimated. Base indicates mode with 100,000 fish and 280 tons of equilibrium catch. Low and high catch are 1/3 more and less than base. Equib R indicates results from a model that decouples the equilibrium recruitment from the S/R curve.

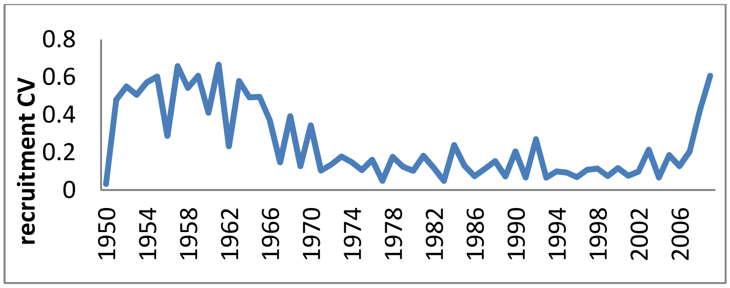


Figure 9. Estimated recruitment deviation uncertainty from 1952-2010.

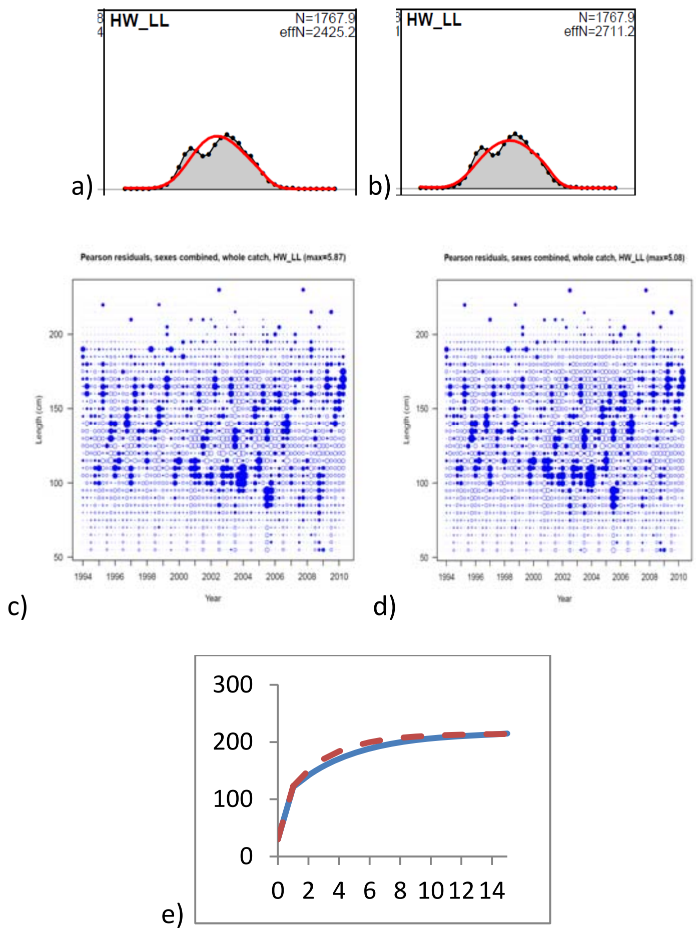


Figure 10. Observed and expected proportion at length from the Hawaii Longline fishery. Results are from a model that included all information (Full) and a model that included only the Hawaii longline proportion at size information and CPUE in fitting (Reduced). In the Reduced model growth was also estimated. Panel a) shows the combined (across years and seasons) composition fit to the size data from the Full model and Panel b) shows corresponding results from the Reduced model. Panel c) shows the Pearson residuals for the Hawaii longline composition from the Full model and Panel d) gives residuals from the Reduced model. Panel e) depicts the assumed growth in the Full model (solid) and the estimated growth from the reduced model (dashed).

