Investigating the influence of length-frequency data on the stock assessment of Indian Ocean bigeye tuna

Rishi Sharma
Indian Ocean Tuna Commission (IOTC),
PO Box 1011,
Le Chantier Mall,
Victoria , Seychelles

Adam Langley,
IOTC Consultant,
7 Van Diemen Street,
Nelson, New Zealand

Miguel Herrera
Indian Ocean Tuna Commission (IOTC),
PO Box 1011,
Le Chantier Mall,
Victoria , Seychelles

James Geehan,
Indian Ocean Tuna Commission (IOTC),
PO Box 1011,
Le Chantier Mall,
Victoria , Seychelles

Abstract

Most integrated stock assessment models are fit to alternative sources of data like indices of abundance and length/age composition of catches in specific fisheries. While indices of abundance are often standardized over time, not much attention is paid to the temporal stability of the length/age data. A sequential approach to fitting model outputs to all sources of data, varying the weight given to the length composition data, for Indian Ocean bigeye tuna (*Thunnus obesus*) is examined in this paper. The sensitivity of the bigeye tuna stock assessment to assumptions regarding the size-selectivity of key fisheries and the relative weight of size frequency data in the stock assessment is examined. Logistic, double normal, and cubic spline selectivity functions are used to model the size composition of catches in the main industrial fisheries (longline and purse seine). Overall, there is a poor fit of stock assessment models to the individual length frequency observations collected from these fisheries, although marginal improvements of fit were made when temporally variable selectivity was implemented in the SS-III framework using the above described functions. The most influential factor in the assessment was the weighting of the length composition data relative to the indices of stock abundance. Contradictory signals between these two data sources have a large effect on spawning biomass dynamics, and inference based on these weightings can produce different management conclusions. We examined alternative hypotheses and discuss the merits of fitting to all sources of data, or discounting some information if it has been unreliably collected over time. We emphasize that understanding the data is key to performing a well-calibrated stock assessment, and further refinements to the approach pursued here are discussed.

Introduction

Integrated stock assessment models have been used in fisheries management for the past three decades (Fournier and Archibald 1982). However, blindly fitting models to all available sources of data may lead to inaccurate results, as has been discussed extensively by Francis (2011). Often fisheries change over time due to shifts in fishery selectivity or the catchability of different age or size classes of fish. When this type of trend is apparent in the data, the modeller is faced with two choices: 1) model these as separate fisheries over time using different catchability estimates, or 2) model these as one fishery with changing selectivity over time. Often the latter approach is used (Gavaris and Ianelli 2002), as it easier to implement.

While most modellers and fisheries management practitioners understand the relative importance of size selectivity and its interaction with biomass and effect on optimal yield and stock status (Haddon 2011, Hilborn and Walters 1992), it is an extremely difficult process to estimate (Hilborn and Walters 1992). Most assessments use external sources to justify the general shape of the selectivity curve (i.e., whether it is asymptotic or dome-shaped) and estimate the specific shape (i.e., the parameters of the functional form) by fitting the model to age and/or length-composition data. In most tuna assessments age data are limited, so length frequency data are the primary source of information. Tagging data can also provide some information on selectivity, although in the case of the Indian Ocean region these data are of limited use, primarily because tag mixing assumptions are violated (Langley et al. 2012). In the case of a highly migratory species like bigeye tuna (*Thunnus obesus*; hereafter referred to as BET), modellers assume similar functional forms for selectivity used by tuna Regional Fisheries Management Organizations (RFMOs) across the globe. Note that estimated selectivity is not independent of other parameters, particularly natural mortality (M), which is kept constant at different ages through time (M=0.8 annually for the early ages in this assessment, and declines to M=0.4 by age 3). These levels of natural mortality are comparable to those used by the Inter-American Tropical Tuna Commission (IATTC) and Western Central Pacific Fisheries Commission (WCPFC) BET stock assessments (Kolody et al. 2010).

Here we present a sequential analysis to account for temporal size-selective patterns in BET catch data by modelling different functional forms of selectivity, along with temporal changes in selectivity over time. The models examined use the following choices: 1) different selectivity functions are examined for the main fisheries (longline and purse seine fisheries by area) to capture changes in length frequencies over time, 2) different forms of temporal variability in size-selectivity are examined, and 3) the influence of the effective sample size with and without time-varying selectivity is examined. Although this analysis focuses on Indian Ocean bigeye tuna, the approach outlined here would be useful in all stock assessments, and emphasizes that understanding the data and its uncertainty is key to a rigorous and sound stock assessment.

BET accounts for 430K Mt (10% of all tropical tuna) of worldwide tuna catches by volume (2008-2010, Herrera et al. *in press*) and is caught primarily in the Pacific Ocean (54%), followed by the Indian (28%) and Atlantic Oceans (18%).The data used in this analysis are primarily from the longline and purse seine fisheries in the Indian Ocean. The longline fishery (distant water) commenced operation in the Indian Ocean during the early 1950s. BET represents a significant component of the total catch from the longline fishery; catches reached a peak in the late 1990s-early 2000s at 70K to 90K Mt per annum. The purse seine fisheries (primarily log and FAD based fisheries) and fresh-chilled longline fisheries developed in the mid-1980s and total BET catches peaked at around 155K Mt in the late 1990s. Since the mid-2000s, the total annual BET catch has fallen considerably, primarily due to a decline in the longline catch in the western equatorial region in response to the threat of piracy off the Somali coast. In 2011 the total annual catch was estimated to be around 91K Mt (Herrera et al. 2012). Small scale fisheries that encounter BET are the Maldivian pole/line fishery and gillnet fisheries operating in the Indian Ocean. Both major fisheries have collected extensive length frequency data over time and area, which are used in the model fitting exercise described here.

Initial assessments of BET have generally applied non-equilibrium surplus production models (Nishida and Rademeyer 2011) and integrated stock assessment methods that are fit to length composition data, indices of abundance, and tagging data (Methot and Wetzel 2013, Kolody et al. 2010, Shono et al. 2009). Both assessment approaches indicate that the Indian Ocean BET stock has not been overfished, although Kolody et al. (2010) highlighted the high level of uncertainty (both derived parameter uncertainty and structural uncertainty, Punt and Hilborn 1997, Quinn 2003) associated with key model parameters that resulted in a range of contrasting estimates of stock status, some of which indicated that the stock is in an overfished state.

The approach pursued here builds on some of the structural uncertainty examined in Kolody et al. (2010) and Langley et al. (2012). The main focus of this work is examining how different functional forms of selectivity can affect stock assessment, and the impact of considering time-invariant vs. temporally-variable selectivity. In addition, we examine the effect of differentially weighting size composition data on model outcomes. If information from the fisheries suggests that the size composition of the catch has changed in response to a change in fishery operations (for example, a change in the spatial or seasonal operation of the fishery), it may be appropriate to estimate temporal variation in the selectivity parameters as we demonstrate here. We examine some alternative hypotheses regarding selectivity forms and weights used for fitting the length-composition data, and the effect of these on reference points (or derived model outputs). The merits of fitting models to all sources of data, or discounting some information if unreliably collected over time, are also discussed.

Materials and Methods

Construction and fitting of integrated stock assessment models combines all available data in a statistical maximum likelihood estimation framework (Fournier and Archibald1982; Maunder and Punt 2013; Methot and Wetzel 2013). These models simultaneously estimate numerous parameters to give the best fit to observed data (Deriso and Parma 1987, Fournier and Archibald 1982). For this analysis we used the Stock Synthesis software (Methot 1989, Methot and Wetzel 2013) that adapts the basic age-structured algorithms presented in Fournier and Archibald (1982). Stock Synthesis has considerable flexibility in the parameterization of length- and age-based selectivity, including a range of functional forms, formulation of priors for the main selectivity parameters, and temporal variability (in blocks, as parameter deviates, or linked to an exogenous variable).

Catch and Biological Data: Model Inputs

Temporal units

Data were disaggregated by calendar quarter (quarter 1 = Jan-Mar), and the model was iterated on quarterly time steps in order to capture potentially important seasonal dynamics over the period 1952-2011. The model was aggregated by sex; age class bins were yearly with a plus group at age 10 (the model worked on quarterly age increments resulting in 40 different quarterly ages in the model).

Spatial structure, fishery dynamics, catch and CPUE data

Spatial structure of population dynamics in a given fishery is an important consideration when constructing a stock assessment model. BET population dynamics are spatially aggregated for the purposes of these model runs; the main fisheries operating in three areas are shown in Table 1, Figures 1 and 2. The primary fisheries modeled are longline in each of the three areas (LL 1-3), and the Purse Seine Free School (PSFS) and Purse Seine Log School (PSLS) in Areas 1 and 2. Three other fisheries with a shorter catch history are included: the fresh tuna fishery (LL in Area 2) a hook and line fishery in Area 2, and a bait boat (BB) fishery in Area 1. The “other” category includes all other catches, again in Areas 1 and 2, primarily the non-industrial fisheries.

Sampling coverage of the length frequencies of catches varies considerably by type of fishery (multiple fleets make a fishery in one area; for example, the Japanese longline fleet is a subset of the longline fishery), and by area. In the case of longline fisheries, sampling from log books for some fleets can be as high as 80% (defined in terms of the number of fish sampled for length compared to the number of fish recorded for catch and effort by vessels operating in the area). The purse seine fleets report similar levels of coverage. However, in the case of the “other” fisheries the sampling coverage is much lower, generally less than 5%, while bait boat coverage is in the range of 30%. The sampling coverage of length-frequencies are high (greater than 30% of the total fish caught) for a number of fleets (e.g., Taiwanese longline and European purse seine fleet), and low for other fleets (less than 1%). Some investigators have noted inconsistencies over time in methods of collection of length-frequency data for the longline fleet (Geehan et al. 2013).

Fishery catches are provided in Herrera et al. (2012). The main fisheries used in this analysis include the LL fleets and the PS fleets which account for 90-95% of the catch (Figure 3). The length composition data (Figure 3) for each fishery tended to be a composite of length data from a number of different fleets, consisting of a mixture of commercial vessels, training vessels, and scientific observers. Using aggregated fleet data may not capture differential operations in the individual fleets (e.g., Japanese longline vessels compared to Taiwanese longline vessels). Japanese longline vessels are the primary source of the length frequency data for BET in the early 1970s and 1980s, while in the 1990s Taiwan replaced Japan as the main longline fleet reporting size data. According to the Overseas Fisheries Development Council (OFDC) of Taiwan, size measurements between 2000 and 2011 were recorded for over 4.4 million BET specimens from longline vessels. Between 2003 and 2005 alone – i.e., the years of highest sampling – length measurements were recorded for over 1.7 million BET samples by Taiwan. In addition to these differential fleet characteristics, the spatial distribution of the operation of these fleets has also changed considerably over time

The primary abundance data used in the BET assessment models are taken from standardized CPUE indices derived from the Japanese longline fleet, which has been in operation since the late 1950s. This is the only such standardized data set collected systematically and, like other tuna RFMOs around the world, we use it as a primary source of data in fitting these models. Standardized CPUE indices for the entire Indian Ocean were derived from the Japanese longline fleet using a generalized linear model (GLM, Satoh and Okamoto 2012, Hoyle and Okamoto 2011). The indices are derived by year and quarter for 1960–2011 (Figure 4). The overall Indian Ocean CPUE indices are very similar to the CPUE indices from the western equatorial region (Satoh and Okamoto 2012), and were all assumed to have a coefficient of variation (CV) of 10%. The high level of precision was assumed because the CPUE indices are the primary indicator of relative abundance in the assessment model and so the derived trends in stock abundance should be generally consistent with these indices. For all models examined in this analysis, catchability for the main longline fisheries was assumed to be temporally invariant.

Given all these complexities, it seemed appropriate to use differential selectivity patterns by time and fishery (stratified by area) to better capture some of the fishery-specific temporal and spatial dynamics, as there is some evidence based on the 5 by 5 degree data (latitude and longitude degrees on the Indian Ocean) that suggests that the magnitude of catch can vary by a factor of two across different 5 by 5 degree areas over times (IOTC 2012).

Model Structural Assumptions

Biological parameters

Recent estimates of Indian Ocean BET growth derived from otolith and tag release/recovery studies are available from Eveson et al. (2012). Growth estimates are available for both sexes combined. The quarterly growth deviates from a von Bertalanffy growth function, with considerably lower growth for quarterly age classes 4–8 (Figure 5). Maximum average length (L_∞) was estimated by Eveson et al. (2012) at 150.9 cm fork length (FL). The growth model was unable to reliably estimate the standard deviation of length-at-age; however, the most appropriate level of variation in length for all age classes was considered to be represented by a coefficient of variation of 0.10 (P. Eveson , CSIRO Marine Research, GPO Box 1538, Hobart, Tasmania 7001, Australia, pers. comm.). The growth function was modelled in SS using age-specific deviates on the *k* growth parameter. Because this feature has only recently been implemented in SS, it is currently not documented (R. Methot, NWFSC, NOAA Fisheries, Seattle, WA 98112 USA pers. comm.) and is one of the first applications within SS to use this feature (Figure 5).

The size at sexual maturity used in the model was equivalent to that applied by Shono et al. (2009) and Kolody et al. (2010). Female fish were assumed to attain sexual maturity at 100 cm FL with full sexual maturity at about 125 cm FL. The length-weight relationship was equivalent to that previously used by Shono et al. (2009) and Kolody et al. (2010) and was originally derived by Nakamura and Uchiyama (1966). Fish weight was determined using the allometric relationship , $a\text{length}^b$, with a = 3.661×10^{-5}, and b = 2.901, where weight is in kilograms and length is in centimetres.

