IOTC-2016-WPTT18-20
Received: 10 October 2016

Stock assessment of bigeye tuna in the Indian Ocean for 2016 — model development and evaluation.

Adam Langley, IOTC Consultant

1 Introduction

The distant-water longline fishery commenced operation in the Indian Ocean during the early 1950s. Bigeye tuna represented a significant component of the total catch from the longline fishery and catches increased steadily over the subsequent decades, reaching a peak in the late 1990s–early 2000s (IOTC 2015). The purse-seine fisheries and fresh-chilled longline fisheries developed from the mid-1980s and total bigeye tuna catches peaked at about 150,000 mt in the late 1990s. During the mid-2000s, the total annual bigeye catch declined 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. The total annual catch declined to 85,111 mt in 2010 but recovered somewhat over the following years, reaching 119,554 mt in 2012. The total annual catch in 2014 was 100,231 (IOTC 2015).

Previous assessments of the Indian Ocean bigeye tuna stock have been conducted using Stock Synthesis (Shono *et al* 2009, Kolody *et al* 2010, Langley *et al* 2013a) and ASPM (Nishida & Rademeyer 2011) software. Kolody *et al* (2010) conducted a thorough examination of the sensitivity of the assessment to a range of the key model assumptions. The resulting model scenarios highlighted the high level of uncertainty associated with the stock assessment, encompassing a range of contrasting estimates of stock status.

Langley *et al* (2013a) updated the IO bigeye assessment, extending the model time period to 2011. The modelling investigated a wide range of model sensitivities particularly regarding the spatial configuration of the model, the treatment of the available tagging data and the relative influence of the individual length frequency data sets. The assessment was subsequently updated during WPTT15, extending the catch, CPUE indices and size frequency data to the end of the 2012 year (Langley *et al* 2013b). Overall, the results of the 2013 assessment were consistent with the earlier assessment of Kolody *et al* (2010).

The SS model results provided the basis for the management advice for bigeye tuna formulated by WPTT15 (IOTC 2013, Langley *et al* 2013b). Uncertainty in the assessment was characterised using a grid of model runs that incorporated the key model assumptions (natural mortality, SRR steepness and longline catchability). For most model options, recent fishing mortality rates were below the F_{MSY} reference level ($F_{current}/F_{MSY} < 1$) and that the stock was not in an overfished state ($SB_{current}/SB_{MSY} > 1$). The median value of *MSY* from the model runs was 132,000 t (range 98,000–207,000 t) (IOTC 2013).

The final assessment models adopted by WPTT15 did not incorporate the available IO bigeye tag release/recovery data. Preliminary model options that incorporated these data were reviewed by WPTT15 (Langley *et al* 2013a). However, the spatial dynamics of these models were not considered to adequately represent the dynamics of the IO bigeye tuna stock, specifically the regional distribution of biomass and the movement dynamics. The model results were sensitive to the spatial structure of the model (1 or 3 regions), the tag mixing period (4 quarters vs 8 quarters) and the relative weighting of the length frequency data. The WPTT15 recommended that further model development be undertaken to utilise the tagging data in future IO bigeye stock assessments.

This report documents the next iteration of the stock assessment of the Indian Ocean bigeye tuna stock for consideration at 18th WPTT meeting. The terms of reference for the bigeye tuna stock assessment included a review of the spatial stratification and parameterization of the assessment model to enable the integration of the tagging data into the assessment model. The IOTC Secretariat also requested that a range of model sensitivities be conducted to investigate model assumptions, specifically related to natural mortality, selectivity and SRR steepness. This report documents the preliminary results of the assessment for presentation to WPTT18.

2 Data compilation

The data sets included in the stock assessment were fishery-specific catch and length frequency data, standardised longline CPUE indices and tag release/recovery data from the Indian Ocean RTTP. The stock assessment was implemented in Stock Synthesis (version V3.24z) and data were compiled in accordance with the SS data structures (Methot 2015).

2.1 Spatial structure

Most of the bigeye tuna catch from the Indian Ocean is caught within the latitudinal range 35° S to 10° N (Figure 1). This area of the Indian Ocean was used as the spatial domain of the assessment model (Figure 2). The small amount of bigeye catch from higher latitudes was reassigned to fisheries of the appropriate method within the model domain.

Preliminary modelling adopted a regional structure that was equivalent to the spatial stratification of the 2013 bigeye assessment (Langley *et al* 2013a). The spatial domain of the model was stratified into three regions: western equatorial region (region 1), eastern equatorial region (region 2) and southern region (region 3). The regions were defined to partition the spatial distributions of the main fisheries — most of the longline catch is taken within the two equatorial regions (15° S to 10° N), while the purse seine catch is predominantly taken within the western equatorial region. A seasonal longline fishery operates in the southern region. The longitudinal partitioning of the equatorial area subdivides the distribution of tagged fish recoveries from releases that occurred in the western area of region 1. There are also some differences in the temporal trends in the longline CPUE indices from the three regions.

The final model options further subdivided the western equatorial region (region 1) into two regions to account for differences in the distribution of tags within this region (see Section 2.5). The region was partitioned at the equator; the area south of the equator and the north and south regions were denoted Region 1N and Region 1S, respectively.

2.2 Catch data

Preliminary model options included a total of 12 fisheries defined based on the fishing gear type and the regional stratification. In the final model options, the Region 1 fisheries were partitioned based on the revised spatial structure (north and south of the equator) resulting in three additional fisheries. The fishery definitions are presented in Table 1. Bigeye tuna catches (in mt) were compiled by fishery and year/quarter.

Longline, distant-water (LL 1N, LL1S, LL2, LL3). The longline fishery operates throughout the Indian Ocean although catches are concentrated in the equatorial region (Figure 1). Catches are primarily from the Japanese, Korean and Taiwanese distant-water longline fleets. Most (62%) of the distant-water longline catch has been taken from the western equatorial region and annual catches from the LL1 fishery (LL1S and LL 1N) steadily increased from the early 1950s to reach a peak of 64,000 mt in 2003–2004. Catches of about 55,000 mt were maintained during 2005–07, declined rapidly to about 15,000 mt in 2010–2011, recovered to about 50,000 mt in 2012 and then declined to about 22,000 mt in 2014–2015 (Figure 3).

Annual catches from the LL2 fishery fluctuated between about 10,000–15,000 mt during 1975–2011. In the subsequent years, annual catches declined steadily to about 4,700 mt in 2015.

Annual longline catches from the southern area were comparatively low, averaging about 3,000 mt, from 1960 to 1990 (Figure 3). Catches then increased to a peak of 22,000 mt in 1995, declined steadily to about 5,000 mt in 2007 and remained at that level over subsequent years.

Longline, fresh tuna fleet (FL2). The fishery developed in the late 1980s and annual catches rapidly increased to reach a peak of about 30–35,000 mt in the late 1990s–early 2000s. Catches declined sharply in 2003 and then again in 2010. During 2011–2015, annual catches were about 12,000–15,000 mt per annum (Figure 3). The length compositions of the fish sampled from this fishery differ somewhat from the distant-water longline fisheries (LL1-3).