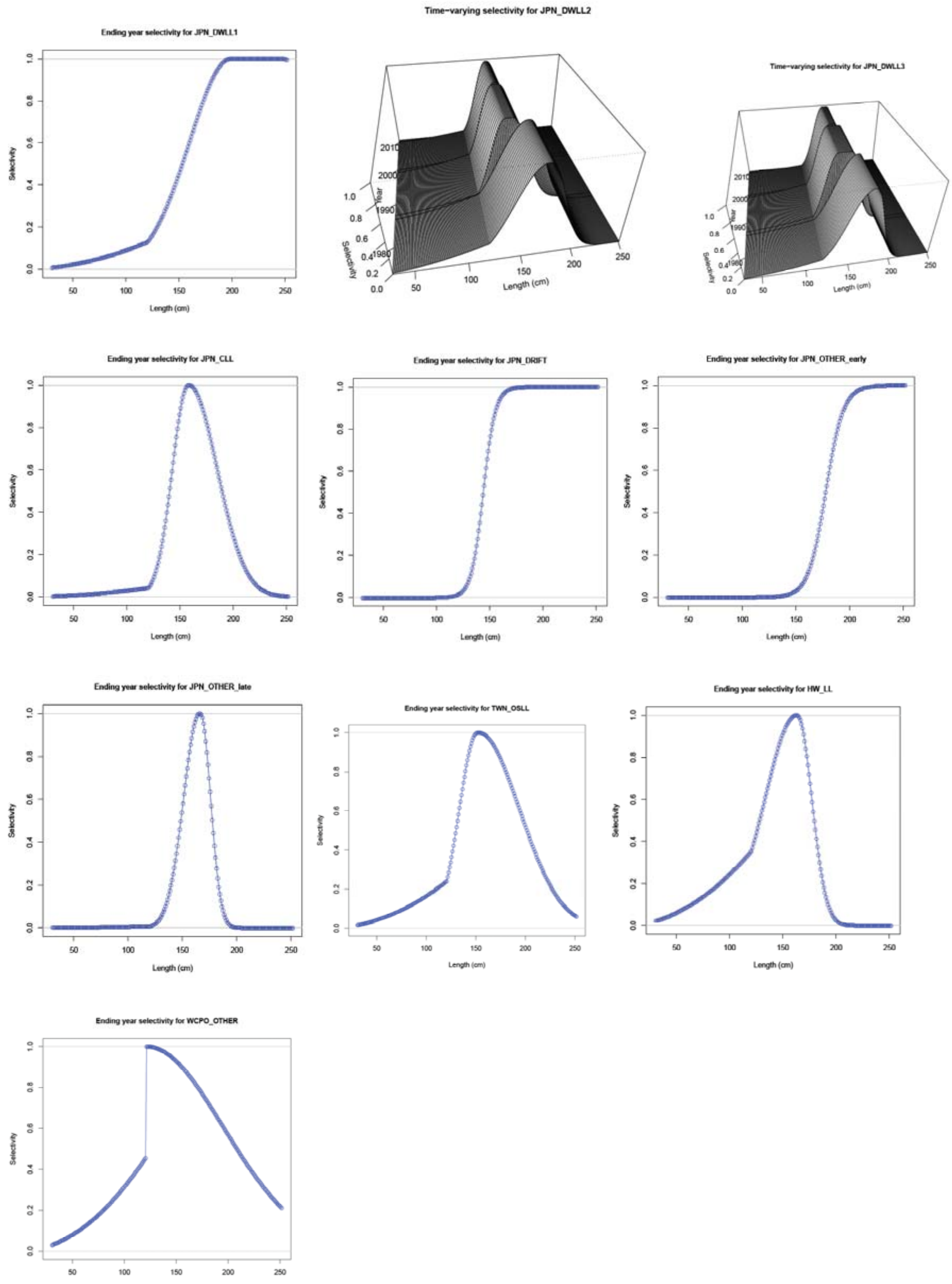


Figure 11. Estimated selectivity patterns by fishery (model 4). Fisheries with time varying selectivity patterns are displayed in 3-D plots.