Age-specific natural mortality was equivalent to the schedule used by Shono et al. (2009) and Kolody et al. (2010) where M is 0.8 annually for the early ages and declines to 0.4 by age 3 (quarterly estimates of 0.2 and 0.1 were used in the model). The levels of natural mortality are comparable to IATTC and WCPFC bigeye tuna stock assessments with relatively high natural mortality for the younger age classes and natural mortality of about 0.1 per quarter for the adult age classes.

Recruitment

Recruitment occurs in each quarterly time step of the model. Recruitment was estimated as deviates from the Beverton-Holt stock recruitment relationship (SRR), although deviates were estimated for 1964–2009 only (184 deviates). Recruitment deviates were not estimated for the earlier period of the model due to the lack of longline CPUE indices prior to 1960 and the lack of length frequency data prior to 1965, and recruitment deviates were not estimated for the last eight quarters in the model as there is insufficient CPUE abundance index data to estimate recruitment in later years. Recruitment deviates are assumed to have a standard deviation (σ_R) of 0.6. The steepness (*h*) parameter of the SRR was fixed at a value of 0.8, an intermediate value from the range of values proposed by Harley (2011).

<u>Fishery dynamics and selectivity assumptions</u>

All selectivities were incorporated using the SS-III functionality (see Methot and Wetzel 2013) which can account for different forms of the selectivity function and aggregate fisheries into one form. For all fisheries, selectivity was estimated as an age-based process. Double normal selectivity was assumed for Areas 2 and 3 and logistic selectivity was assumed for Area 1 and the FL2 longline fishery. Selectivity is expressed as:

$$S_{y,f,l} = 1 + e^{\left(-\ln(19)(l_t - \beta_{1,y,f})/\beta_{2,y,f}\right)^{-1}} \qquad \text{(eq. 1)}$$

where S is the proportion selected to the gear, y is the year, f is the fishery, and β_1 and β_2 are parameters related to the size at which 50% (β_1) selectivity occurs for fishery f in year y. β_2 is the difference between the size at 95% selectivity and 50% selectivity for the same fishery and year (Methot 2009).

The selectivities of the LL2, LL3, PSLS and BB fisheries were estimated using a double normal functional form to account for the bimodal length composition of the catch from the PSFS fishery, and the selectivity was modelled using a cubic spline with 6 nodes (Haerdle 1990). Limited data were available to estimate the selectivity of either the PSLS2 or PSFS2 fisheries (i.e., purse seine fisheries operating in Area 2). The selectivity of these fisheries was constrained to be equivalent to the corresponding fishery selectivity in the western region. Selectivity in this case was calculated as:

$$S_{y,f,l} = asc_{y,f,l}\left(1 - j_{1,y,f,l}\right) + j_{1,y,f,l}\left(\left(1 - j_{2,y,f,l}\right) + j_{2,y,f,\gamma,l}\, dsc_{y,f,l}\right) \qquad \text{(eq.2)}$$

where *asc* and *dsc* are the ascending and descending functions of the normal and are described by the joiner functions, $j_{1,y,f,l}$, and $j_{2,y,f,l}$ respectively (for more details see Methot 2009). Five parameters are used to estimate this function and non-informative normal priors are used to bound the parameter estimates. Note that both eq. 1 and 2 could be sex-specific but are not in this case.

Limited size data are available from the “other” fisheries. Initial attempts to estimate independent selectivities for these fisheries were not successful, partly due to the variability in length composition among samples. In aggregate, the length compositions are bimodal and similar to the length composition from the PSFS fishery. On that basis, the selectivities for the two "other" fisheries (OT1 and OT2) were assumed to be equivalent to the PSFS fishery. Similarly, limited length data are available for the mixed gears (hand-line, gillnet/longline combination) fishery in Area 2 (LINE2, Table 1), so the selectivity was assumed to be equivalent to the main longline fishery. Finally, for the single region model, the CPUE indices are linked to the selectivity of the LL1 fishery.

Selectivity Sensitivity Model Runs

A range of models was configured to investigate how selectivity affects the assessment (see Table 2 for more details). The following models were examined:

1) The base case (Scenario 1) model uses a logistic selectivity in Area 1 for the longline fishery and fresh tuna longline fishery. Double normal selectivity functions were used for the longline fisheries in Area 2 and Area 3, and the PSLS in Areas 1 and 2. This function was applied to the BB fishery as well. Finally the PSFS used a cubic spline function. The other categories were modeled with the same cubic spline function, as some of the catches appeared to have the same bimodal functionality as the PSFS. A model run with a larger effective sample size for the length composition data from the LL and PS fisheries was conducted as well (Scenario 5).
2) In scenario 2 and 6, the only change from the Scenario 1 parameterization was that logistic functions were assessed for all longline fisheries, to contrast with the base model which parameterized the selectivity of LL 2 and 3 using the double normal functional form.
3) In scenarios 3 and 7, selectivity was varied over time in three discrete time blocks (1952-1972, 1972-2001, and 2002-2011) to capture some of the changes in size composition observed in the LL fisheries over time (see Figure 8, where some evidence is suggested for changes in the 1970s and then again in the 2000s, Scenarios 3 and 7). The logistic selectivity function was still used in Area 1 for the longline fishery and fresh tuna longline fishery, and the double normal selectivity functions were used for the longline fisheries in Area 2 and Area 3, but time period blocks were used instead of the time-invariant approaches used in Scenario 1 and 5. This approach would translate into three different selectivities for the LL fishery by area and time (a total of 9 fisheries, though Area 2 LL and Area 3 LL share parameters; Scenarios 3 and 7 use these forms and only change the LL selectivity by time using).
4) Scenarios 4 and 8 used time varying selectivity for both the LL and PS fisheries by time and area for the LL and PSLS fisheries (3 different selectivities for the LL fishery by area and time make a total of 9 fisheries, though Area 2 LL and Area 3 LL share parameters, and 3 selectivities for PSLS for a grand total of 12 new fisheries as Area 1 and Area 2 share the

same selectivity function). Note that scenario 4 and 8 analyze the LL and PSLS simultaneously.

5) We also constructed a simple model aggregating all gear-specific fisheries operating in multiple areas into one fishery across the entire Indian Ocean; a total of four fisheries was included in this analysis (LL, PSFS, PSLS and "other"). Cubic spline functions were used for the selectivities, including LL (scenario 9). This model only used recruitment deviates from the period1985-2007, and operated on an annual time step with seasons (at this resolution this is 22 recruitment deviates), while the other models operated on a quarterly time step and estimated deviates for 46 years (at a quarterly resolution this is 184 recruitment deviates).

We examined the effect of weighting the length composition data heavily versus discounting these data by contrasting scenarios 1-4 with scenarios 4-8. An effective sample size of 10 was used across all fisheries in the case of the base run, and an effective sample size of 100 for the LL and PS fisheries was used for weighting the length composition data more heavily (Scenarios 5-9). For the "other" fisheries an ESS of 10 was applied for all scenarios. In all the scenarios examined, the CV for the index of abundance was held constant at 0.1. We compare the MLE estimates of current biomass versus initial biomass over time to demonstrate how these alternative assumptions affect selectivity.

Likelihood profile Analysis

Profile likelihood (Edwards 1992) techniques were used to examine the effect of one parameter, recruitment at virgin biomass levels (R0), and its interaction with selectivity on the overall model fit using the base case scenario as this is informative since M and growth are fixed. Since selectivity is tightly constrained by the LL1 logistic function, different estimates of R0 will reflect the fishing mortality and temporal trends in recruitment from the two main data sets examined in this paper.

Results

For the aggregated fishery (Table 3, Scenario 9), fits to the aggregated length composition data (over all time periods for a single fishery) were reasonable while the fits to individual length observations were very poor, thus this model was disregarded as it failed to account for the length frequencies observed in any of the fisheries over time.

The base model (Scenario 1) exhibits a relatively good fit to the abundance indices (Figure 6a; Table 2, Scenario 1), and a reasonable fit to the length composition data aggregated over time for the main fisheries (Figure 6b; Table 2, Scenario 1). While the average trend is captured reasonably well in the main fisheries, i.e., the longline and purse seine fisheries, the temporal variation is not captured as well (Figure 9). Figure 9 displays the Pearsons residuals where the circles are positive (dark) or negative (light) residuals between the model estimates of length composition and the observed length frequency of the catch in the particular fishery and area. The fits are poor although typical for longline size data, with serious structural change problems through time (Figure 9, Table 2). The time-varying selectivity has very little visible effect on the residual pattern (Figure 9, Table 2).

Table 2 lists diagnostics measuring the goodness of fit, namely the negative log-likelihood values of the different components of the model, the length frequency component and the survey component (the catch component is left out as it is often negligible in these models). Note that the model of the four aggregated fisheries (Scenario 9 in Table 2 which had 52 parameters) is not comparable to the other models presented (Table 2) unless we use other statistics like AIC (Akaiki 1983), though we do report reference points obtained from the aggregated model (Table 3, scenario 9).

In the case of using the logistic versus the double normal functions for the LL fishery in Areas 2 and 3 (Figure 7), the fits were comparable (Table 2, Scenario 2) though the base model fit is better than the logistic fit (Table 2) based on the overall likelihood values. While the logistic function may not result in as good a fit as the double normal function, it also has fewer parameters to estimate. In addition, the logistic function does not have any problems with denoting a cryptic biomass as all age classes are fully vulnerable after a certain age. This is not the case with a dome-shaped selectivity curve like the double normal curve. In our case, this is not a problem as there is still one fishery with logistic selectivity within the model domain. While the logistic function fits the data because full selectivity occurs at a much younger age, the double normal function achieves this fit by including fewer older fish in the catch.

Due to variation in the length composition data over time (Figure 8) among fleets, using temporally variable selectivity with a higher weight on length frequency data gave marginally better fits for many fisheries (Figure 9) but may have resulted in poorer fits for some fisheries (PSLS1). The fits compare models with lower effective sample size, i.e. in this case the base model with lower effective sample size, and the time varying selectivity models with lower effective sample size (Scenario 1 and 4 respectively). This was done so we could take the confounding effect of effective sample size out of the time varying selectivity component. A similar exercise was conducted for models with higher effective sample size, but the results showed similar trends as we have here.

In essence the overall shape of the selectivity curve does not change much over time (Figure 10), though the LL fishery appears to be selecting older fish over time. This pattern does not change when heavier weight is given to size-selectivity data. Note, however thatif these data were weighted with a large effective sample size, it would have a large impact on the overall assessment. Even though the temporally-variable component was added in block format, the residual patterns still remain quite similar (Figure 9). The only noticeable change is that the magnitude of the residuals decreases when the time varying component is added for the LL fishery in the 1970s and after 2003, when different selectivity blocks are introduced in the fitting procedure. It could be concluded that the time varying selectivity blocks are not adequate to account for the variation in size over time and the apparent inconsistency between the length-frequency data and the index of abundance.

To examine the influence of the different sets of length-composition data on the estimation of recruitment at virgin biomass levels R0, we conducted a likelihood profile analysis. Since M and growth are fixed and selectivity tightly constrained by the LL1 logistic function, different estimates of R0 will reflect the fishing mortality and temporal trends in recruitment from the two main data sets. It is evident from the likelihood profiles that there is somewhat conflicting information about R0 from the two data sets, with the overall length data indicating lower overall stock size (Figure 11). Further examination indicates that the longline and fresh tuna longline fisheries in these areas (primarily in Areas 1 and 2, figure not shown) are having the largest influence on the R0 parameter, primarily because there are few older fish in the populations. Any biases in these length frequencies, mis-specification of the selectivity curve, or use of incorrect M, can have huge effects on the assessment (Figure11) as is evident with the estimates of target reference points S_{MSY} and Yield (Table 3, Figure 12).

A comparison of the different sets of models with similar effective sample size weightings (Table 2) shows marginal improvements in the overall log-likelihood values for models with the same effective sample size (sensitivity 1 to 4 and sensitivity 5 to 8 respectively). The log-likelihood values from models with different effective sample size weightings cannot be directly compared, but these diagnostics can be used in a qualitative manner to assess how the fits may be improving (accounting for the differences in parameters estimated) for cases with the same effective sample size for the length frequency data. A large effect on the assessment dynamics (Figure 12) is observed by weighting the length composition data higher (assigning a sensitivity of 5 versus 1). Adding temporally-variable selectivity (Table 2, sensitivity 3, 4 or sensitivity 7, 8) or different functional forms (Table 2, sensitivity 2 or sensitivity 6), resulting in a marginal improvement in fit to the length frequency data, has minor effects on management parameters obtained from the assessment (for scenarios with the same effective sample size for the length frequency data).

Finally, when we contrasted the model runs with higher and lower effective sample sizes (Figure 13) and examined how this weighting relates to the fits to the different datasets used in the model, we found contradictory signals between the index of abundance and the length frequency data. While in the analysis presented here, the effect of these different formulations on the assessment is negligible in terms of the current stock status, this effect may not be negligible in other global assessments. All scenarios investigated indicate that the current stock size (B_{2011}) is greater than B_{MSY} (low effective sample size scenarios; between 1.3-2.3), and current fishing mortality rates are below F_{MSY}. Weighting the length-composition data more heavily produces a more pessimistic picture of the stock, i.e., B_{2011}/B_{MSY} is between 1.14 and 1.94, and $\boldsymbol{F_{2011}/F_{MSY}}$ is between 0.41 and 0.66 for all combinations examined. Consequentially target optimal yields are 160-270 kT from the lower effective sample size combinations versus 93kT -115 kT from the higher effective sample size combinations.