Purse seine (PSFS1N, PSFS1S, PSFS2, PSLS1N, PSLS1S, PSLS2). Almost all of the industrial purse-seine catch is taken within the western equatorial region (Figure 1) and catches are dominated by the fishery on associated schools (PSLS1N and PSLS1S) (Figure 3). Annual catches from the PSLS1 (N and S) fisheries reached a peak of about 30,000 mt in the late 1990s and have fluctuated at 15–25,000 mt over the last decade, with the exception of a lower catch in 2012. Since the late 1990s, annual catches from the purse-seine free-school fishery (PSFS1N and PSFS1S) have fluctuated between 5,000–10,000 mt.

Relatively minor catches were taken intermittently by the purse-seine fisheries in the eastern equatorial region (PSFS2 and PSLS2) (Figure 3).

Baitboat (BB1). Bigeye catches from the Maldives baitboat fishery are estimated to have increased steadily from a minimal level in the late 1970s to about 6–7,000 mt in recent years (Figure 3). This fishery was assigned to the northern equatorial region (Region 1N).

Line (LINE2). The LINE2 fishery includes small scale fisheries using handlines, small longlines and the gillnet/longline combination fishery of Sri Lanka. Annual catches are estimated to have increased steadily from a minimal level in the 1970s to about 8,000 mt in recent years (Figure 3).

Other (OT1 and OT2). The “Other” fisheries include gillnet, trolling and other minor artisanal gears. The fisheries are aggregated by region for the two equatorial areas. Within the western region the OT1 fishery is primarily comprised of the Iranian driftnet fishery operating in the high seas. Total catches were negligible prior to 2005 but subsequently increased to about 2,500 mt per annum (Figure 3). This fishery was assigned to the northern equatorial region (Region 1N).

For the Other 2 (OT2) fishery, recent catches were primarily taken by the Indonesian troll and gillnet fleets. Annual catches increased steadily from the early 1990s to reach a peak of about 5,000 t in 2011–2013 (Figure 3).

2.3 Length frequency data

Bigeye length frequency data were compiled by fishery and year/quarter.

Longline, distant-water (LL 1–3). Size frequency data are available for the LL1–3 fisheries from 1965 to 2015. Length and weight data were collected from sampling aboard Japanese commercial, research and training vessels. Weight frequency data collected from the fleet have been converted to length frequency data via a processed weight-whole weight conversion factor and a weight-length key. Length frequency data from the Taiwanese longline fleet are also available from 1980–2015. In recent years, length frequency data were also collected from locally based longline fleets (e.g. Seychelles).

Length frequency data from all sources were aggregated to provide a composite length composition for each year/quarter. Prior to 1995, the length compositions were dominated by sampling from the Japanese longline fleet, while in the subsequent period the size data were increasingly dominated by data collected from the Taiwanese distant-water longline fleet.

Previous assessments of IO bigeye tuna have highlighted the temporal variability in the length composition data from the main longline fisheries. For the current assessment, the longline length data were examined to investigate potential sources of variation in fish length. For each model region, the length data were restricted to the main area of catch from the bigeye tuna longline fishery (core area). It was considered that restricting the sampling data to the core area would minimise potential variation in length composition attributable to the collection of length samples from the periphery of the fishery.

The core areas for each region were defined as follows.

LL1	Latitude range 10°S to 10° N, longitude 50° E to 70° E
LL2	Latitude range 15°S to 5° S, longitude 85° E to 110° E

LL3 Latitude range 35–25° S

Temporal trends in average fish length (by quarter) were compared amongst the main longline fleets in each region (Appendix 1). The main effect of applying the core area criteria was to exclude a considerable proportion of the Japanese length samples collected from the LL fisheries. This is because the Japanese longline length frequency data are aggregated at a coarse spatial scale (longitude 20°, latitude 10° cells) that does not correspond well with the core area definitions. Applying the core area criteria to the Taiwanese length composition data did not fundamentally change the temporal patterns in the fish size.

For the LL1 area, there were marked differences in the sizes of fish sampled from the various longline fleets during the late 1980s and early 2000s. There were also divergent temporal trends in the lengths of fish sampled amongst the fleets (Figure A1 Appendix 1).

The lengths of fish sampled from the Taiwanese fleet increased markedly during the early 2000s and the length compositions of the samples from catches of most fleets were comprised of larger fish during 2005–2015 (Figure A1 Appendix 1). The increase in Taiwanese fish sizes during the period coincided with a large shift in the ratio of the Taiwanese(?) bigeye and yellowfin longline catches in the region during the same period; the ratio of bigeye in the longline catch increased during the late 2000s and remained at a higher level in the subsequent years (Hoyle et al 2015, see Figure 20). The lengths of fish sampled from the Taiwanese fleet during 2007–2015 were comparable to the lengths of fish sampled from other longline fleets operating in the region (Figure A1 Appendix 1).

A similar trend is also apparent in the length composition data from the eastern equatorial region (Figure A3 Appendix 1). The lengths of bigeye tuna sampled from the Taiwanese longline fleet increased considerably during the early 2000s, followed 3 years later by an increase in the catch of bigeye tuna relative to yellowfin tuna from the Taiwanese longline fleet (Hoyle et al 2015, see Figure 21).

Nonetheless, a recent review of the recent Taiwanese length composition data by Geehan & Hoyle (2013) concluded that "*changes from around 2003 in the size frequencies of BET, YTF, and ALB samples recorded by Taiwan,China are likely to result from a change in the sampling protocols or processing and management of the data. Alternative explanations (changes in the population structure or changes in fishery selectivity) are unlikely given the nature of the changes and the fact that similar changes are observed for all three species*". Further, Geehan & Hoyle (2013) recommended "*excluding from stock assessments the size data for BET, YFT and ALB from the Taiwanese DWLL fleet after the early-2000s, until the cause of changes in the size frequency data have been determined by the WPTT*".

Based on that recommendation, the recent length frequency data from the Taiwanese longline fleet were excluded from the final length frequency data sets incorporated in the current assessment. The period of data exclusion was also extended to 1997–2015, encompassing the period of considerable change in the length composition from the region 1 longline fishery.

For the final data sets, the length compositions of the LL1–3 fisheries are dominated by fish in the 90–150 cm length range (Figure 6). The aggregated length compositions are comparable for the three regions, although the average lengths of fish in the sampled catch fluctuated over the study period (Figure 7). For LL1, average fish length declined during the early 1990s. Limited samples were available from LL1 during 1997–2005 but fish sampled in the subsequent years (2006–2015) were considerably larger than sampled during the preceding period (Figure 7). These samples were predominantly obtained from the locally based longline fleets. Relatively few samples were available from LL2 and LL3 during the latter period.

Longline, fresh tuna fleet (FL2). Length data are available from 1998−2008; however, no length data are available from the earlier period of the fishery (1973−1997) (Figure 5). Length and weight data were collected during the unloading of catches at several ports, primarily from fresh-tuna longline vessels flagged in Indonesia and Taiwan/China (IOTC-OFCF sampling).

The composite length composition of the catch is similar to the distant water longline fleet (Figure 6) and remained relatively stable over the sampling period (Figure 7).

Purse seine (PSFS1, PSFS2, PSLS1, PSLS2). Length-frequency samples from purse seiners have been collected from a variety of port sampling programmes since the mid-1980s. The samples are comprised of large numbers of individual fish measurements and represent comprehensive sampling of the main period of the fishery (Figure 5). Limited size data are available from the purse-seine fisheries within region 2.