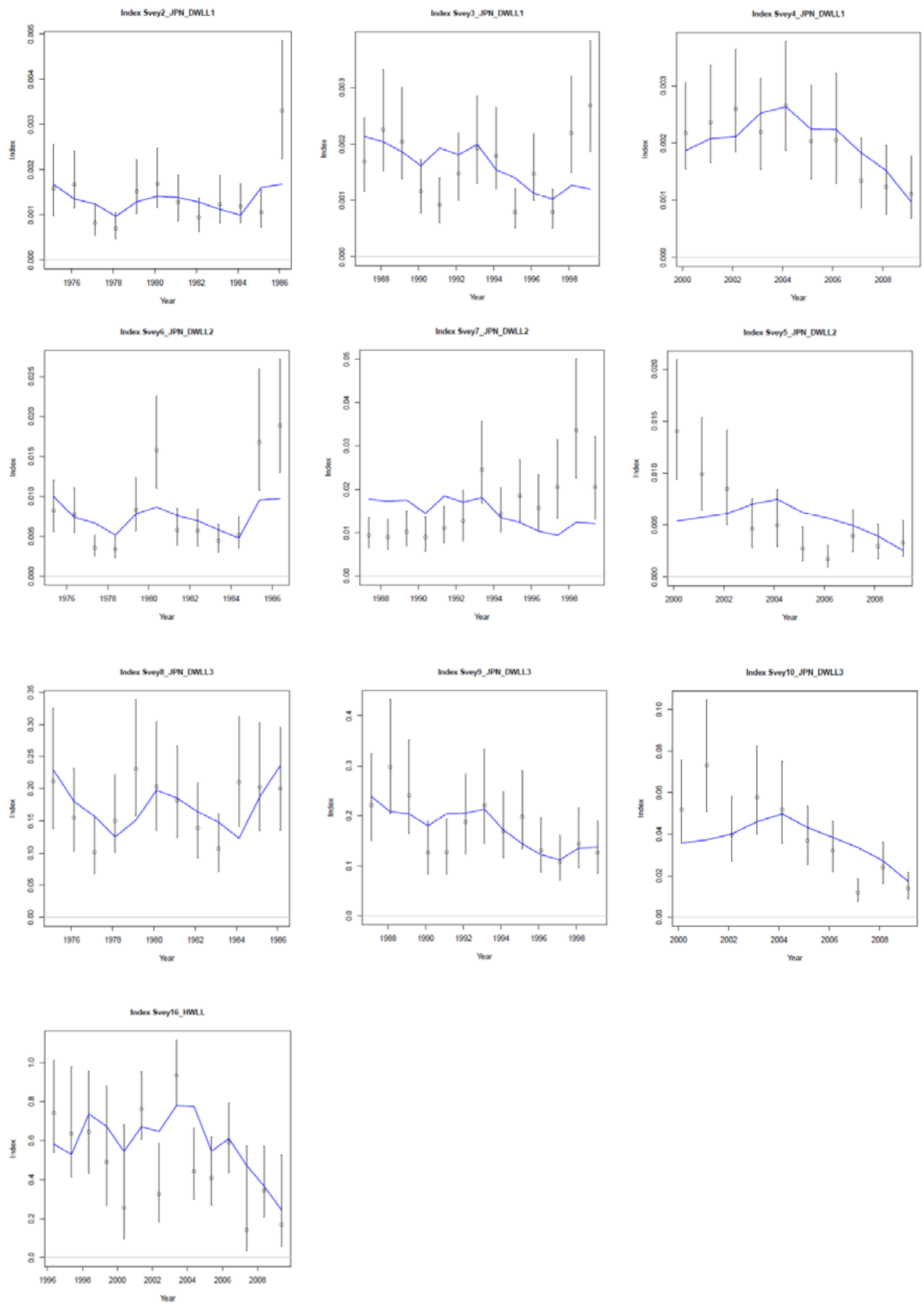


Figure 12. Observed (dots) and predicted CPUE (lines) for the Japanese DWLL fleets areas 1-3 and the Hawaiian LL fleet (model 4). The vertical lines represent the inputted CV around the observations.