Discussion

Francis (2011) suggests that in cases where there are contradictory signals between the length-composition data and the index of abundance, it is almost always more important to capture the index of abundance over the length composition data. Fitting to the length-composition rather than the abundance data may be misleading, as this approach gives more weight to processes of which we have a poor understanding, and which are also correlated with the overall abundance signal (Francis 2011). In most integrated models used across the globe, it is the nature of the modeler/scientist to use all the data. However, as shown in this case, it is extremely important to understand the data before they are used in the assessment. If the length frequency data are incorrect, then the resulting model will be biased, and management advice resulting from the model will be incorrect as well.

This paper illustrates a sequential approach on how to model selectivity, the potential impacts of different selectivity assumptions, and effects of these assumptions on stock assessment outputs. Three different critical issues for analysing selectivity are illustrated here: 1) the form of the curve, 2) temporally-variable selectivity and its appropriate use, and 3) the effective sample size of the length-composition data that affects the estimation of key parameters in the assessment and in turn their effect on derived parameters from the assessment. While most of the results were similar when comparing changes in selectivity assumptions, the analysis highlights the importance of weighting the length-composition data. The weighting issue is also related to fishing mortality assumed in the assessment, and recruitment deviation estimated in the model, but this question is not examined here extensively.

While developing the models here, a simpler approach was initially taken (Table 1, Scenario 9). However, this model was abandoned as it did not capture changes in length composition over time. This outcome was not surprising as the fisheries were aggregated, and using a cubic spline function for selectivity, while allowing flexibility that captures bimodal length frequencies, still fails to account for all the complexities in the temporal dynamics of the fishery. Thus, keeping the fisheries at least separated geographically may more accurately reflect fishery dynamics, provided the data are unbiased and consistently sampled over time, and so this approach was used in subsequent model runs (Crone et al. this edition).

Shape of the Selectivity Curve Used and its Effects

We conclude from our analyses that while the shape of the selectivity function is important in these fisheries, it does not have a significant effect on either the fits to the length-composition data or to the abundance index data, unless it is weighted heavily. Our model run outputs are likely too variable between time steps and among fisheries to reliably fit the contradictory signals, and the estimated selectivity simply reflects the average pattern of fishery exploitation. While a possible improvement may be to fit the length-composition data using selectivity deviates for each time step, that approach might be tantamount to fitting the model to the error in the observations. The use of a logistic function vs. a double normal function has only marginal effects on the overall results, unless we impose a larger effective sample size on the data (Table 3, Figure 7) which forces the model to fit to the length-frequency component. Even then the shape does not differ substantially as is evident in Figure 10.

The issue of whether we use temporally variable selectivity is intricately confounded by the length-frequency data sampled. It is highly unlikely that the fisheries, particularly the sizes targeted, have remained unchanged in the last 50 years, therefore using time-invariant selectivity will probably bias our estimates. However, as is demonstrated here, if the length-composition data are inconsistently measured over time (Figure 14), then putting unduly heavy weight on this parameter can also bias the assessment. Finally, adding temporally variable selectivity gives an overly optimistic picture of the stock status in the case when we don’t weight the length composition data (Figure 13, Scenario 3 and 4; Table 3), and a pessimistic view of the stock in the case where the length-composition data are weighted heavily (Figure 13, Scenario 7 and 8, Table 3). In our analysis we chose to put a lower weight on the length composition data rather than give undue weight to the length composition data based on criteria stated in Francis (2011). As a result of this approach we estimate different levels of biomass than if we chose to weight the length-composition data higher (Figure 13).This discrepancy occurs primarily because the length frequency data from the longline and fresh tuna fishery in Areas 1 and 2 are highly influential in the model. Since natural mortality (M) and growth are fixed, and selectivity tightly constrained by either the logistic or double normal function, a different R0 and recruitment deviates are estimated to better fit to the length observations. Hence we observe the lower R0 values, an increase in recruitment in the 1970s, and increases in recruitment in the 2000s.

The modeled biomass trajectories resulting from the different combinations of input data (Figure 13) have different outcomes as far as management advice is concerned. The issue with the estimated selectivity curves (Figure 7 and Figure 10) that indicate a sharp decline in stock at around age 7 is common when using the double normal function as this approach can cause convergence issues. This problem typically occurs as the growth curve flattens off by age 7 (Figure 5). In addition, we see very few fish over this size in the length composition data (Figure 8), causing these curves to exhibit the declining limb, which would mean we still have a large number of older fish in the population (M values used imply enough older fish available in the population).

Data Issues Related to Estimated Length Compositions and Selectivity

It is evident that either changes are occurring in the fishery over time (Figure 8 and Figure 15), or there are problems with the sampling coverage after the year 2000. It is more likely that something changed in the sampling regime after 2000 (Geehan et al. 2013). Further examination of the data (Figures 14 and 15) indicates that the sampling coverage from the main longline fisheries has changed during the period examined. In the 1970s and 1980s the main source of information for the length frequency data (Figure 14b) was the Japanese longline fleet. However, in recent years these data have been almost entirely contributed by Taiwan (Figure 14a). In addition the Taiwanese data collection techniques and systems changed in the late 1990s (Chang and Wang 1998, Dr. S. Wang personal comm.). Based on the sampling data it appears that larger fish are now encountered in larger proportions than prior to the 2000s, and the overall variance in length-composition in the catch data has also decreased. These observations indicate that there are inconsistencies in the length composition data over time, thus the data series may not be entirely reliable. Because close to 90% of the overall catch comes from these fleets, bias in these data has a large effect on the results of the overall assessment.

Based on these observations, two alternative hypothesis may be proposed: 1) the length frequency data are reliable and the stock is much less abundant than previously estimated (Table 3, Figure 13), or 2) the stock is somewhat more abundant, the abundance index data are more reliable, and the length frequency data are not as reliable (Figure 13, Table 3). A third alternative could also be examined (though not in this paper): the data are reliable until the early 2000s when the sampling regime changed. Under this scenario, we could fit the model to the length-frequency data until 2000, and then only to the abundance index data after that. However, an examination of the earlier fits to the length frequency data prior to 2000 (Figure 9, base case) still reveals a poor fit to the length-frequency samples in those years. Another alternative would be to fix selectivity at the values estimated through the logistic or double normal functions and then fit to the abundance data iteratively (Iwata et al. this edition). This approach would ensure that we accounted for the catch being taken out at the right size and fit to the overall abundance, without worrying about fitting to the length-frequency component as well.

In deciding between alternatives (1), and (2) stated above, Francis (2011) suggests weighting the index data more than the length composition data (which we do, Table 2, Figure 12). Until we have a thorough understanding of the length composition data used in tuning these models, we are forced to rely on the index of abundance data that is compiled and analyzed by the LL fleets of Japan. More effort may also be required to understand why the index of abundance jumped in the mid-1970s. Bigelow et al. (2002) suggest that fleet targeting could be the reason for this increase, meaning that the gear was set at a different depth specifically to catch BET. Alternative standardization procedures using different depths of fishing as a covariate, and the correct assigning of species compositions (these were problematic prior to the 1970s, Herrera et al. 2012) during that period could provide alternative indices of abundance (Bigelow et al. 2002). However, since this approach would require a thorough re-examination of the operational data, it is beyond the scope of this study.

Management Implications of Selectivity and Weighting of Data

While size-selectivity is extremely important and directly related to estimates of optimal yield, target spawning stock sizes, and overall fishing mortality (F), the results from this study indicate that the functional forms have little effect after the majority of the fish become mature. Whether we use a logistic or double normal function (declining limb of selectivity) the overall effect is marginal, other than when we use temporally variable selectivity with lower weights on the length composition data, where the reference points and yields are significantly higher than the other scenarios (Figure 13, Table 3). For convergence issues, one needs at least one asymptotic form, which we used in LL 1. As Wang et al. (this edition) show, the functional form of the selectivity (in their case asymptotic) could cause biases as it is confounded with the maximum length and the observed length-compositions observed in the fishery. While temporally variable selectivity fits the length-composition data better, if the length composition is weighed higher, it has a large effect on R0 (Figure 11) and the overall spawning stock biomass trajectories (Figure 13), indicating a declining recruitment trend and possibly lower optimal spawning stock sizes. The model formulations with the logistic function also estimate a larger fishing mortality rate compared to models with a declining limb in selectivity for obvious reasons.

Further Improvements in Model Formulations of Selectivity

Further refinement of the fisheries by area and jurisdiction could provide the model with more consistent data for length frequencies and indices of abundance. In addition, a length frequency data set that accounts for changes in targeting could also improve the fit of the model, as the stock could be modeled as separate fisheries based on when these targeting changes occurred. If selectivity of the fleets has changed over time (which is likely), using the temporally-variable component makes sense. However, if there is no drastic change in the length-composition data, and we use the same functional form, we get very slight differences in these curves as seen in Figures 7 and 10. Modeling these as entirely different fisheries, with different functional forms (double normal versus logistic), could possibly produce improvements in the assessment, provided we have some external source of information indicating that these fisheries operated very differently in these two periods.

Alternative model assumptions such as considering the appropriateness of age- vs. length-based selectivity should also be undertaken (though an analysis to look at these effects, not presented here, gave very similar results). It is also important to consider reliability of other key fixed model parameters, especially growth and M-at-age, which have a large impact on selectivity and available biomass at any given time. Other key processes that may be affecting this assessment are regional stock structure and spatial differences in recruitment processes; these should be examined along with the selectivity assumptions presented here.

Finally, the use of iterative reweighting approaches as suggested by Iwata et al. (this edition) is another alternative to improving the selectivity and not fitting to the catch so that we can take the catch out at the appropriate age without fitting to the length-composition data. This approach may not be helpful if sampling is indeed biased, as is plausible in this case (Figure 15), as the mean length of the catch appears to increase over time while the variability decreases over time (kurtosis decreases). However, one of our key conclusions is that it is critical to first understand whether the changes in length-frequencies are real or biased. This is a critical step, given the importance of these data in influencing the overall assessment results.

Acknowledgements

The authors would like to thank Mark Maunder and Alex Aires Da Silva for helpful comments. In addition the authors would like to acknowledge Ian Taylor and Rick Methot on help with SS-III.

References

Akaiki, H. 1983. Information measures and model selection. Int. Stat. Inst. 22: 277-291.

Bigelow, K.A., Hampton, J., and Miyabe, N. 2002. Application of a habitat-based model to estimate effective longline fishing effort and relative abundance of Pacific bigeye tuna (*Thunnus obesus*). Fish. Oceanogr. 11:3 143–155.

Chung, S.K. and Wang, S. B. 1998. Review of data collection and processing system, and revision of statistics for the Tawianese deep-sea longline fishery operated in the Indian Ocean. IOTC Proceedings No. 1.

Crone, P. R., M. N. Maunder, B. X. Semmens, and J. L. Valero (editors). *In preparation*. Selectivity: theory, estimation and application in fishery stock assessment models. Fisheries Research.

Deriso, R.B. and Parma, A.M. 1987. Dynamics of Age and Size for a Stochastic Population Model. Can. J. Fish. Aquat. Sci. 45: 1054 1068.

Edwards.A. W. F. .1992. Likelihood. Johns Hopkins University Press. 296 p.

Eveson, P., Million, J., Sardenne, F., Le Croizier, G. 2012. Updated growth estimates for skipjack, yellowfin and bigeye tuna in the Indian Ocean using the most recent tag-recapture and otolith data. IOTC-2012-WPTT14-23.

Fournier, D. and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39: 1195-1207.

Fournier, D. A., H. J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen, and J. Sibert. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw. 27:233-249. Visit the ADMB project website: www.ADMB-project.org.

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68:1,124-1,138.

Gavaris, S. and Ianelli, J. 2002. Statistical issues in fisheries stock assessments. Scandinavian Jr. of Stats. Vol. 29: 245-267.

Geehan, J. Hoyle, S. and Herrera, M. 2013. Review of length-composition data of Taiwanese Longline Fisheries. IOTC–WPTT15.

Harley, S.J. 2011. Preliminary examination of steepness in tunas based on stock assessment results. WCPFC SC7 SA IP-8, Pohnpei, Federated States of Micronesia, 9–17 August 2011.

Herrera,M., Gillett, B. and Pepperell, J. (*in Press*). Tunas and Their Fisheries: safeguarding sustainability in the 21st century *In*” Small-scale fisheries (artisanal, local subsistence and recreational fishing) *Eds* Gillett,B. and Pepperell, J. Wiley-Blackwell.

Herrera, M. , Pierre, L., Geehan, J. and Million, J. 2012. Review of the Statistical Data and Fishery trends for Tropical Tunas. IOTC-2012- WPTT14–07.

Haddon, M. 2011. Modelling and Quantitative Methods in Fisheries. Chapman & Hall, Inc. USA

Haerdle, W. 1990. Smoothing techniques with implementation in S. Springer-Verlag, New York.

Hilborn R, Walters C. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, Inc., USA.

Hoyle, S.D, and Okamoto, H. 2011. Analyses of Japanese longline operational catch and effort for bigeye and yellowfin tuna in the WCPO. WCPFC-SC7-2011/SA IP-01. Scientific Committee Seventh regular session, 9-17 August 2011. Pohnpei, Federated States of Micronesia.