The associated purse-seine fishery (PSFS) primarily catches smaller bigeye tuna, while the size composition of the catch from the free-school fishery is bimodal, being comprised of the smaller size range of bigeye and a broad mode of larger fish (Figure 6). There was a general decline in the average length of fish caught by the PSLS1 fishery from 1990 to 2015 (Figure 7). The average size of fish sampled from the free-school fishery was variable among quarters, although fish tended to be smaller during the late 2000s, increased during the early–mid 2000s and declined in more recent years. It is unknown whether the trends in the length composition of the purse seine catch are representative of the population or reflect changes in the operation of the fishery.

Baitboat (BB1). Limited length samples are available from the fishery (Figure 5) and the sampled catch was dominated by fish in the smaller length classes (50–70 cm) (Figure 6 and Figure 7).

Line (LINE2). Negligible length frequency data are available from the fishery although the available data indicate that the catch was predominantly composed of larger fish (Figure 6). Fish sampled from this fishery are considerably larger than the fish sampled from the other main longline fisheries.

Other (OT1 and OT2). While catches from the OT1 fishery are dominated by the Iranian driftnet fleet, there are no length samples available from this component of the fishery. Instead, the available OT1 length samples were collected from the 'other' fisheries that operated prior to 2005 (Figure 5). The aggregate length samples encompass a broad length range (Figure 6).

For the Other 2 (OT2) fishery, limited length samples were collected from the Indonesian small purse seine and troll fisheries (Figure 5). The aggregate length frequency data available include two size modes from the small scale purse seine samples (Figure 6). This is probably due to different sizes of fish taken by different modes of fishing (e.g. fishing at night with light, around anchored FADs, etc.).

For each of the defined fisheries, the length-frequency data were aggregated by year/quarter into 95 2-cm size classes (10–12 cm to 198–200 cm). Each length frequency observation from the purse seine fishery represents the number of fish sampled raised to the sampling units (sets in the fish compartment). For the non-purse seine fisheries, each observation represents the actual number of fish measured.

Each aggregated length sample was assigned an initial sample size. The sample size was determined based on the number of fish included in the aggregated sample, up to a maximum of 1000. The sample size was then divided by 100 resulting in a maximum initial sample size of 10. Purse seine length samples were also assigned an initial sample size of 10.

2.4 CPUE indices

Standardised CPUE indices were derived using generalized linear models (GLM) from operational longline catch and effort data provided by Japan, Korea, and Taiwan,China (Hoyle *et al* 2016). Cluster analyses of species composition data by vessel-month for each fleet were used to separate datasets into fisheries understood to target different species. Selected clusters were then combined and standardized using generalized linear models. Bigeye catch (numbers of fish) was the dependent variable of the positive catch model (lognormal error structure), while the presence/absence of bigeye tuna in the catch was the dependent variable in the binomial model. In addition to the year-quarter, models included covariates for vessel identity, 5° square location, number of hooks, and either cluster (for region 3) or HBF (for regions 1 and 2).

For the 2016 assessment, three sets of CPUE indices were derived based on different treatment of the fishing vessel variable in the CPUE modelling (Hoyle *et al* 2016). The assessment modelling incorporated the *boat_allyears* set of CPUE indices on the basis that the indices represented the longest time series (1953–2015) and incorporated vessel effects for the period when individual vessel identifiers were available (1979–2015).

The following Table specifies the individual sets of CPUE indices used for each model region.

Region	Model variables	Indices series name
1	No cluster, HBF	*Joint_regB2_R1_dellog_boat_allyrs*
2	No cluster, HBF	*Joint_regB2_R2_dellog_boat_allyrs*
3	Cluster, no HBF	*Joint_regB2_R3_dellog_boat_allyrs*

The long term trends in the regional CPUE indices are broadly comparable to the Japanese longline CPUE indices used in the previous stock assessments (Matsumoto *et al* 2013). A notable feature of both sets of CPUE indices from the equatorial regions (Regions 1 and 2) is the large increase in the indices in the late 1970s; the indices increase by 2–3 fold during 1976–1978 and then decline sharply between 1978 and 1979 stabilising at a level considerably higher than the 1974–1976 level (Figure 8).

These rapid and widespread changes in the CPUE indices are not compatible with the population dynamics of bigeye tuna. Previous stock assessments incorporated the longer term CPUE indices (Kolody et al 2010, Langley et al 2013) and the assessment models estimated strong temporal trends in the recruitment deviates in an attempt to fit the large changes in the abundance indices during the late 1970s.

The analysis of the operational level data from the Japanese longline fleet has not been able to account for the large change in CPUE during the late 1970s. The transition in CPUE during this period could be explained by a marked change in the reporting of the bigeye tuna catch by the Japanese fleet, although no evidence is available to support this assertion. There was an exceptionally high bigeye tuna catch reported by the fleet in 1978 (especially the first quarter).

The marked change in the magnitude of the CPUE indices between the two periods (pre 1976 and post 1979) corresponds to a period when there was a large shift in the gear configuration of the Japanese longline gear in the equatorial regions (Hoyle & Okamoto 2015). Prior to 1976, most of the sets were conducted using 5–6 HBF, while from 1980 most of the sets used 10–12 HBF (see Figure 7 of Hoyle & Okamoto 2015). There is very limited overlap in the HBF between the two time blocks and, consequently, the CPUE standardisation procedure may not be able to adequately account for the change in efficiency related to the introduction of the deeper setting longline gear (i.e., HBF and year effects may be confounded).

On that basis, the CPUE indices for the period prior to 1979 were excluded from the stock assessment data set for the main model options. There is scope to revisit this decision at a later time and reinstate the earlier (pre 1976) indices. However, in that case it is recommended that the indices from the earlier period be treated as a separate set of abundance indices.

The truncated sets of CPUE indices from 1979–2015 were normalised to the average of each series. The CPUE indices from region 3 exhibit considerable seasonal variation, with lower CPUE in the first quarter (Jan–Mar) and relatively high CPUE in the third quarter (Jul–Sep) (Figure 10). Stock Synthesis does not have the flexibility to estimate seasonal catchability or movement dynamics when the model is configured based on a quarterly time step. Consequently, to account for the seasonal variation in the CPUE indices, the Region 3 CPUE indices were incorporated in the model as four separate sets of abundance indices (i.e. one series for each quarter).

The CPUE indices from the three regions exhibit broadly comparable trends, declining by about 65–70% from the early 1980s to 2010s (Figure 10). The CPUE indices fluctuated over the time period with higher CPUE in each region during the late 1980s and mid-1990s. From the late 2000s, the CPUE from the three regions deviated; the CPUE indices from region 2 declined during this period and then fluctuated about the lower level, while the CPUE indices from region 1 were maintained at about the same level during the late 2000s, were considerably higher in 2011–2012 and then declined during 2013–2014 (Figure 10).

The deviations in the trends between the region 1 and 2 CPUE indices were further investigated by comparing the ratio of the two indices with the Indian Ocean Dipole Index (IODI) (Figure 11). There is some indication that the differential in CPUE in each region is correlated with the IODI; the strong positive IODI during 2006–2012 is correlated with the higher CPUE in region 1 (relative to region 2), while the sharp decline in CPUE in region 1 during 2013–2014 corresponded to generally negative IODI (Figure 11).