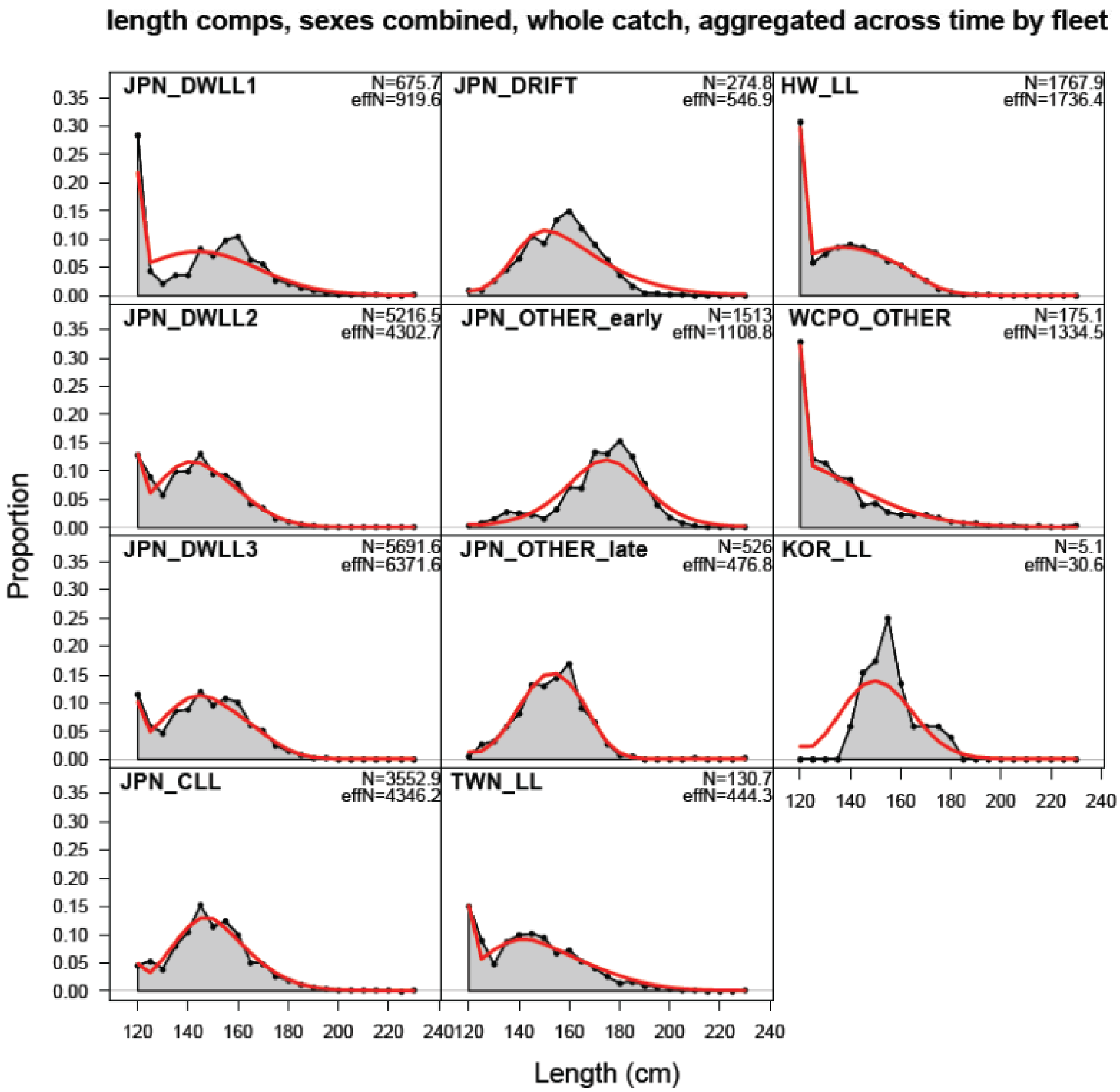
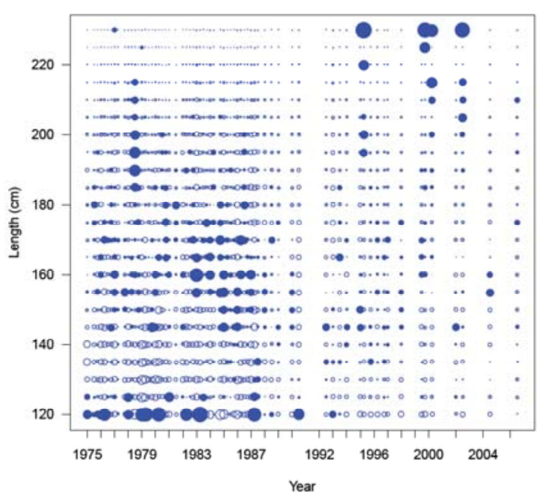
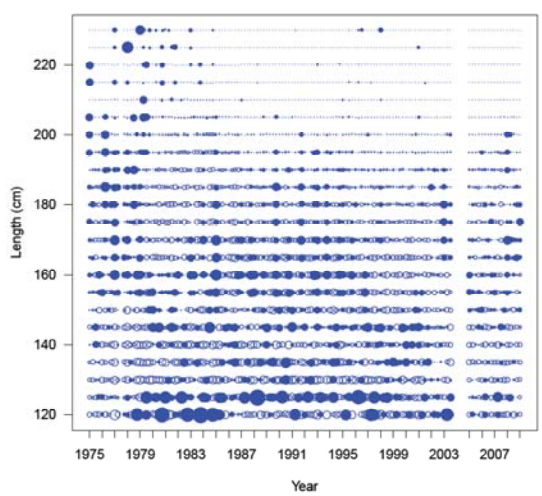


Figure 13. Plot of combined (across year and season) observed and predicted size composition by fleet (model 4).

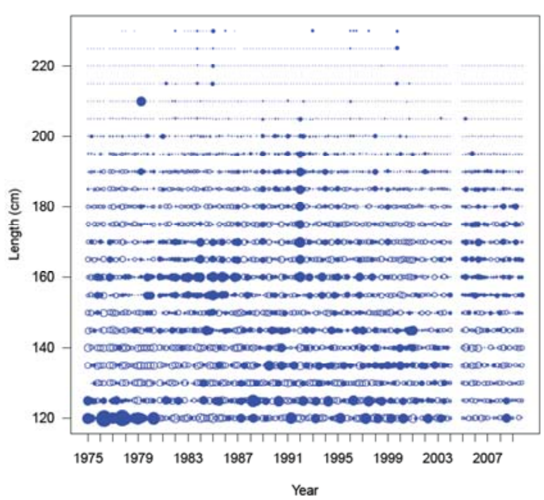
Pearson residuals, sexes combined, whole catch, JPN_DWLL1 (max=7.12)