IOTC 2011. Report of the Thirteenth Session of the IOTC Working Party on Tropical Tunas. IOTC–2011–WPTT13–R.

IOTC 2012. Report of the Fourteenth Session of the IOTC Working Party on Tropical Tunas. IOTC–2012–WPTT14–R.

Iwata, S., Kitakado, T. and Takeuchi, Y. 2013. Estimation of selectivity in Stock Synthesis: lessons learned from the tuna stock assessment. *Fisheries Research this edition*

Kolody, D., Herrera, M., Million, J., (2010). Exploration of Indian Ocean Bigeye Tuna Stock Assessment Sensitivities 1952-2008 using Stock Synthesis (updated to include 2009). IOTC-2012-WPTT10-4.

Langley, A. Herrera, M. and Sharma, R. 2013. Stock Assessment for Bigeye Tuna in the Indian Ocean for 2012. IOTC Working Party Document.

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

Methot, R. D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. ***In*** Mathematical analysis of fish stock dynamics: reviews, evaluations, and current applications, E. F. Edwards and B. A.Megrey, eds. ***Ap.*** Fish. SOC. Symp. Ser. 6, pp. 66-82.

Methot, R.D. 2009. User Manual for Stock Synthesis. NOAA Fisheries, Seattle, USA.

Nakamura, E.L. and J.H. Uchiyama. 1966. Length-weight relations of Pacific tunas. *In* Manar, T.A. (editor), Proc., Governor’s [Hawaii] Conf. Cent. Pacif. Fish. Resources: 197-201.

Nishida, T., and Rademeyer, R. (2011). Stock and risk assessments on bigeye tuna (*Thunnus obesus*) in the Indian Ocean based on AD Model Builder implemented Age-Structured Production Model (ASPM). IOTC-2011-WPTT-42.

Punt, A. and Hilborn, R. 1997. Fisheries Stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries 7: 35-63.

Quinn, T. J. 2003. Ruminations on the development and future of population dynamics models in fisheries. Nat. Res. Modelling 16: 341-391.

Satoh, K.; Okamoto, H. 2012. Updated Japanese longline CPUE for bigeye tuna in the Indian Ocean up to 2011 standardized by GLM. IOTC-2012-WPTT14-26.

Shono, H., K. Satoh, H. Okamoto, and T. Nishida. (2009). Updated stock assessment for bigeye tuna in the Indian Ocean up to 2008 using Stock Synthesis III (SS3). IOTC-2009-WPTT-20.

Wang. S. Maunder, M. and Aires-Da-Silva, A. 2013. Selectivity's distortion of the production function and its influence on management advice. *Fisheries Research this edition.*

List of Tables

List of Figures

Figure 14. Length frequency data trends in average catch and sample sizes from Japan (fig. 14a) and Taiwan, China (fig. 14b) longline fleets over the time period used in the assessment.

Figure 15**:** Comparison of BET length frequency distributions of Tawianese and Japanese Longline fleets. The size of the green proportional circles indicates the sampling coverage in each year, which are scaled relative to the minimum sampling coverage of one fish per tonne of catch recommended by the Indian Ocean Tuna Commission (denoted by the black circle at the centre of each proportional circle). Years with proportional circles larger than the minimum sampling standard indicate relatively high sampling coverage in a given year; the larger the circle the higher the sampling coverage. The crosshairs indicate the average length of sampled cohorts in each year class.

Table 1. Definitions of the individual model fisheries

Code	Method	Region
FL2	Longline, fresh tuna fisheries	2
LL1	Longline, distant water	1
LL2	Longline, distant water	2
LL3	Longline, distant water	3
PSFS1	Purse seine, free school	1
PSFS2	Purse seine, free school	2
PSLS1	Purse seine, associated sets	1
PSLS2	Purse seine, associated sets	2
BB1	Baitboat and small scale encircling gears (PSS, RN)	1
LINE2	Mixed gears (hand-line, gillnet/longline combination)	2
OT1	Other (trolling, gillnet, unclassified)	1
OT2	Other (trolling, gillnet, unclassified)	2

Table 2. Negative Log likelihood values of the different selectivity sensitivity runs of the different components based on the effective sample sizes of the length composition data.

Scenario Number	Description	Log Likelihood	Survey LL	Length Comp -LL	No. of parameters estimated
Scenario 1	Base case	4257	304	4005	207
Scenario 2	Logistic LL Sel (Low Eff SS)	4316	303	4065	204
Scenario 3	Time varying Sel-LL (Low Eff SS)	4126	289	3877	223
Scenario 4	Time Varying LL and PSLS Fishery (Low Eff SS)	4112	288	3863	235
Scenario 5	Base case (higher eff SS)	21737	709	21003	207
Scenario 6	Logistic LL Sel (High Eff SS)	22149	730	21377	204
Scenario 7	Time varying Sel-LL (High Eff SS)	20940	524	20940	223
Scenario 8	Time varying Sel-LL & PS(High Eff SS)	20777	522	20201	235
Scenario 9	Aggregated Fishery Model	3866	882	2502	52

Table 3. Maximum Posterior Density (MPD) estimates from the final set of model options and associated model sensitivities. The preferred (reference) model option is highlighted.

Scenario	Selectivity Sensitivity Run	SB_0	SB_{MSY}	SB_{2011}	SB_{2011}/SB_0	SB_{2011}/SB_{MSY}	F_{2011}/F_{MSY}	MSY
1	***Base case***	***1606040***	***446970***	***606830.3***	***0.38***	1.36	***0.31***	179370
2	Logistic LL Sel (Low Eff SS)	1455880	409173	534456	0.37	1.31	0.35	163248
3	Time varying Sel-LL (Low Eff SS)	2493330	705780	1622145	0.65	2.30	0.14	273752
4	LL and PSLS Fishery (Low Eff SS)	2432020	696212	1598100	0.66	2.30	0.15	259886
5	Base case (higher eff SS)	1082610	300837	343326.8	0.32	1.14	0.51	114591
6	Logistic LL Sel (High Eff SS)	1323390	243942	343327	0.26	1.41	0.66	93973
7	Time varying Sel-LL (High Eff SS)	948687	270865	525487	0.55	1.94	0.39	106121
8	Time varying Sel-LL & PS(High Eff SS)	929923	268817	528458	0.57	1.97	0.41	96629
9	Aggregated Fisheries	938225	165640	421377	0.45	2.54	0.24	193240

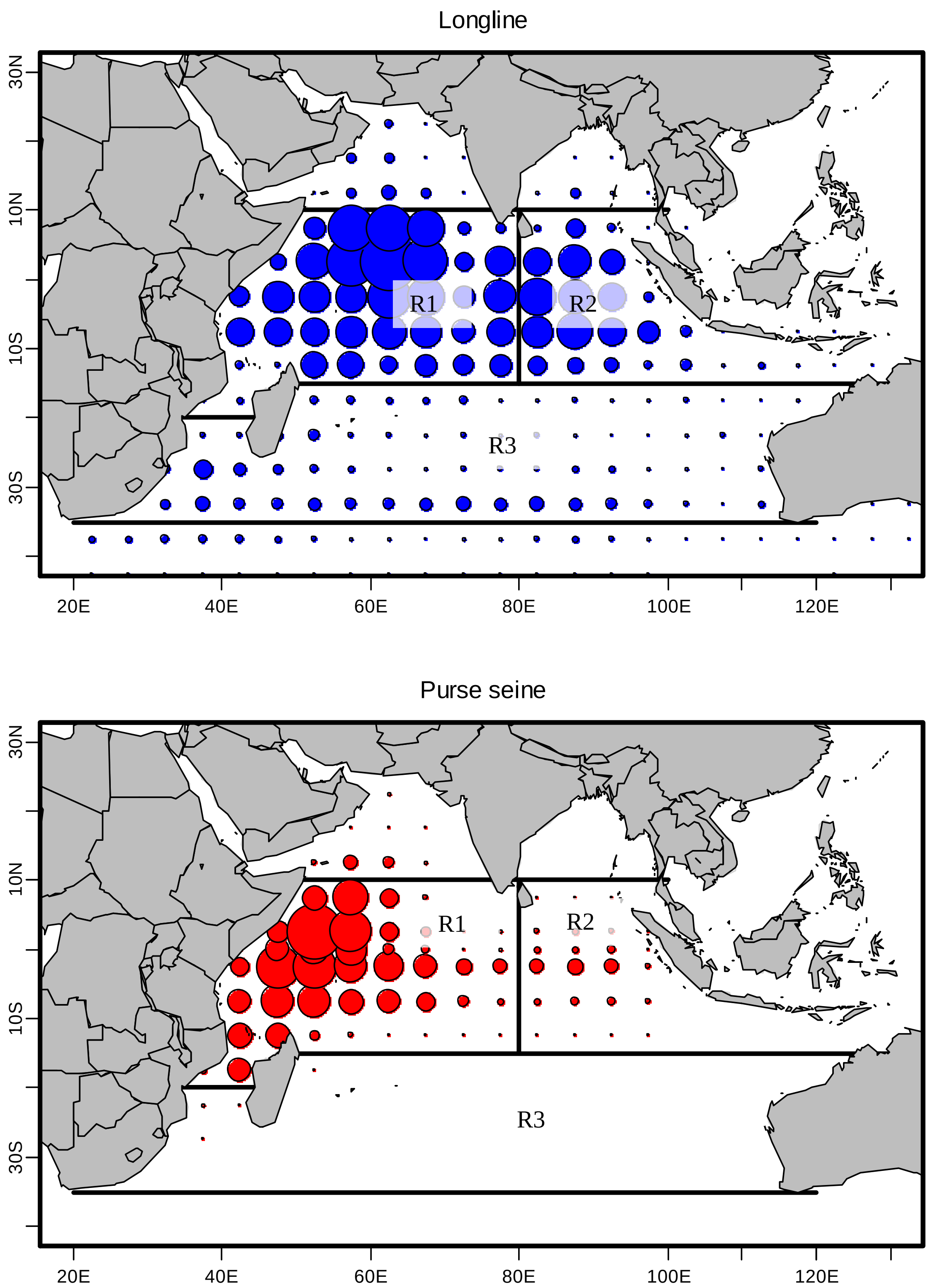


Figure 1. Aggregate LL and PS catch (max bet catch in 5 deg cell aggregated over time 70159.62 mt)

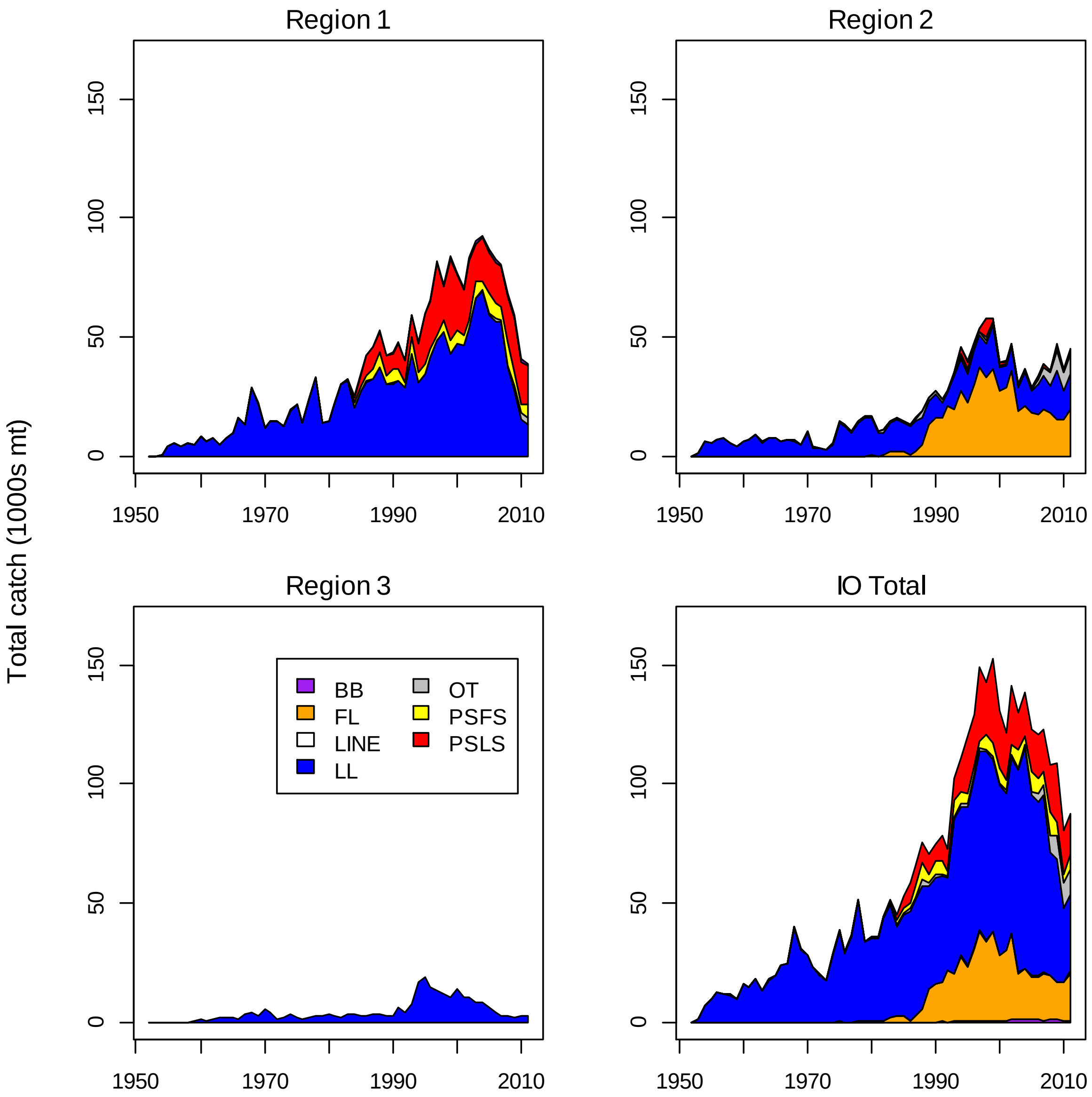


Figure 2. Annual catches by fishery and region.