The CPUE standardization procedure incorporates a number of factors, including changes in gear configuration, that may account for changes in the performance of the longline fishery over time. In addition, the operational level CPUE indices also incorporates individual vessel effects to account for changes in the efficiency of the fleet over time. These individual vessel effects were not included in the standardised CPUE indices incorporated in the previous stock assessments of IO bigeye tuna (Matsumoto *et al* 2013). The previous assessment attempted to account for gross increases in the efficiency of the fishing fleet by applying an increase in catchability to the time series of CPUE indices at a rate of 1% per annum. The resulting CPUE indices were included in the range of model sensitivity options.

The estimate of the increase in longline catchability (1% per annum) included in the previous assessment was derived by comparing CPUE indices that incorporated individual vessel effects with CPUE indices without the vessel effects included. The CPUE indices incorporated in the current assessment included vessel effects in the standardization procedure. Therefore, it is considered that the increase in fishing efficiency associated with the modernisation of the fleet has been accounted for in the CPUE indices. However, there are likely to be a range of other factors that have improved the fishing efficiency for bigeye, particularly as the fleet has shifted to target the species. These factors are likely to be partially accounted for in the CPUE standardization process (e.g. change in distribution of fishing effort and gear configuration) although it is considered likely that the CPUE indices under-estimate the magnitude of the change in fishing efficiency. There is no information available to indicate the magnitude of the additional (unaccounted) increase in the catchability of bigeye by the longline fleet and, hence, the assessment modelling has not included model options that account for an increase in longline catchability.

Separate longline CPUE indices were not available for the two sub-regions of the western equatorial region (Regions 1N and 1S). Consequently, for the four region model it was necessary to incorporate the LL1 CPUE indices in both sub-regions (Regions 1N and 1S). In future, separate CPUE indices should be derived for Regions 1N and 1S.

For the final model options, a common catchability coefficient was estimated in the assessment model, thereby, linking the respective CPUE indices among regions. This significantly increases the power of the model to estimate the relative (and absolute) level of biomass among regions. However, as CPUE indices are essentially density estimates it is necessary to scale the CPUE indices to account for the relative abundance of the stock among regions. For example, a relatively small region with a very high average catch rate may have a lower level of total biomass than a large region with a moderate level of CPUE.

The approach used was to determine regional scaling factors that incorporated both the size of the region and the relative catch rate to determine the relative level of exploitable longline biomass among regions. The scaling factors were derived for the original three regions from the composite longline CPUE data from 1981–2000. The relative scaling factors thus calculated for regions 1–3 were 1.00, 0.87, and 0.83, respectively (S. Hoyle *pers. comm.*). For the four region model, the Region 1 scaling factor was apportioned equally between the two sub-regions (Regions 1N and 1S) reflecting the relatively even distribution of longline catch between the two areas. The final regional scaling factors were: R1N 0.50, R1S 0.50, R2 0.87, R3 0.83.

2.5 Tag release/recovery data

A considerable amount of tagging data was available for inclusion in the assessment model. The data used consisted of bigeye tuna tag releases and returns from the Regional Tuna Tagging Project-Indian Ocean (RTTP-IO) phase of the Indian Ocean Tuna Tagging Programme (IOTTP). Tags were released during 2005–2007 and recoveries were monitored by the IOTTP during 2005–2009 and by the IOTC in the subsequent years.

All the bigeye tag releases of the RTTP-IO occurred in a localized area off the Tanzania coast within the western equatorial region (region 1S) (Figure 12). Most of the releases occurred during the second and third quarters of 2006 and the third quarter of 2007 (Figure 13).

The model data set included all tag recoveries to mid-2015. The spatial distribution of tag recoveries is presented in Figure 14. A relatively high proportion of tag recoveries occurred in the vicinity of the main release location. There was also a relatively large number of tags recovered from bigeye tuna catches from the Mozambique Channel (Figure 14). Most of the recaptures at these two locations were from purse seine associated sets during 2007 and were comprised of tagged fish at liberty for 6–12 months (Figure 15).

Tag recoveries from the purse seine fishery for fish at liberty for at least 12 months were more evenly distributed over the main area fished by the purse seine fleet (Figure 16 and Figure 17). However, further examination of the tag recoveries from the PSLS fishery identified considerable differences in the recovery rate (number of tags per tonne of catch) amongst latitudinal zones for tags at liberty for at least 12 months. (Figure 18). Tag recovery rates from the main area of the fishery south of 2°S were consistently higher than from the fishery in the area north of 2°S, especially during the main recovery period (2007–2008). During the main recovery period, PSLS catches were dominated by the component of the fishery in the area north of 2°S.

The difference in tag recovery rates between the two main areas of the fishery indicates that the dispersal of tagged bigeye during the 12 month “mixing period” was insufficient to redistribute tagged fish throughout the bigeye population resident within the western equatorial region (Region 1). Consequently, the distribution of PSLS fishery effort (and catch) would have strongly influenced the number of tags recovered from the fishery. In an attempt to account for the incomplete mixing of tagged fish, the spatial structure of the assessment model was refined by partitioning the western equatorial region into two regions (Region 1N and Region 1S).

For inclusion in the assessment model, tag releases were aggregated into release groups defined by release region (region 1S only), time period of release (quarter) and quarterly age class. The numbers of fish in each age at release were determined by applying an age-length key to the length composition of the tagged fish. The age-length key was derived by assuming an equilibrium population age-length structure based on the age-specific natural mortality, average length-at-age from the bigeye growth function and the standard deviation of length-at-age (CV 0.1) (see Section 3). Fish aged 15 quarters and older were aggregated in a single age group.

A total of 34,572 releases were classified into 86 tag release groups. Release groups with less than five (5) fish were excluded, representing the removal of a small number of tags. The number of tags in each release group was reduced by 30% to account for initial tag mortality; the initial tag mortality estimate of 20.5% (Hoyle *et al* 2015) was increased by a further 10% to account for an assumed level of tag mortality associated with the best (base) tagger (S. Hoyle *pers. comm.*). This reduced the number of tags in the final release data set to 24,103. Most of the tag releases were in the 4−8 quarter age classes (Figure 13).

The number of tags recovered from each tag release group were then classified by recapture fishery and recapture time period (quarter). Tag seeding experiments conducted during 2005–2008 revealed considerable temporal variability in tag reporting rates from the IO purse-seine fishery (Hillary *et al.* 2008). Reporting rates were lower in 2005 (57%) compared to 2006 (89%) and 2007 (94%). The improvement in reporting was attributed to the development of publicity campaign and tag recovery scheme raising the awareness of the stakeholders, *i.e.* stevedores and crew.

To account for the temporal changes in reporting rate, the number of tag returns from the purse-seine fishery in each stratum (tag group, year/quarter, and length class) were adjusted using the respective estimate of the annual reporting rate. A reporting rate of 94% was assumed for adjusting the 2008−2015 purse seine tag recoveries.

The number of tag recoveries from the purse seine fisheries was increased by an additional 10% to account for the proportion of the catch that was not landed in Seychelles and considered to have not been examined for tags. The numbers of tag recoveries were also adjusted for long-term tag loss (tag shedding) based on an analysis by Gaertner and Hallier (2015). Tag shedding rates for bigeye tuna were estimated to be approximately 1.7% per annum.