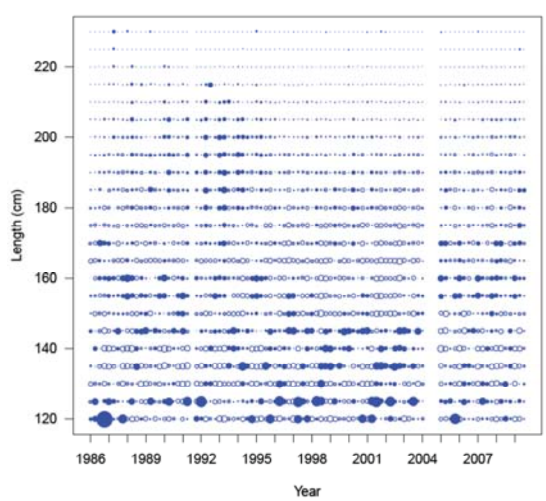
Pearson residuals, sexes combined, whole catch, JPN_DWLL2 (max=8.83)



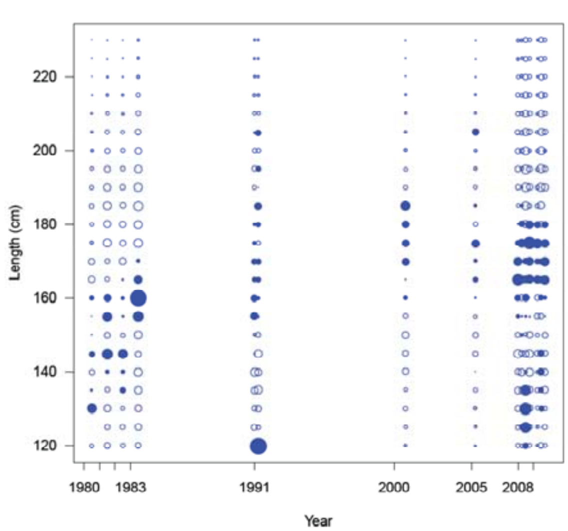
Pearson residuals, sexes combined, whole catch, JPN_DWLL3 (max=10.17)



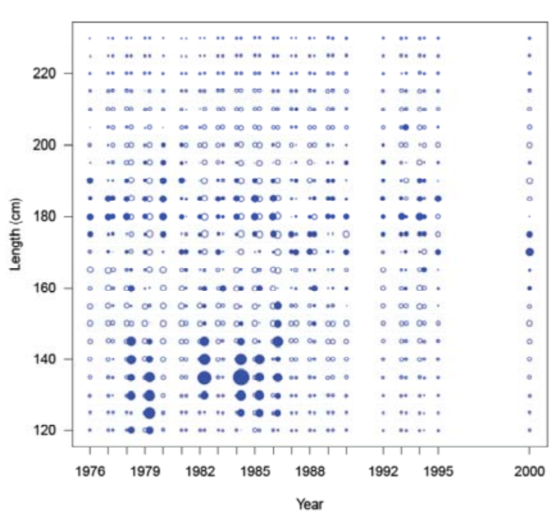
Pearson residuals, sexes combined, whole catch, JPN_CLL (max=12.8)



Pearson residuals, sexes combined, whole catch, JPN_DRIFT (max=4.13)



Pearson residuals, sexes combined, whole catch, JPN_OTHER_early (max=15.58)