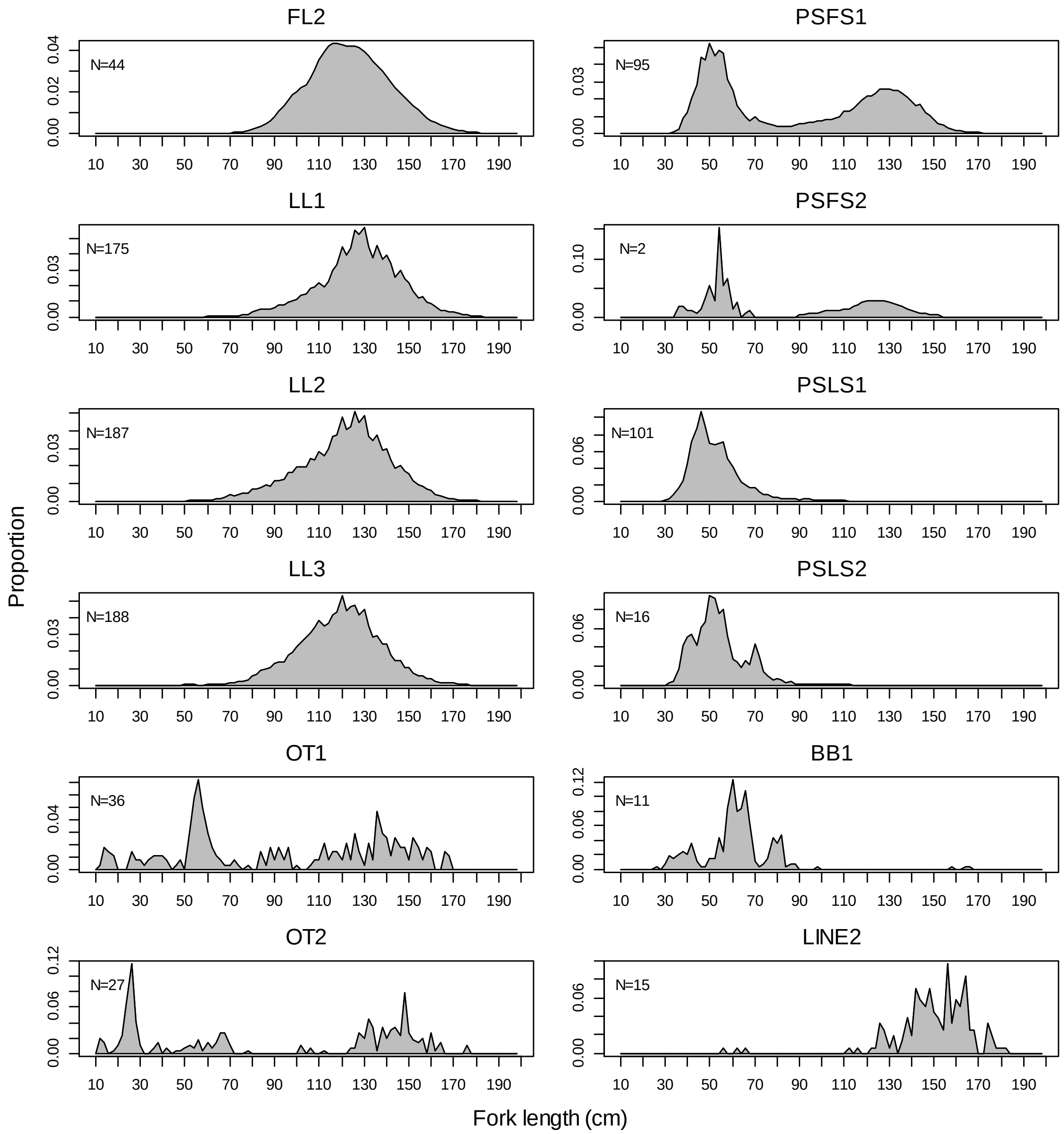


Figure 3. Length compositions of bigeye tuna samples aggregated by fishery. N represents the number of quarterly samples.

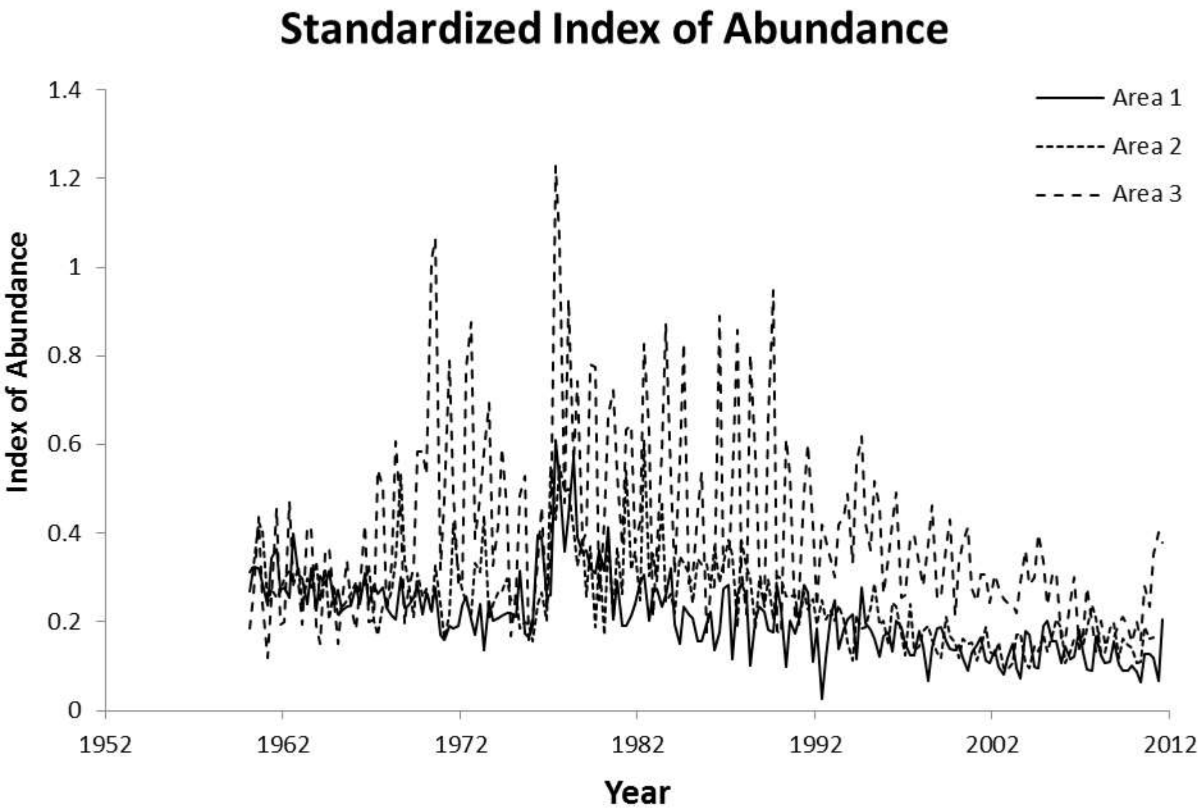


Figure 4. Standardised longline CPUE indices for each region. The longline series are standardised among regions

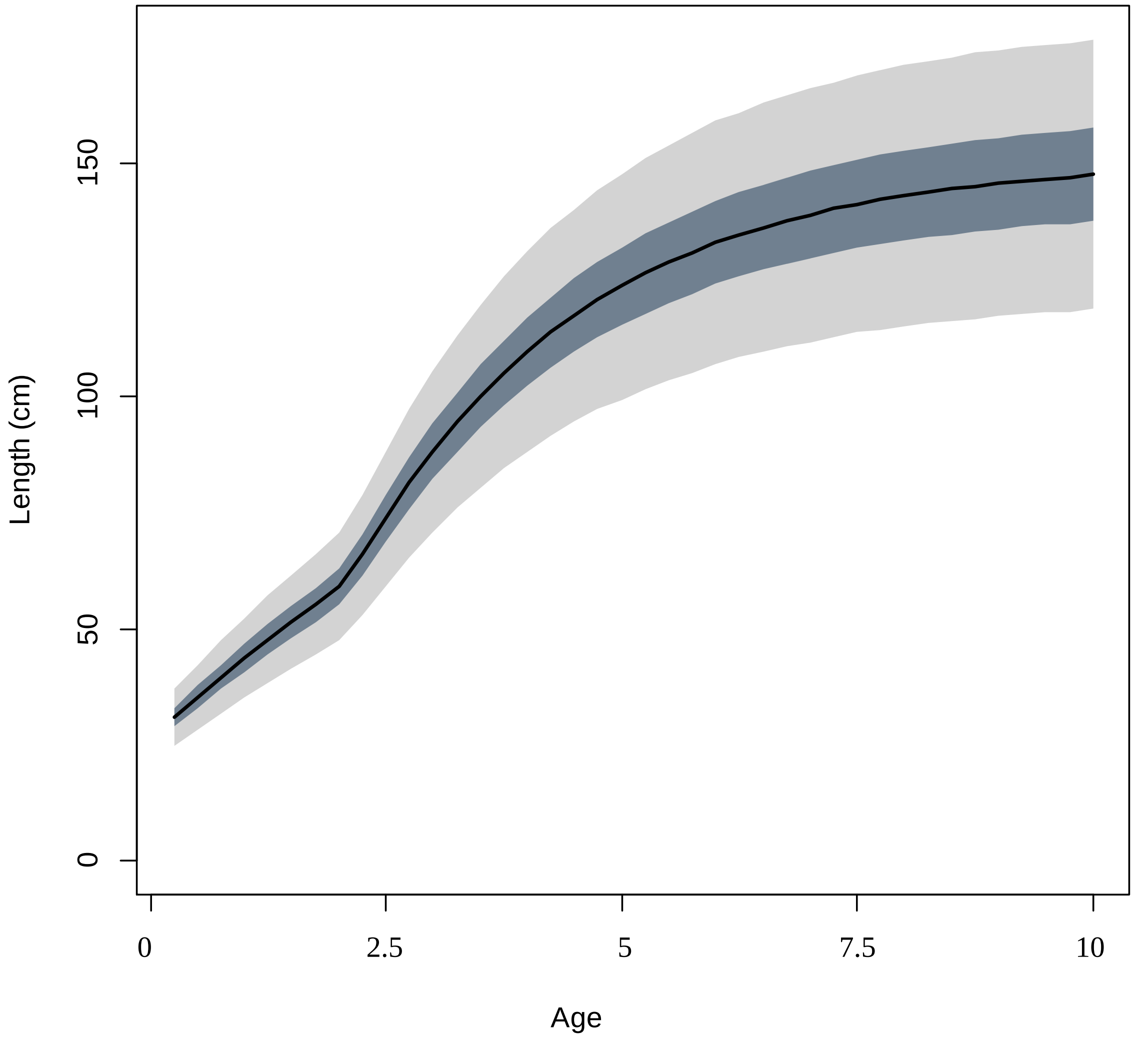


Figure 5. Growth of Indian Ocean bigeye tuna (following Everson et al 2012). The dark grey region represents the quartile range of the distribution of length-at-age and the light grey region represents the 95% confidence interval.