A total of 5,671 actual tag recoveries were included in the tagging data set. The cumulative effect of processing the tag recovery data increased the number of recoveries to 6,578 tags. A significant proportion of the tag returns from purse seiners were not accompanied by information concerning the set type. These tag recoveries were assigned to either the free-school or FAD fishery based on the assumed age of the fish at the time of recapture; i.e. based on the age assigned to the release group and the period at liberty. Fish "older" than 12 quarters were assumed to be recaptured by the free-school fishery; "younger" fish were assumed to be recovered by the FAD set fishery.

Most of the tag recoveries were from the purse seine FAD fishery within region 1, primarily in the first 12 months following release (Table 3). Overall, 74% of bigeye tag recoveries occurred within the first 12 months. Tagged fish at liberty for longer periods were predominantly recovered from the purse seine FAD fishery (region 1), purse seine school set fishery (region 1), and the longline fisheries in regions 1 and 3 (LL1 and LL3) (Table 3).

For the purse-seine fisheries, the tag dataset was adjusted for reporting rates (as described above) and hence the fishery reporting rates were fixed at a value of 1.0 in the assessment models (see Section 4.5). No information is available regarding the reporting rates of tagged fish caught the other (non purse-seine) fisheries.

3 Biological parameters

Recent estimates of Indian Ocean bigeye tuna growth derived from otolith age data and tag release/recovery 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 19). Maximum average length (L_∞) was estimated by Eveson *et al* (2012) at 150.9 cm (fork length). 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, pers. comm.). The growth function was implemented in SS using age-specific deviates on the k growth parameter.

Preliminary modelling investigated two alternative levels of age-specific natural mortality. The higher level of natural mortality is 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 (Figure 20). A lower level of natural mortality was proposed based on a Lorenz curve analysis (A. Fonteneau pers. comm.) with a lower natural mortality for the adult age classes (0.0625 per quarter) (Figure 20). This is comparable to the level of natural mortality assumed for Atlantic bigeye tuna in the recent ICCAT stock assessment by (ICCAT 2015).

From the RTTP, a considerable number of tagged fish were captured after 7–8 years at liberty, indicating a considerable proportion of the tagged fish had reached an age of 8–10 years; 8 tags were recovered after 10 years at liberty and a few tags were recovered during the most recent year (2015), corresponding to an age at recovery of 11–12 years. The higher level of natural mortality would result in a very small proportion of the tagged fish reaching 12 years of age, suggesting that the lower level of natural mortality may be more plausible.

The preliminary assessment modelling produced considerably lower tag likelihood values for the model options that included the lower level of natural mortality, indicating that the number of tag recoveries, primarily from the PSLS fishery, were more consistent with lower levels of M. Nonetheless, considerable caution should be placed on model based inferences of natural mortality due to potential biases in the tag recovery process, particularly related to the extent of tag mixing, and potential confounding between the estimation of natural mortality and other model parameters (especially selectivity, recruitment and movement).

For the final set of model options, the lower level of natural mortality was given priority based on the minimum estimates of longevity obtained from the maximum periods at liberty for tagged fish.

The size of sexual maturity was equivalent to that applied by Shono et al (2009) and used in the subsequent assessments. Female fish were assumed to attain sexual maturity from 100 cm (F.L.) with full sexual maturity at about 125 cm (Figure 21). In recent years, additional histological data have been collected from IO bigeye tuna and these data may enable a re-evaluation of the maturity ogive (Emmanuel Chassot *pers. comm.*)

The length-weight relationship was equivalent to that incorporated in previous assessments (Shono *et al* 2009, Kolody *et al* 2010, Langley *et al* 2013) and was originally derived by Nakamura and Uchiyama (1966). Fish weight = $a.\text{length}^b$, a = 3.661 x 10^{-5} , b = 2.901 where weight is in kilograms and length is in centimetres.

4 Population Model — structure and assumptions

4.1 Population structure and initial conditions

The model population structure is comprised of 41 quarterly age classes; the first age class represents fish aged 0–3 months (age 0) and the last age class accumulates all fish age 40+ quarters. The population is aggregated by sex and partitioned by region (3 or 4 spatial strata).

The model commences in 1975 and extends to the end of 2015 in quarterly intervals (164 time steps).

The initial (1975) age structure of the population was assumed to be in an exploited, equilibrium state. The four main LL fisheries were operating prior to 1975 and initial fishing mortality parameters were estimated for each of these fisheries. The resulting fishing mortality rates are applied to determine the initial numbers-at-age in each model region.

4.2 Recruitment

Recruitment of 0 age fish occurs in each quarterly time step of the model. Recruitment was estimated as deviates from the BH stock recruitment relationship (SRR) for 1985–2014 (120 deviates). The recruitment deviates were estimated for the period that corresponds to the operation of the PSLS fishery which provides catch and length data for the smaller fish and, hence, may be informative regarding the variation in recruitment. Recruitment deviates were not estimated for the last four quarters in the model as the relative abundance of these recruitments is not monitored by the model abundance index (longline CPUE). Recruitment deviates were assumed to have a standard deviation (σ_R) of 0.6. Spawning potential is defined as the biomass of mature female fish (based on an assumed sex ratio of 50:50).

Recruitment was assumed to occur in the three equatorial regions only. The overall proportional distribution of recruitment among the three regions was estimated. There is no information to indicate that there are significant differences in the pattern of recruitment between the regions; i.e. the CPUE trends are broadly comparable between the equatorial regions and the length composition data from the longline fisheries do not appear to be informative regarding recruitment. Length composition data from the small fish fisheries are available from the western equatorial regions only. Consequently, the relative distribution of recruitment between the three regions (1N, 1S and 2) was assumed to be temporally invariant.

The final model options included three (fixed) values of steepness of the BH SRR (h 0.7, 0.8 and 0.9). These values are considered to encompass the plausible range of steepness values for tuna species such as bigeye tuna and are routinely adopted in tuna assessments conducted by other tuna RFMOs.

4.3 Movement

In Stock Synthesis, movement is implemented as the proportional redistribution of fish amongst regions, including the proportion remaining in the home region. The redistribution of fish occurs instantaneously at the end of each model time step (quarter).

Movement of fish was estimated amongst the four model regions. Movement was parameterised to estimate differential movement from young (8 quarters) to old (≥15 quarters) fish to approximate potential changes in movement dynamics associated with maturation. For each movement transition, two separate movement parameters were estimated (for young and older fish). A linear interpolation between the age specific movement rates was applied to determine movement of the intermediate age classes.

Fish younger than age 3 quarters were assumed to remain within the natal region. Movement rates were assumed to be temporally invariant.

4.4 Fishery dynamics

For all fisheries, selectivity was assumed to be an age based process. A common selectivity function was estimated for the four main longline fisheries (LL1N, 1S, 2, 3) using a logistic function. A logistic selectivity was also assumed for the FL2 longline fishery.

The selectivities of the PSLS and BB fisheries were estimated using a double normal functional form. Separate selectivity functions were estimated for the PSLS1N and PSLS1S fisheries.

For the PSLS1N and PSLS1S fisheries, there was a marked shift in the length composition of the fishery catch in the mid-2000s. Preliminary modelling accounted for the apparent change in the selectivity of the PSLS1 fishery by including two time blocks (1975–2005 and 2006–2015) for the estimation of the selectivity parameters related to the age of the peak selectivity and the width of the ascending limb of the selectivity. The temporal shift in selectivity was not included in the final (four region) model options.