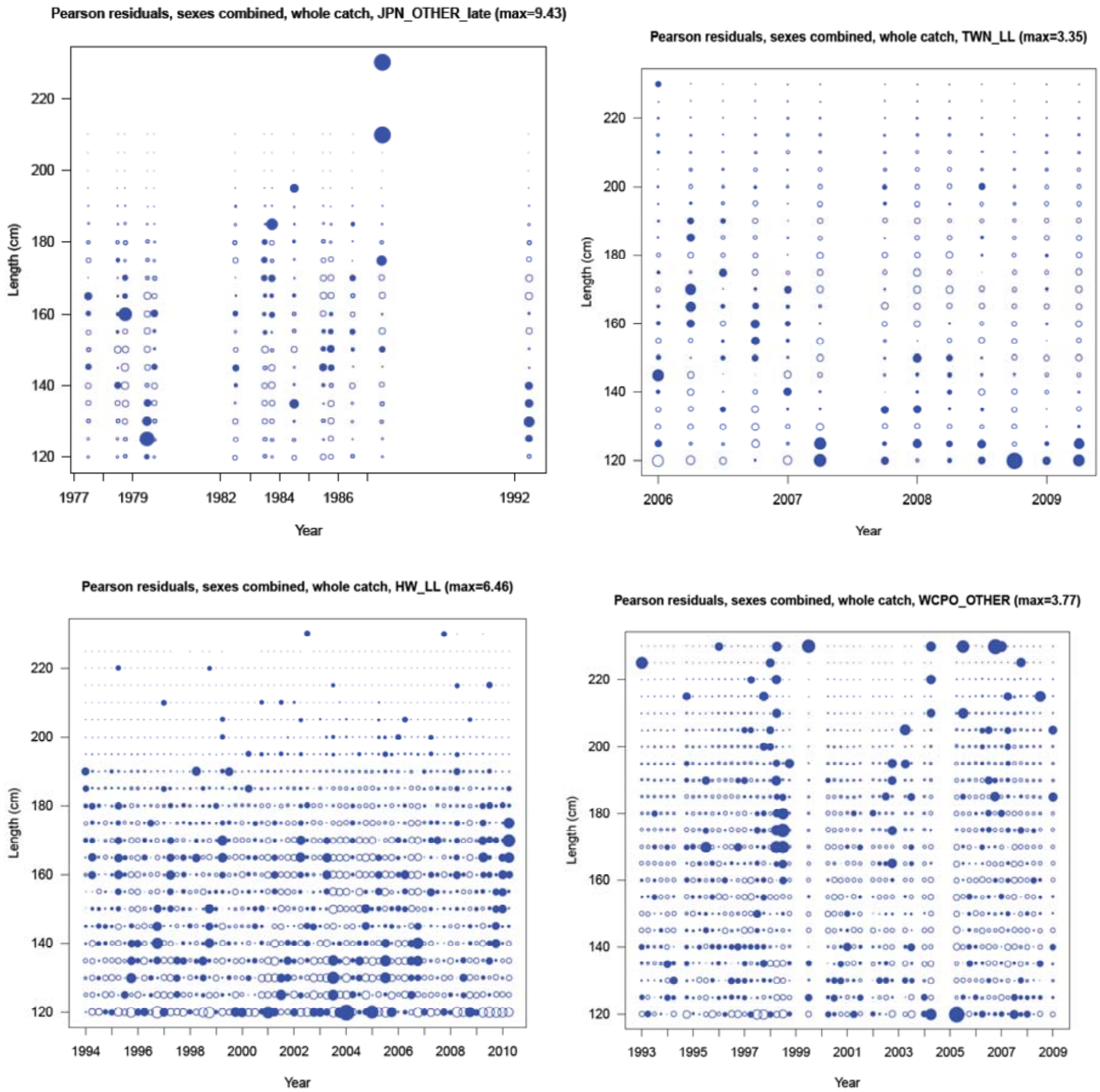


Figure 14. Pearson residuals from the fit to size composition information by fishery, season and year. Solid circles represent observations that are greater than the model predictions and open circles represent observations are less than the model predictions (model 4).

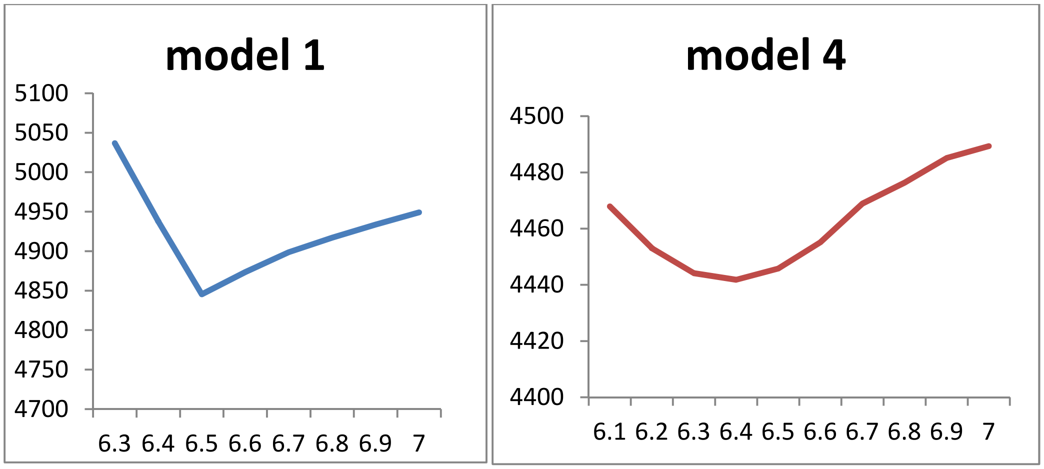


Figure 15. Total likelihood (y-axis) against fixed values of lnR0 (x-axis) from Model 1 and Model 4 .