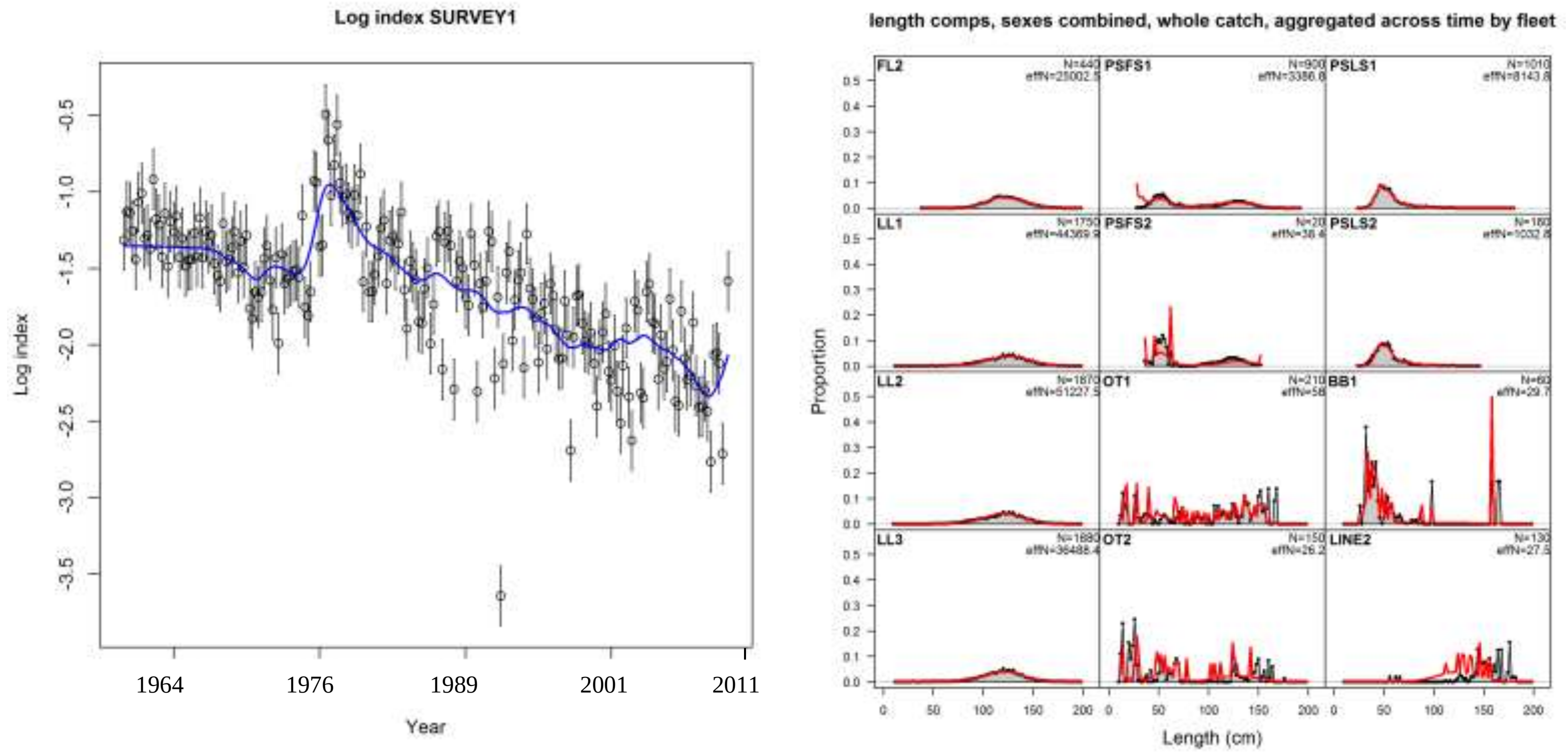


Figure 6. Fit of the base model the index of abundance as obtained from the Japanese logline fishery and the average length compostion from the data aggregated across time by each fishery separately

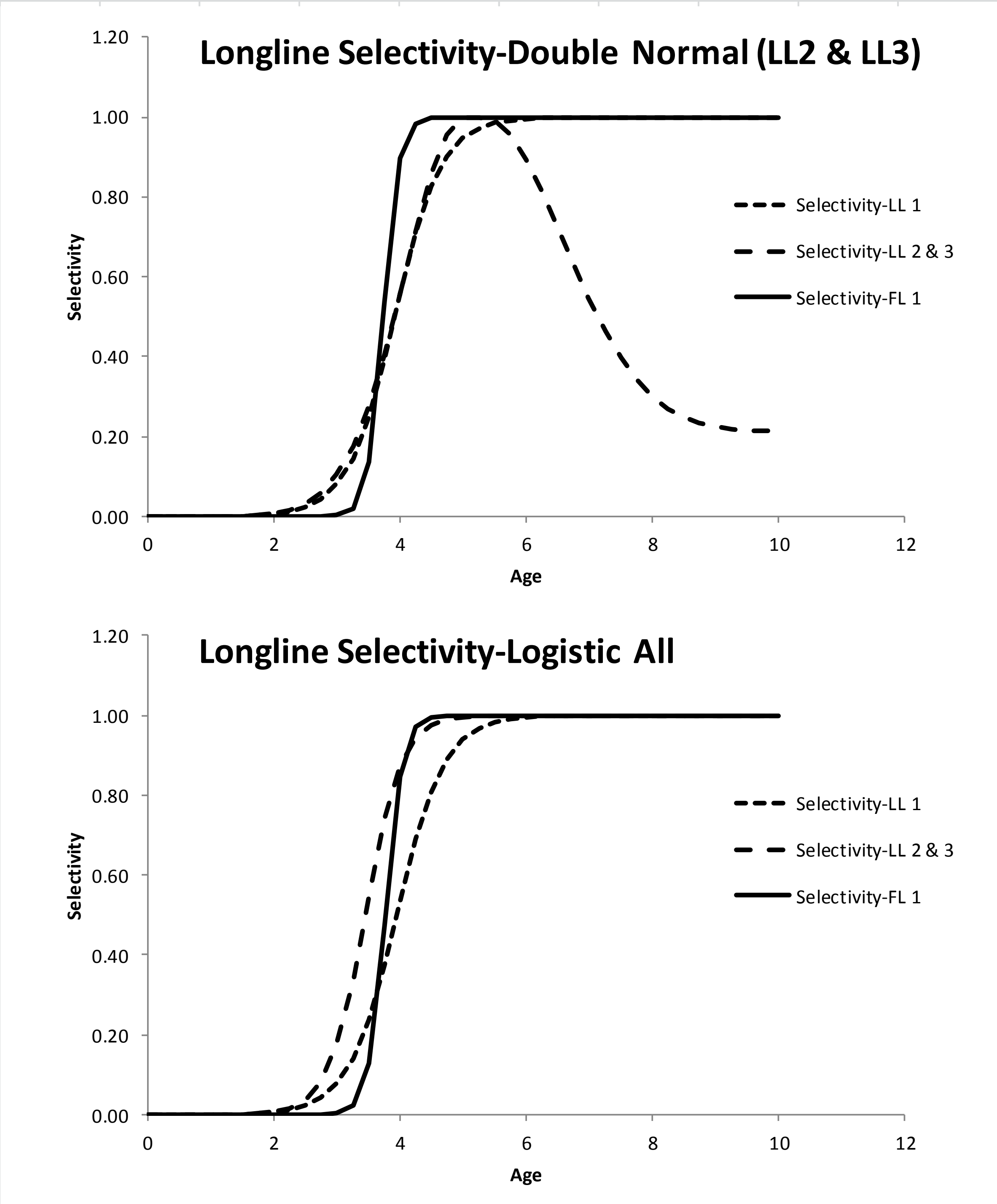


Figure 7. Logistic versus Double Normal Selectivity estimated in LLand Fresh Tuna (FL) fisheries in area 1, 2 and 3. Note LL 2 and 3 selectivity is modelled identically.

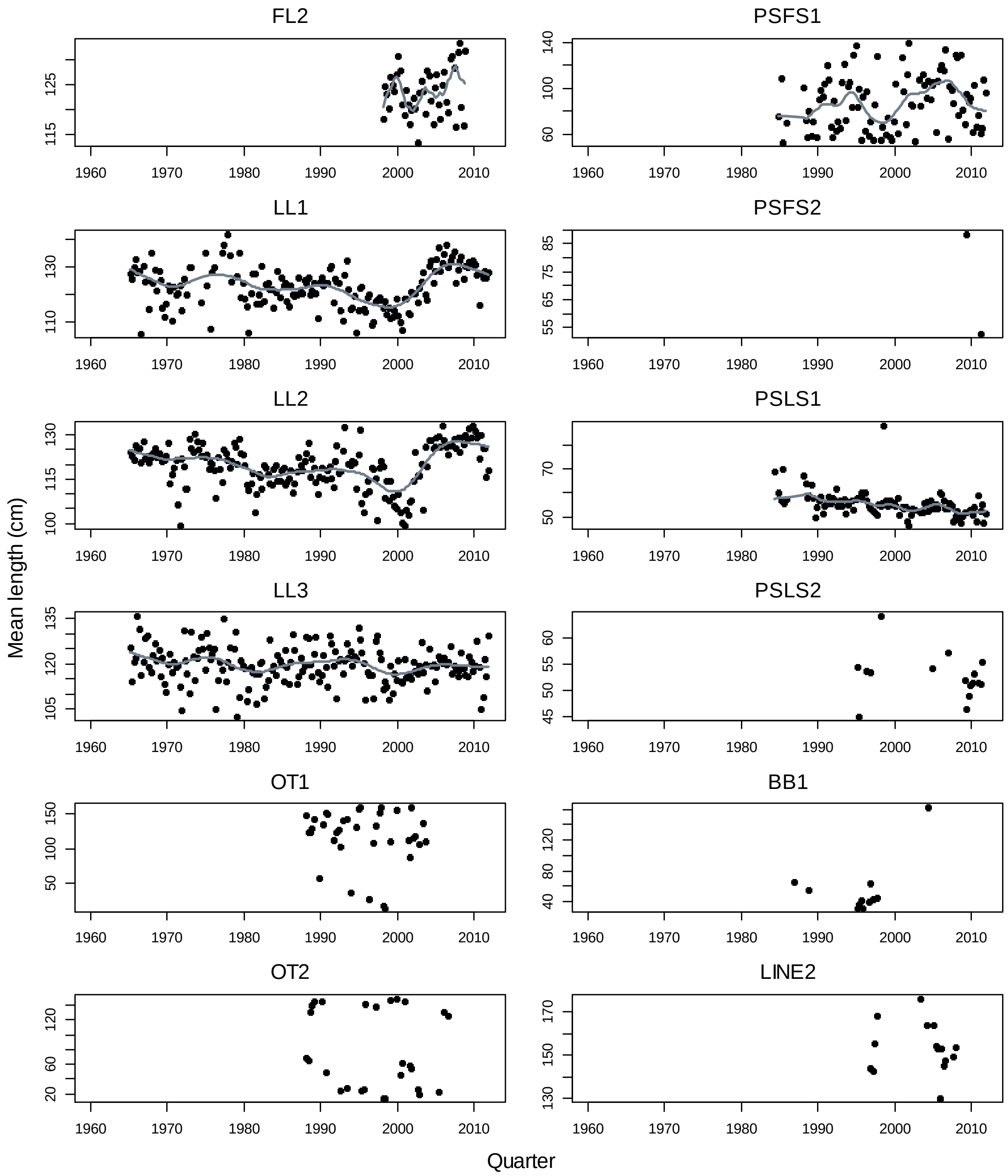


Figure 8. The average length (fork length, cm) of bigeye in the individual samples from each fishery. The grey line represents a lowess smoothed trend. The y-axis differs among the individual plots.

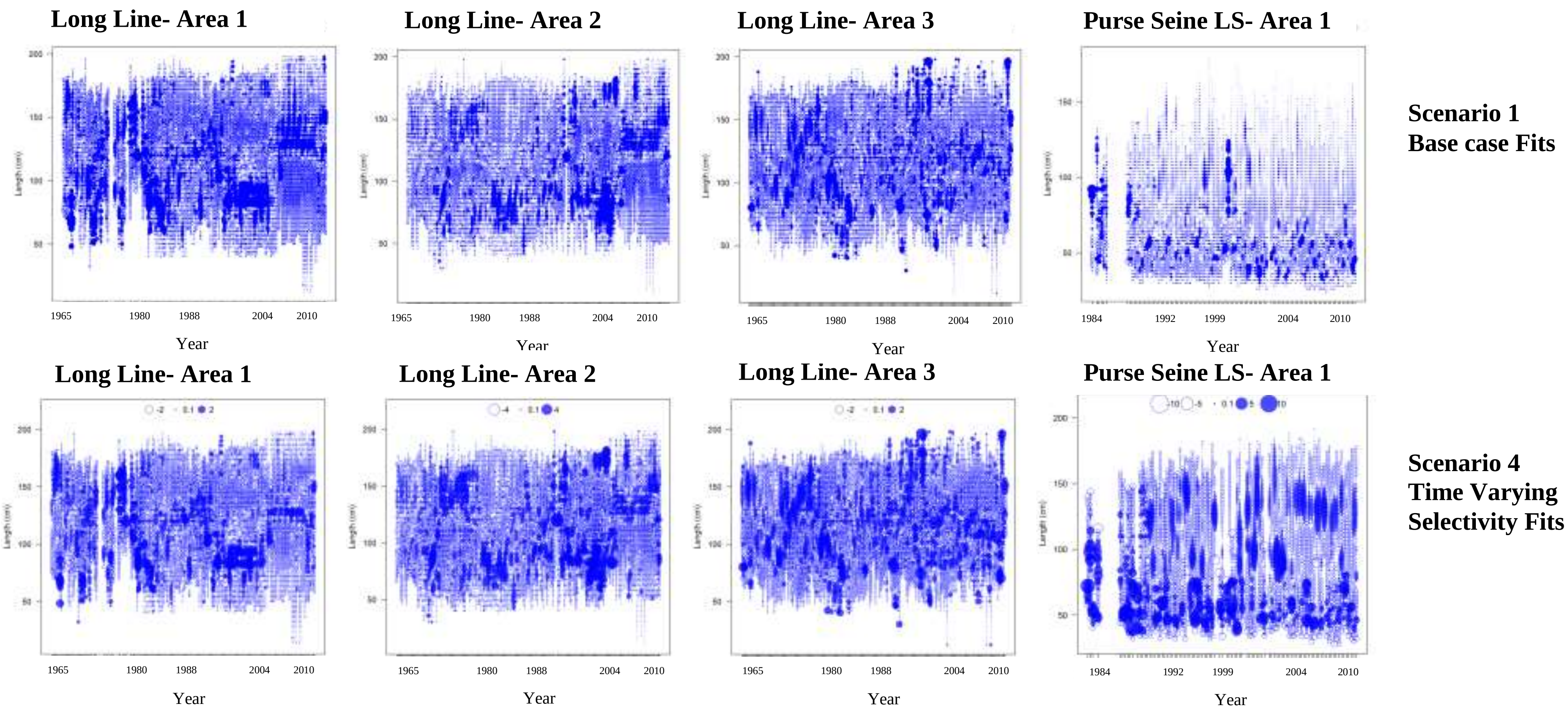


Figure 9: Marginal improvements in Pearsons residuals for length composition data FL2, LL 2 and PSFS 1when fitting time varying selectivity in all the LL and PS fisheries by the blocks 1952- 1971, 1972-2001, and 2002-2011 respectively