To account for the bimodal length composition of the catch from the PSFS fishery, the selectivity was modelled using a cubic spline with 6 nodes. The selectivity of the PSFS1N and PSFS1S fisheries was assumed to be equivalent.

Limited data were available to estimate the selectivity of either the PSLS2 or PSFS2 fisheries. The selectivity of these fisheries was constrained to be equivalent to the corresponding fishery selectivity in the western equatorial region 1S.

Limited size data are available from the “Other” fisheries. During the previous assessment, attempts to estimate independent selectivities for these fisheries were not successful, partly due to the variability in the length composition between 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 LINE2 fishery and the selectivity was assumed to be equivalent to the main longline fishery.

Fishing mortality was modelled using the hybrid method that estimates the harvest rate using Pope’s approximation and then converts it to an approximation of the corresponding F (Methot & Wetzel 2013).

4.5 Tag dynamics

In general, the population dynamics of the tagged and untagged populations are governed by the same model structures and parameters. The tagged populations (tag groups) are monitored over time intervals following release. The predicted number of tags in each region and subsequent time intervals are derived based on the movement parameters, natural mortality and fishing mortality.

For each time interval, the number of tags recovered by a specific fishery is predicted based the modelled number of tags in each age class in the region, the selectivity of the fishery and the fishing mortality of the specific fleet (fishery). The predicted number of tag recoveries is also moderated by the fishery specific reporting rate.

The observed number of tag recoveries for the purse seine fisheries were already adjusted to account for the differential tag reporting rates. On that basis, the reporting rates for the purse seine fisheries were fixed at 1.0. For the other fisheries, there is very limited information is available to indicate the tag reporting rates and fishery specific reporting rates were estimated based on uninformative priors. All fishery reporting rates were assumed to be temporally invariant.

The assessment framework assumes that the probability of capturing a tagged fish is equivalent to the probability of catching an untagged fish. Violation of the assumption of homogeneous mixing of tagged fish at the relevant spatial scale (i.e. region) is likely to introduce a bias in the estimation of fish abundance. In Stock Synthesis, a mixing period is specified which partitions the tag data sets (by release group); tag recoveries (observed and predicted) from the mixing period are excluded from the tag likelihood and therefore do not influence the estimation procedure.

For bigeye tuna, almost all tags were released from a localised area of region 1S. Most of the tag recoveries were from the PSLS1S fishery for which there were reliable estimates of tag reporting rates. The tagged bigeye were predominantly aged 4−8 quarters at release, while the selectivity of the PSFS1S is estimated to be 4−11 quarters.

Consequently, there is likely to be a limited time period (4−8 quarters) during which most of the tagged fish would be available to the PSLS fishery. Thus, a mixing period of four quarters was chosen on the basis that sufficient numbers of tagged fish remained available to the PSLS fishery during the post mixing period, albeit for a relatively limited period. The four quarter mixing period may be sufficient to allow for a reasonable degree of mixing of tagged fish within the south-western equatorial region (Region 1S). The dispersal of tags into the north-western equatorial region (Region 1N) will be mediated by the estimated movement rates (from Region 1S), however, the distribution of tagged fish in this region is unlikely to by homogeneous and it is likely that tag densities would be higher in the southern area of Region 1N (i.e. closer to the release location). Consequently, tag recoveries from the region may be influenced by the spatial distribution of the catch from the fishery.

The model also incorporates the tag recoveries from the other fisheries, most notably the PSFS and LL fisheries. There are no external estimates of tag reporting rates available for the longline fishery and, hence, the tag recoveries from the longline fishery will be considerably less informative about stock abundance.

4.6 Statistical framework

The total likelihood is composed of four main components: catch data, the abundance indices (CPUE), length frequency data and tag release/recovery data. There are also contributions to the total likelihood from the recruitment deviates and priors on the individual model parameters. The model was configured to fit the catch almost exactly so the catch component of the likelihood is very small. There are two components of the tag likelihood: the multinomial likelihood for the distribution of tag recoveries by fleets over time and the negative binomial distribution of expected total recaptures across all regions. Details of the formulation of the individual components of the likelihood are provided in Methot & Wetzel (2013).

The regional CPUE indices are assumed to represent the relative abundance (numbers of fish) of the proportion of the regional population that was vulnerable to the longline fishery (i.e. the selectivity of the LL1–3 fisheries). The weighting of the CPUE indices followed the approach of Francis (2011). An initial (three region) model was implemented that down weighted all the length composition and tag release/recovery data sets. The RMSE of the resulting fit to each set of CPUE indices was determined as a measure of the magnitude of the variation of each set of indices CPUE indices. The resulting RMSEs were relatively low (0.2−0.3) and, on that basis, a CV of 0.2 was assigned to each set of CPUE indices. The level of precision ensured the stock biomass trajectories were generally consistent with the CPUE indices while allowing for a moderate degree of variability in fitting to the indices.

For the four region model, it was considered that the duplication of the LL1 CPUE indices in the two sub-regions (1N and 1S) would over-emphasize the LL1 CPUE indices in the overall model. Consequently, in the final model options the CV for the LL1N and LL1S CPUE indices was increased from 0.20 to 0.25.

The weighting of the tag component of the likelihood was investigated following Francis (in press). The relative weighting of the tagging data was influenced by the magnitude of the over-dispersion parameters assigned to the individual tag release groups. The appropriate value of the over-dispersion parameter was determined using an iterative approach. From an initial model run, the residuals of the fit to the tag recovery data were determined (observed – expected number of tags recovered). The variance of the standardised residuals represents the variability in the tag-recapture data and the over-dispersion parameter should reflect this level of variability. For the initial model, the variance of the standardised residual was approximately 6.0. For comparative purposes, the over-dispersion parameters were estimated separately for each tag release group (n=86), assuming a relatively uninformative beta prior (mean 10, sd 3). The resulting over-dispersion parameters had an average value of 6.7 and a range of 2.8−9.5. The final model options estimated the over-dispersion parameters.

Preliminary modelling revealed that the model results were sensitive to the relative weighting of the length composition data. There is a large quantity of the length composition included in the assessment model; however, there are concerns about the reliability of the size frequency data for some of the key fisheries, notably the LL fisheries. For that reason, it was considered that the length composition data should not be allowed to dominate the model likelihood and, thereby, directly influence the trends in stock abundance.

For all fisheries, with the exception of the PSLS fisheries, the individual length frequency observations were assigned an effective sample size (ESS) of 1; i.e., down-weighting the initial sample size of 10 by an order of magnitude. For the PSLS fisheries an ESS of 10 was assigned to all length observations, The higher weighting of the purse seine PSLS length frequency data reflects the comprehensive nature of the port sampling programme monitoring the catch. There is a high degree of variation in the length composition data from the PSFS fisheries which appears related to the bimodal structure in the fishery length compositions; variation in the individual length samples may be attributable to sampling different proportions of the catch from each length mode. Based on the apparent level of sampling error an ESS of 1 was assigned to the length samples from the PSFS fisheries.