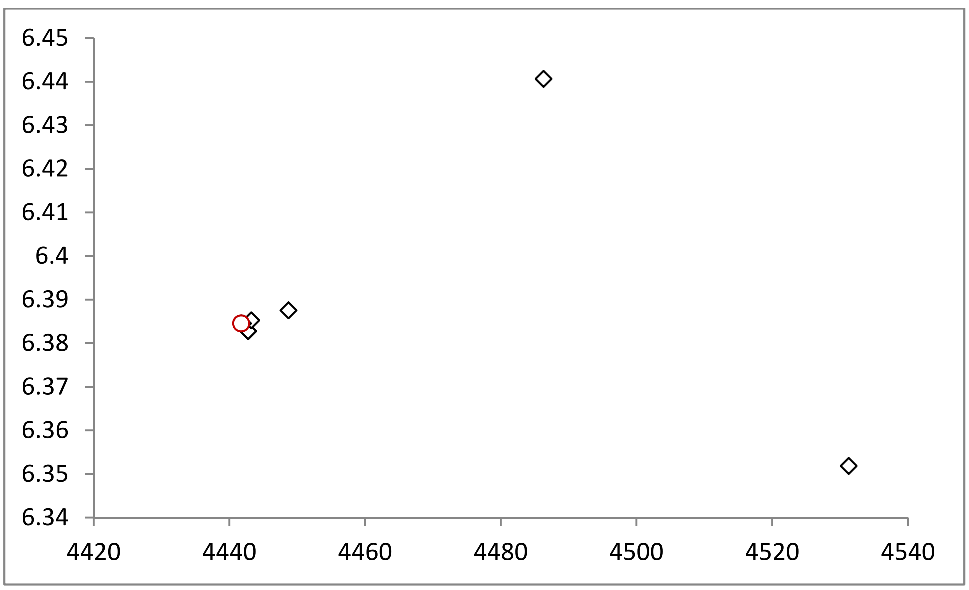


Figure 16. Plot of estimated R0 (y-axis) and total ending likelihood (x-axis) for random starting values of the model (diamonds) and the base model (circle) (model 4).

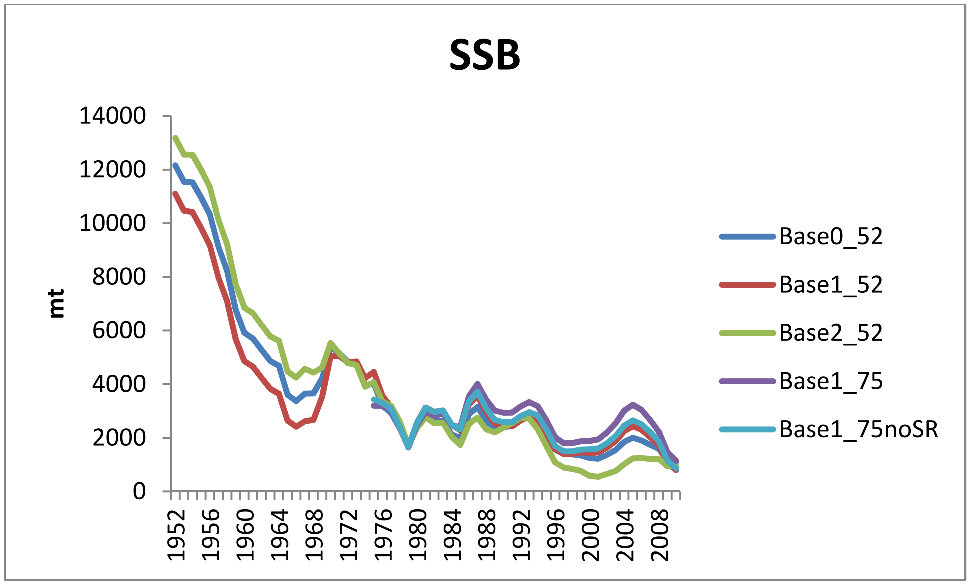


Figure 17. Plot of Spawning biomass (1952 to present) from 5 alternative models (legend shows models 1-5 in order).