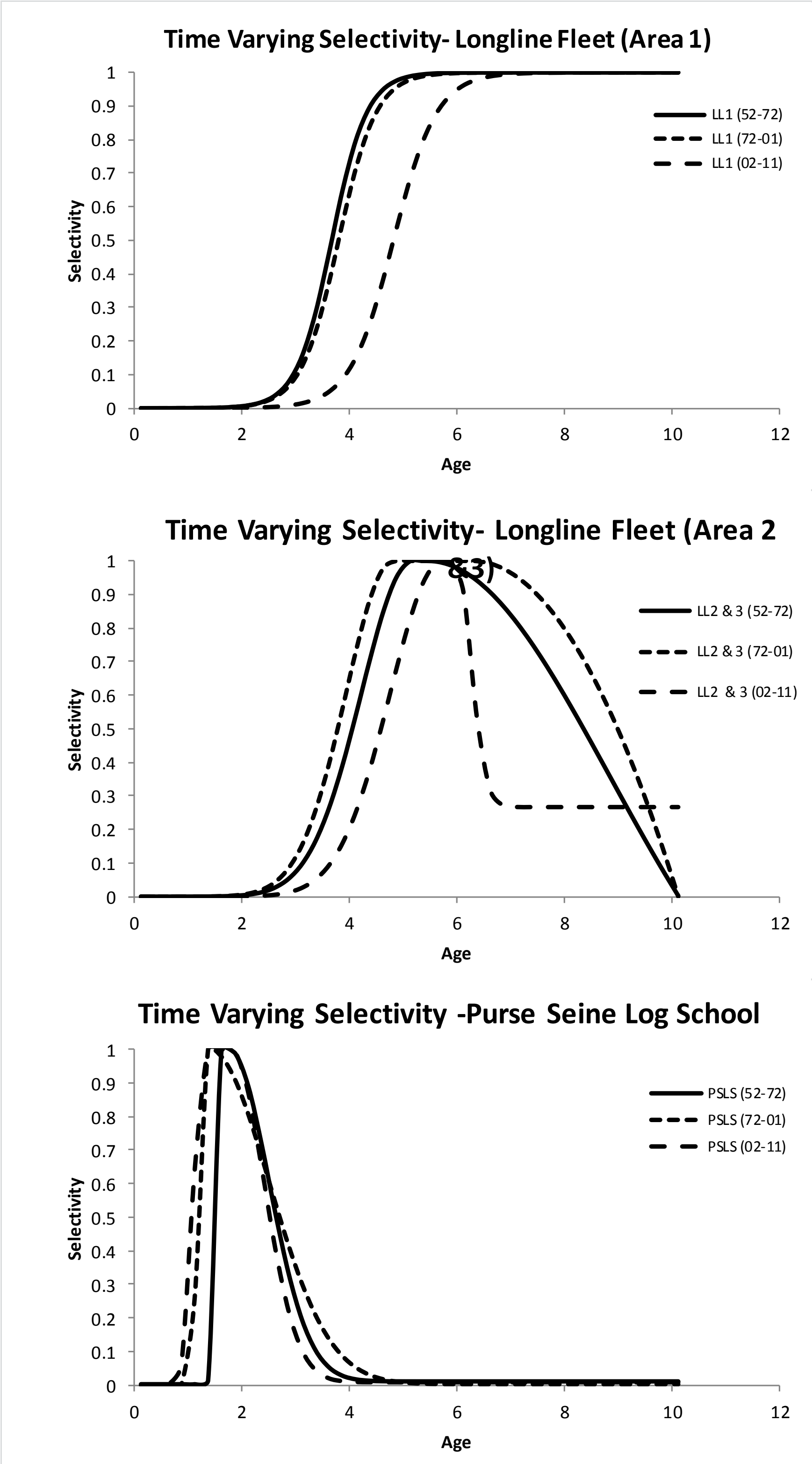


Figure 10: Time Varying selectivity using higher Effective Samples sizes (200), between the Longline and Purse Seine fisheries using the time blocks 1952- 1971, 1972-2001, and 2002-2011 respectively.

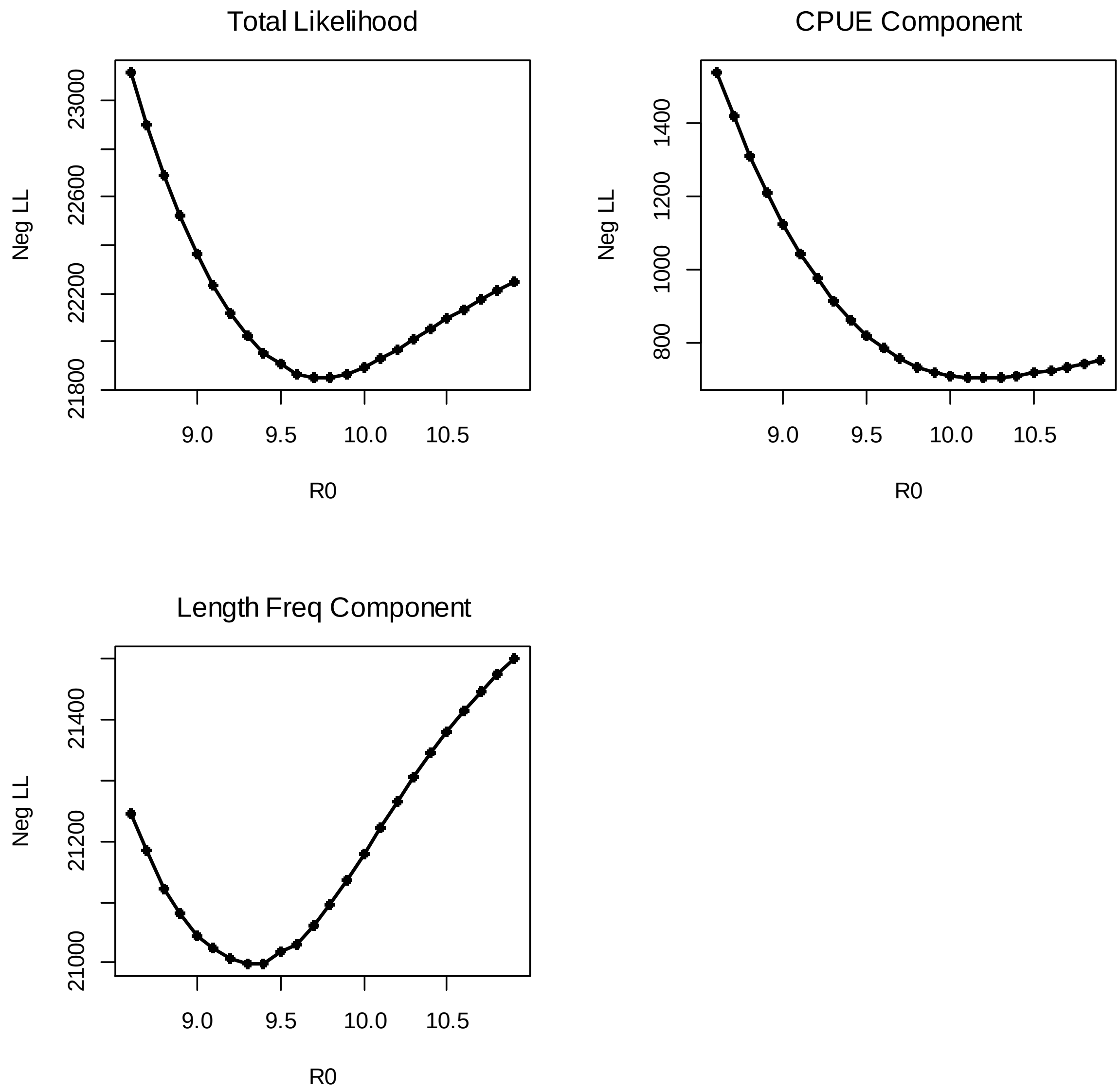


Figure 11: Likelihood profile for the R_0 parameter for the single region model. The likelihood profiles for the CPUE and length frequency components of the total likelihood are also presented.

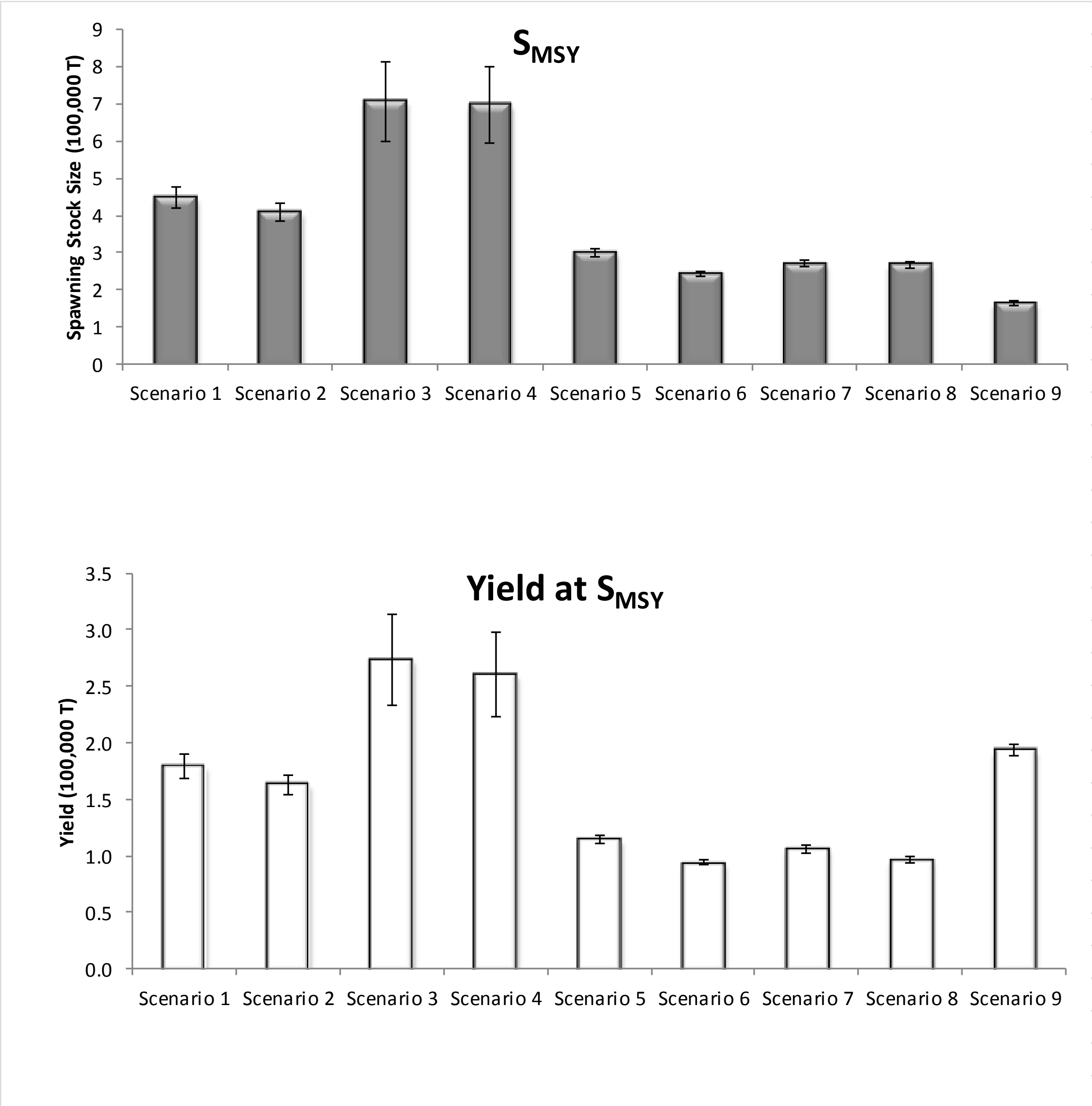


Figure 12. Target reference points as estimated by the different models summarized in Table 2 and Table 3.