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

5 Preliminary modelling

A considerable amount of exploratory modelling was conducted during the development phases of the assessment (Appendix 2). These analyses investigated the sensitivity of the model results to a number of key structural uncertainties and the influence of the main data sets. The formulation of these models was based on the three region model structure. The main conclusions of the analyses are provided in the following paragraphs. Detailed results from specific model options are presented in Appendix 2.

1. ***Model spatial structure***. Three alternative spatial structures were considered: 1) a single region encompassing the entire Indian Ocean, 2) a two region model that amalgamates regions 1 and 3 and retains region 2 as a separate region, and 3) a spatial structure that subdivides the Indian Ocean into three regions. The three model options yielded very similar results, although the single region model estimates of spawning biomass were slightly lower (about 5%) (Figure A9). The level of total spawning biomass appears to be strongly informed by the tag release/recovery data from the late 2000s. For the spatially disaggregated models, a relatively high proportion (about 70%) of the recruitment and spawning biomass was estimated to be within the western equatorial region, with relatively low mixing rates between the two equatorial regions. The western equatorial region accounts for most of the catch of fish in the smaller length classes – those length classes that are vulnerable to the PSLS fishery which accounted for most of the tag recoveries and, hence, is most informative about stock size. For the spatially aggregated model, the tag release/recovery data inform the model about the abundance of the fish in those smaller length classes for the entire IO region (as complete mixing is assumed to occur after 4 quarters). Thus, because most of the small fish catch is taken from the western region, the single region model will yield an estimated level of biomass (in the late 2000s) that is similar to the spatially disaggregated models.
2. ***Initial conditions.*** The reference model commenced in 1975 and assumes equilibrium exploited conditions. An alternative model option commenced in 1952 and included the entire catch history from the history; initial conditions were assumed to be at the unexploited, equilibrium

level. The latter option estimated a decline in stock biomass during 1952–1975 that was very similar to the level of initial depletion estimated by the reference model (Figure A7).

3. ***Data conflict — CPUE indices vs Tagging data***. The influence of the tagging data in determining the scale of the overall biomass is evident when the tagging data were excluded from the total likelihood (Figure A10). This resulted in a substantially higher level (approx. double) of initial biomass and a much lower level of stock depletion. Estimates of *MSY* from the later model are also substantially higher. There is considerable conflict between the tag release/recovery data and the CPUE indices within region 1; the CPUE indices in the region were relatively high during 2007–2012 while the tagging data estimated a lower biomass level that was more consistent with the trend in the CPUE indices in region 2. Thus removing the tagging data from the model resulted in an improvement in the fit to the region 1 CPUE indices and, hence a higher level of overall biomass in the late 2000s–early 2010s and scaled the historical biomass level accordingly.
4. ***Data conflict — Tagging data vs Length frequency data.*** Model options that had a higher weighting to the tagging data sets resulted in a deterioration in the fit to the length composition data from the PSLS fisheries, especially in the more recent period (post 2005) when the size of fish sampled from the PSLS fishery was smaller. The higher weighting of the tagging data shifted the PSLS fishery selectivity towards older fish due to the numbers of older tagged fish that were recovered by the fishery. This may be a function of the age assignment of the tag release data set (to individual age based tag release groups).
5. ***Data conflict — CPUE indices vs Length frequency data.*** Down-weighting the longline CPUE indices (increasing the CV from 0.2 to 0.3) did not substantially change the scale of the spawning biomass during the period of the main tag release/recovery phase; however, higher levels of biomass were estimated for the preceding period and the more recent years (Figure A7). Reducing the weighting of the CPUE indices resulted in more pronounced patterns in the residuals from the region 1 (negative) and region 2 (positive) during the period prior to 2000. There was a corresponding improvement in the fit to the length composition data indicating that these data may be influencing the scale of the stock biomass during the earlier period. This was also evident when the Effective Sampling Size of the longline length frequency data was increased (from 1 to 10) (Figure A8) resulting in an increase in the scale of the overall biomass level. This difference was amplified when the ESS was further increased.
6. ***Regional recruitment dynamics.*** The reference model assumed the regional distribution of recruitment was temporally invariant. However, the resulting model revealed temporal patterns in residuals from the fit to the CPUE indices from region 1 and 2. An alternative model option estimated temporal deviates in the regional recruitment distribution between region 1 and 2. The resulting model estimated an overall recruitment distribution that was substantially different from the reference model with 60% of recruitment associated with region 2 (compared to 30% in the reference model). The resulting model estimated a level of biomass in region 1 that was comparable to the reference model, but estimated a much higher level of biomass in region 2 (350% higher) and region 2 was estimated to account for 60–70% of the total IO biomass. There was a correspondingly higher (about 50% higher) overall level of spawning biomass estimated for the regional recruitment model. The model resulted in an improvement in the fit to the region 2 CPUE indices but not appreciable improvement in the fit to the region 1 CPUE indices. These results reaffirm that the magnitude of the region 1 biomass is strongly informed by the tagging data. The magnitude of the biomass in region 2 is not informed by the tagging data set and is reliant on the decline in the CPUE indices to determine the magnitude of the biomass in the region (especially since movement rates are estimated to be low between the two regions). Thus, estimating variation in the distribution of recruitment relaxes the constraints on the trends in stock abundance between the two regions. However, there are limited data, apart from the CPUE indices, to inform the model about historical trends in recruitment and it is unknown whether the CPUE indices are sufficiently reliable to inform the model regarding these long-term recruitment patterns. The estimation of regional recruitment gives the model considerable flexibility to fit CPUE trends. For region 2, the model estimates a substantial declining trend in recruitment during the early 1990s–mid 2000s followed by a general increase in the regional recruitment, whereas recruitment levels in region 1 were considerably more

stable. The apparent conflict between the tagging data and CPUE data within region 1 highlights potential issues related to the precision of stock biomass levels when solely informed by CPUE indices and historical catch.

7. ***Tagging data weighting***. Clearly, the tag data are highly influential in determining the magnitude of stock biomass. The default weighting of these data attributes a relatively large component of the total likelihood to the tag release/recovery data sets – there are two separate tag likelihood components in the total likelihood and there are a considerable number of tag recoveries included in the post-mix data set. The estimated spawning biomass level was relatively insensitive to a range of treatments of the tagging data set in the preliminary modelling phase (e.g. duration of the mixing phase, exclusion of non-PS tag recoveries) (Figure A10). Most influential was the trial that set the over-dispersion parameters at a very high value (50); however, when more realistic values for the OD parameters were assumed or the parameter values were estimated there was only a relatively small change in stock biomass relative to the reference model.

The results of the preliminary modelling guided the formulation of the final set of stock assessment models (see below).

6 Final model options

A key conclusion from the preliminary modelling was that the tagging data were in conflict with both the LL CPUE indices and the PSLS length composition data from the western equatorial region. A large number of model options were investigated to attempt to resolve the conflict among these data sets without success. The model results tended to be dominated by the influence of the tagging data, expressed by the estimation of relatively low levels of biomass, persistent patterns in the recruitment deviates, biomass trends that deviated from the CPUE indices and a relatively poor fit to the PSLS length composition data.

On that basis, the four region model was adopted to further partition the tag release data set and, potentially, minimise the potential for bias introduced by incomplete mixing of tagged fish within the western equatorial region. The final model options were primarily based on the lower level of natural mortality.