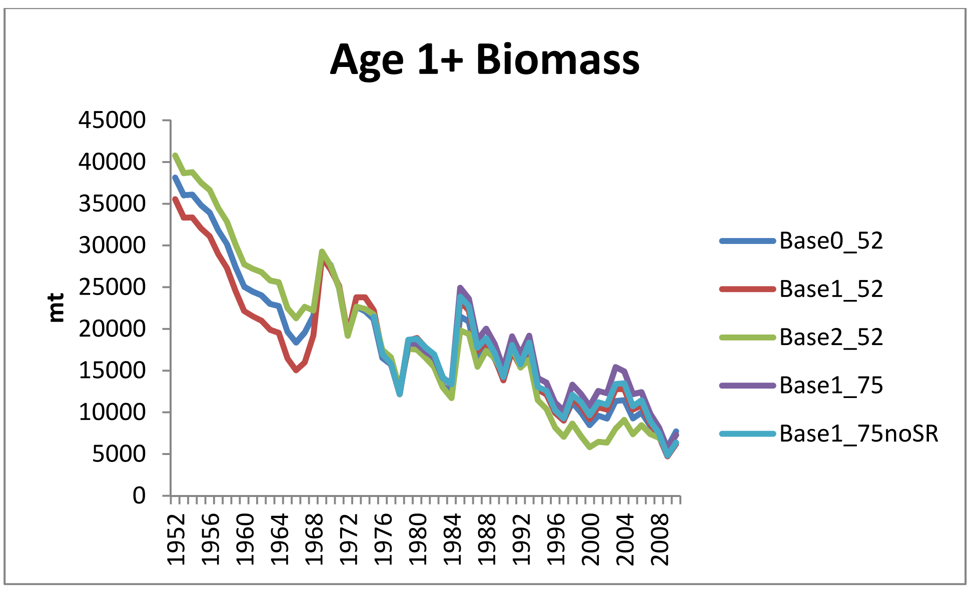


Figure 18. Plot of Age 1^{+} biomass (1952 to present) from 5 alternative models (legend shows models 1-5 in order).

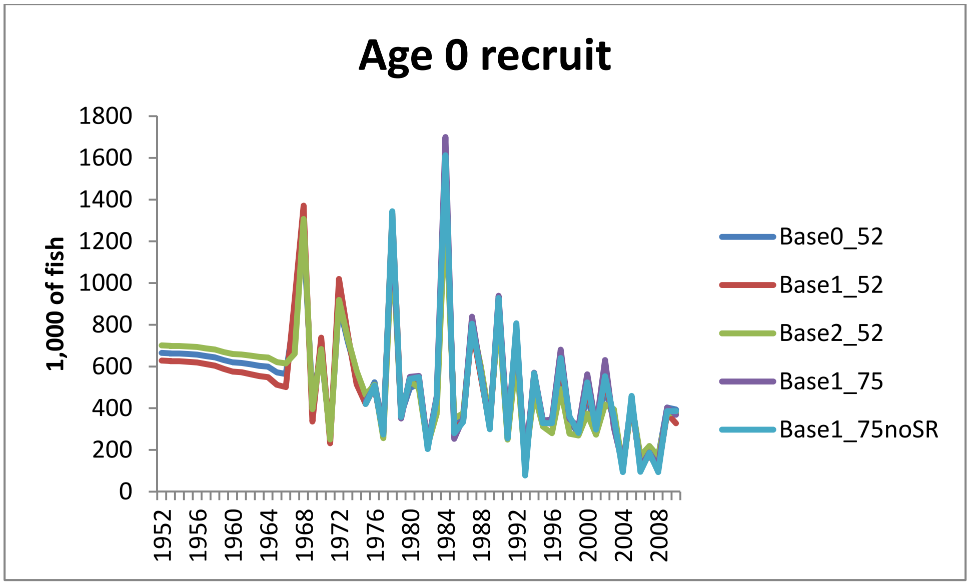


Figure 19. Plot of Recruitment (1952 to present) from 5 alternative models (legend shows models 1-5 in order).

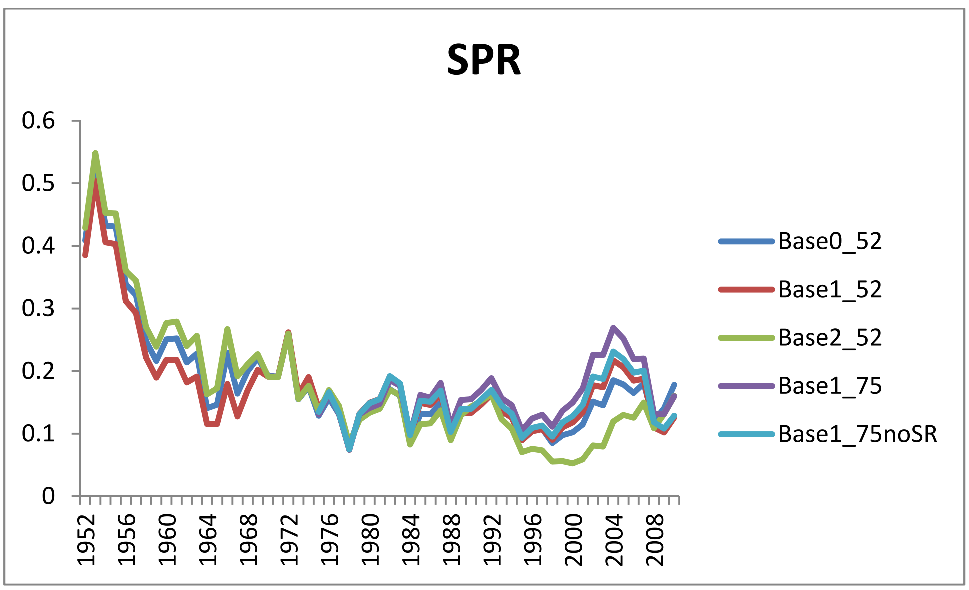


Figure 20. Plot of fishing intensity as SPR from 5 alternative models (1952-present).