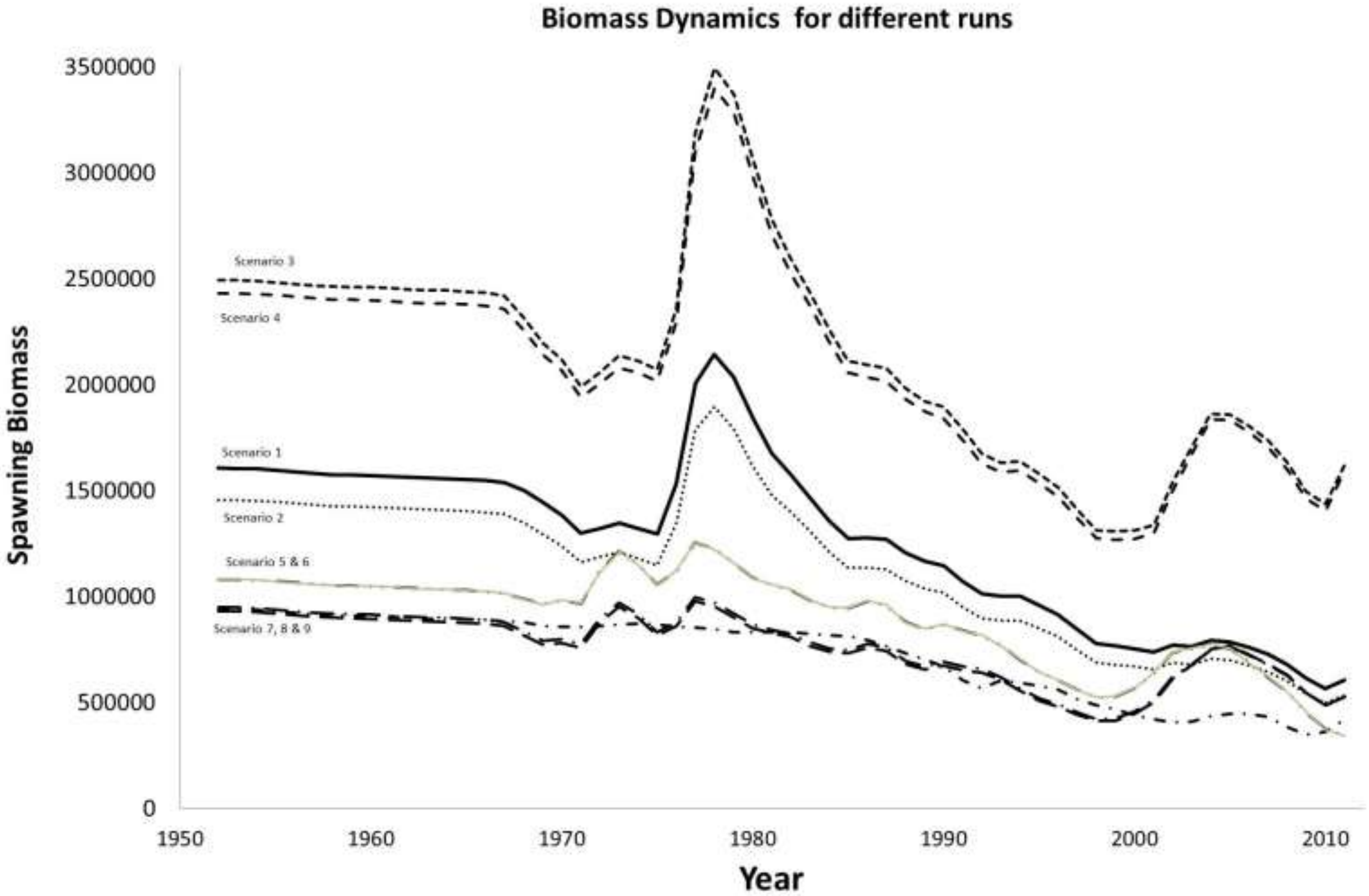
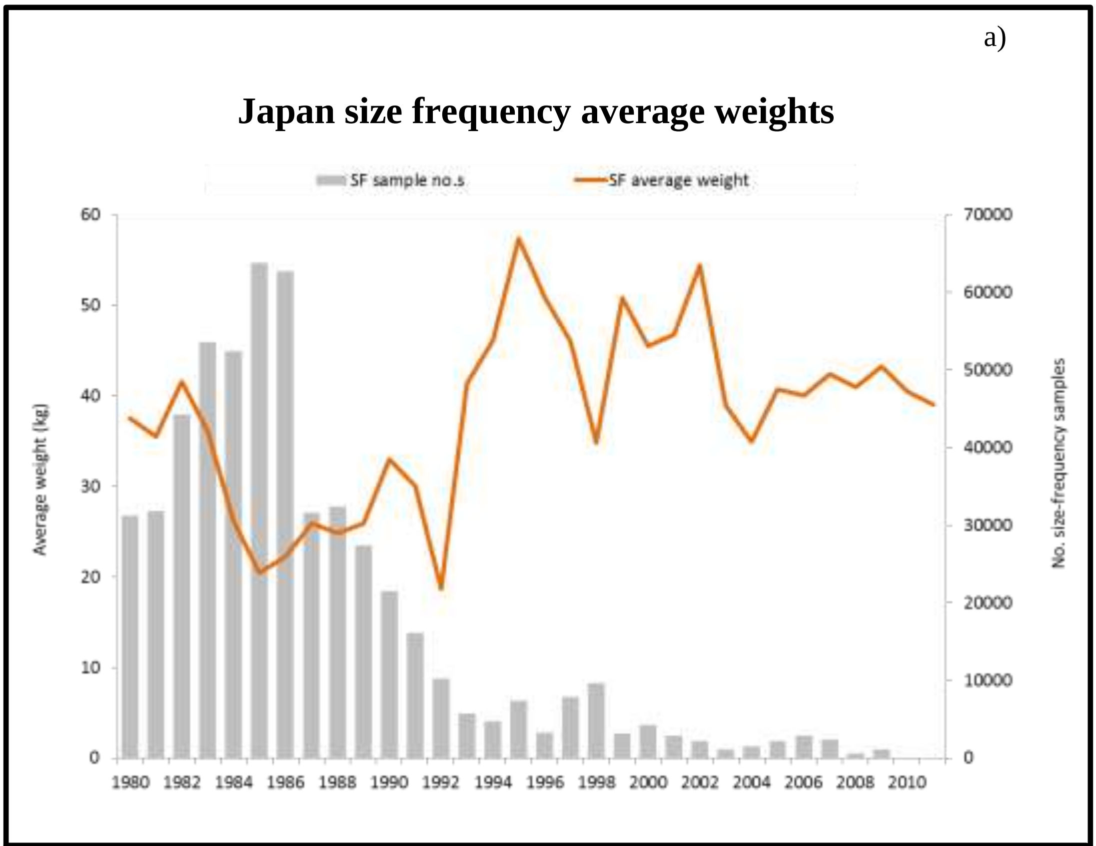


Figure 13. The different Biomass trajectories estimated from the single region models using time-varying selectivity with contrasting weighting of the length frequency data (different effective sample sizes).



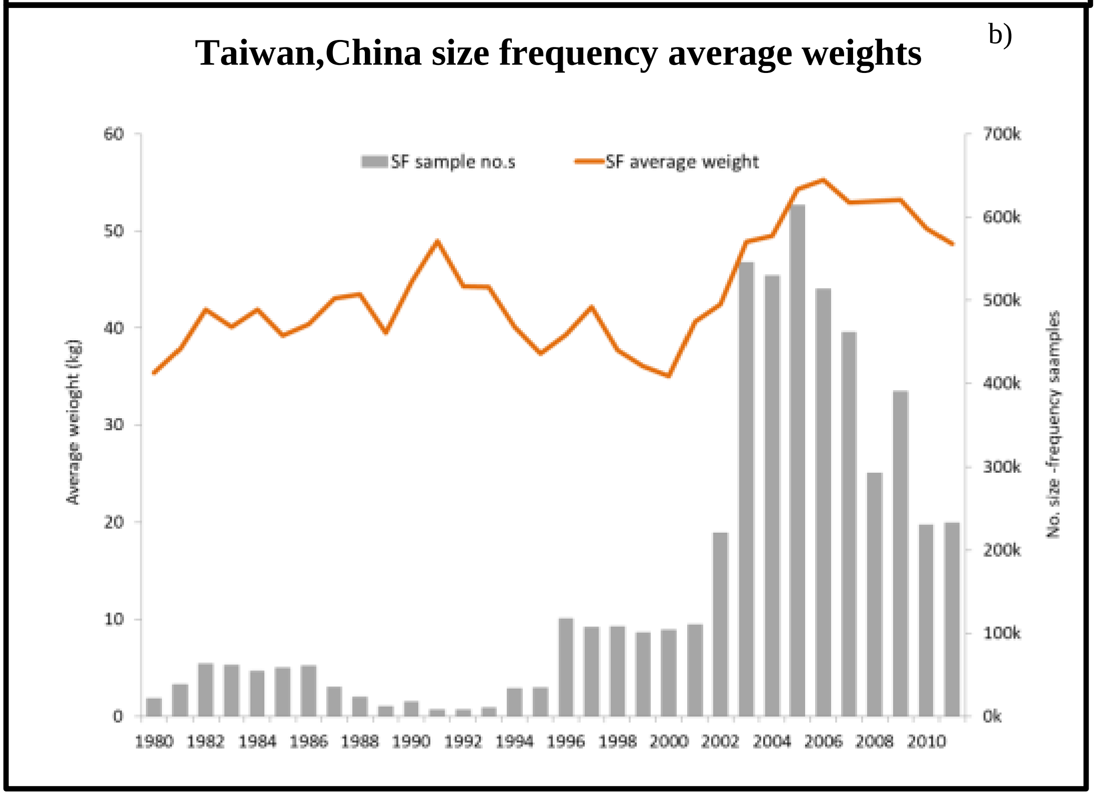


Figure 14. Length frequency data trends in average catch and sample sizes from Japan (fig. 14a) and Taiwan, China (fig. 14b) longline fleets over the time period used in the assessment.

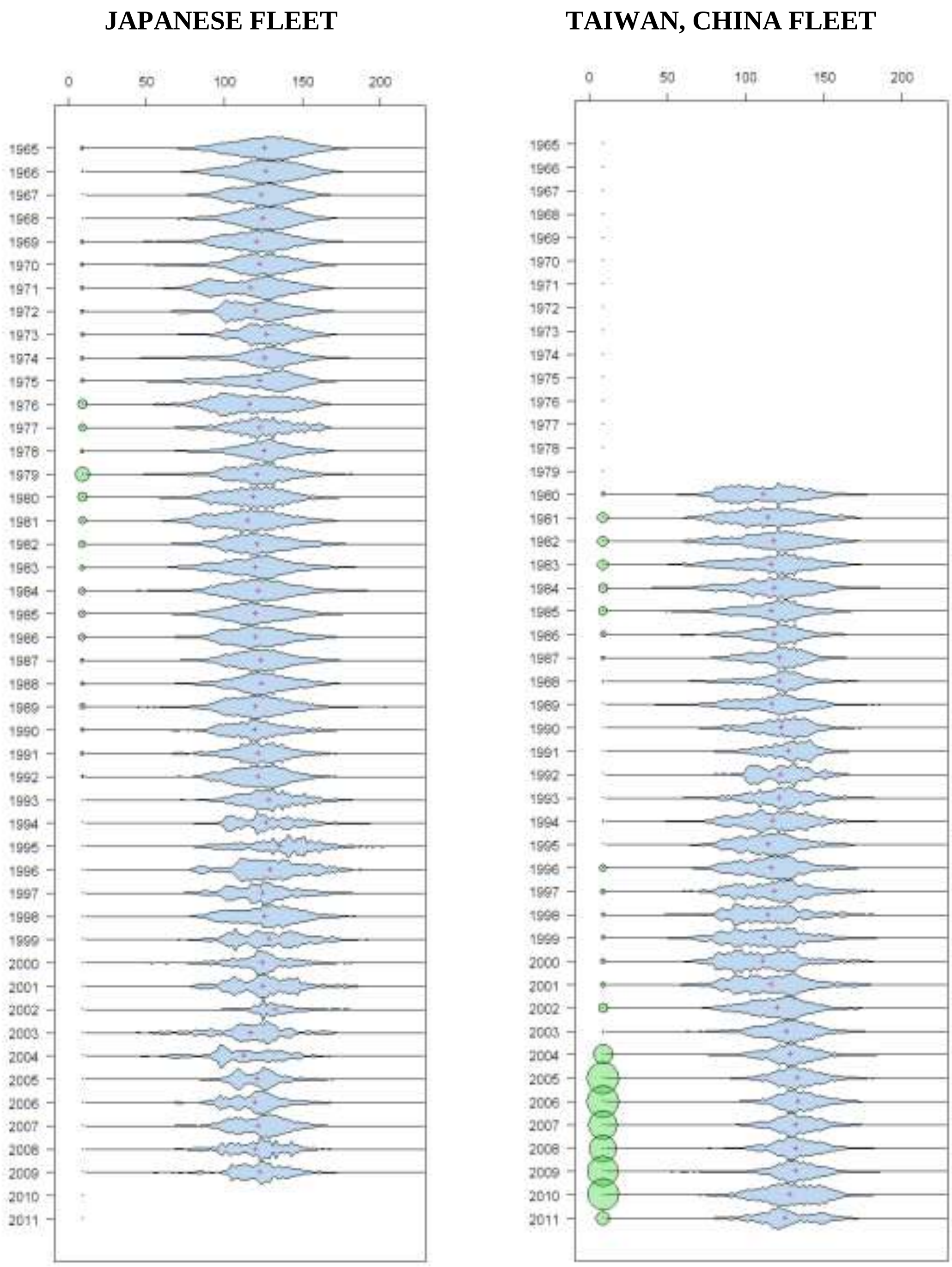


Figure 15: Comparison of BET length frequency distributions of Tawianese and Japanese Longline fleets. The size of the green proportional circles indicates the sampling coverage in each year, which are scaled relative to the minimum sampling coverage of one fish per tonne of catch recommended by the Indian Ocean Tuna Commission (denoted by the black circle at the centre of each proportional circle). Years with proportional circles larger than the minimum sampling standard indicate relatively high sampling coverage in a given year; the larger the circle the higher the sampling coverage. The crosshairs indicate the average length of sampled cohorts in each year class.