For comparative purposes, three different weighting assumptions were applied to the tagging data set. The weightings were applied by the values assigned to the proportion (lambda) of the two components of the tag likelihood included in the total model likelihood. A lambda of 1.0 represents the native weighting of the tag recovery data and gives the tagging data a relatively high weighting in the overall model due to the relatively large number of tag recoveries in the “post mix” data set; a lambda of 0.01 means that the tagging data have a very low weight in the model, approximating a model that excluded the tag data entirely; a lambda of 0.1 represented an intermediate weighting of the tagging data.

The intermediate level of weighting of the tagging data (lambda 0.1) is an arbitrary weighting between two extremes. This level of weighting estimated levels of biomass that were approximately halfway between the biomass levels estimated from the other two weighting options. Levels of weighting higher or lower than the current intermediate level of weighting will result in biomass levels that are closer to the lower or high levels of biomass, respectively.

For presentation purposes, the model option with the intermediate weighting of the tagging data was selected to characterise the main features of the assessment. The specifications of this reference model are presented in Table 4. Key differences in the parameterization of the model among the different tag weighting assumptions are also highlighted. The results of additional model sensitivities are also presented.

6.1 Parameter estimates

The estimated parameters in the assessment model: the overall population scale parameter *R0*, the time series of recruitment deviates, the overall distribution of recruitment among regions, age specific movement parameters, the fishery selectivity parameters, fishery tag reporting rates and the catchability parameters for the CPUE indices.

The age-specific selectivity functions are presented in Figure 22. For the main longline fisheries (LL1N,1S,2,3), full selectivity is attained at about age 18 (quarters). Peak selectivity for the PSLS1N and PSLS1N fisheries occurs at ages 5–8 quarters. (Figure 22). For the PSFS fisheries, selectivity was estimated to be bimodal with a similar level of selectivity for the younger and older modes.

Recruitment deviates were estimated for the quarterly time steps from 1985–2014 (Figure 23). There are longer term trends evident in the recruitment deviates with higher than average recruitment estimated for the mid 1990s–early 2000s and lower recruitment during the mid–late 2000s. These trends correspond to period of higher and lower catches from the PSLS fishery.

The variation in the estimated recruitment deviates is broadly comparable with the assumed level of variation (*SigmaR* of 0.6) (Figure 23). Overall, there is a relatively low level of precision associated with the estimates of the individual recruitment deviates (SE 0.40–0.55).

Recruitment deviates during 2005–2007 differed appreciably amongst the model options with different weightings for the tag likelihoods (Figure 24). For the high tag weighting (lambda 1), recruitment deviates were considerably more variable during this period with a higher recruitment deviate in 2007 Q3 and lower deviates in the two subsequent quarters. This time period corresponds to the period immediately prior to the main tag recovery period.

The reference model assumes a constant distribution of recruitment amongst the model regions. Recruitment was assumed to occur in the equatorial regions only and the distribution of recruitment was estimated to be apportioned 9% to Region 1N, 9% Region 1S and 82% to Region 2. Higher weighting of the tag data substantially increased the proportion of the recruitment in Region 1S (to 55%) at the expense of Region 2.

Movement rates were estimated amongst the model regions. Region 2 was estimated to provide a source of recruitment to the western equatorial regions (1N and 1S) (Figure 25). Limited mixing was estimated to occur between Regions 1N and 1S. The model estimates low movement rates of mature fish amongst the regions, with the exception of reciprocal movement between Region 1S and Region 3 (Figure 25). The estimated population dynamics resulted in the (equilibrium, unexploited) biomass in all regions being dominated by recruitment from Region 2 (Figure 26).

Tag reporting rates were estimated for the non purse-seine fisheries. For some of these fisheries, the estimated reporting rates are unlikely to be influential in the overall assessment as the reported tags were predominately recovered during the tag mix period. However, a considerable proportion of the tag recoveries from the LL1S and LL3 fisheries occurred during the post mixing phase and, hence, the tag reporting rates will have some influence in the model likelihood. For these fisheries, tag reporting rates were estimated at 0.36 and 0.28, respectively (Figure 27), while the reporting rate for the LL1N fishery was estimated to be considerably lower (0.08).

The LL2 fishery reporting rate was estimated at 0.03 although the parameter is poorly determined due to the very small number of tag recoveries from the regional longline fishery (LL2).

6.2 Model diagnostics

Overall, the assessment model fits the general decline in the CPUE indices for the three main fisheries (LL1N, LL1S & LL2) and the seasonal CPUE indices from LL3, particularly during the main fishery period (quarters 2 and 3) (Figure 28). However, the trends in the LL1N&S and LL2 CPUE indices deviate during 1998–2015 and the model is unable to simultaneously fit both sets of indices. Consequently, there is a period of positive residuals in the fit to the LL2 CPUE indices during the late 1990s and negative residuals during the mid-2000s (Figure 29).

For the LL1N&S CPUE indices, there is a period of positive residuals during the late-2000s–early 2010s when the CPUE indices were consistently higher than predicted by the model. Conversely, the CPUE indices were markedly lower than predicted by the model during 2014 and 2015 (Figure 29).

Overall, the variation in the residuals (RMSE 0.23, 0.23 and 0.21, for LL1N, LL1S and LL2) was broadly comparable to the S.E. initially assigned to the CPUE indices.

The fit to the observed number of tag recoveries was examined for those fisheries which accounted for most of the tag returns. The fit to the number of tag recoveries was examined by recombining the tags into individual release periods (i.e. aggregating the releases by age class) and excluding those recoveries that occurred during the mixing period. The fit to the tag recoveries was then examined by time period (quarters) and by age at recovery, for the individual release periods and the aggregated data set (all releases combined).

Most of the observed tag recoveries in the post mix period were from the PSLS1S fishery (Figure 30) and the number of tag recoveries varied considerably amongst the release periods. A high proportion of the total releases occurred during the first four quarters following the mixing period. The fit to the number of recoveries during this period varied amongst release periods. Nonetheless, overall the model substantially under-estimated the number of tag recoveries during the four quarter period (Figure 30). Longer-term recoveries were less vulnerable to the PSLS1S fishery (due to the age specific selectivity) and, hence, numbers of recoveries declined considerably (Figure 30).

There was also a poor fit to the tag recoveries from the PSLS1S fishery by age class (at recovery) (Figure 31). The model substantially under-estimated the overall number of tags recovered from the fishery in the older age classes (11–14 quarters) (Figure 31). The model also under-estimated the number of tags recovered from the older fish (11–14 quarters) in the PSLS1N fishery (Figure 35). This result indicates that the age-specific recoveries are inconsistent with the fishery selectivity functions. These observations may indicate that the initial assignment of tags to the individual age classes did not accurately represent the actual age of the tagged fish at the time of release.

Limited numbers of tags were also recovered by the region 1S PSFS fishery and longline fishery (LL1S) after longer periods at liberty (compared to PSLS1S). Overall, there was a reasonable fit to the tag recoveries from these two fisheries during the main tag recovery period (to 2011). However, the model over-estimates the number of longer term tag recoveries from the older age classes (at recovery), i.e. those fish at liberty for a longer period (Figure 32 and Figure 37), particularly for the longline fishery (Figure 37).

Tag recoveries were also aggregated by region of recovery and time period (Figure 38). The model provides a poor fit to the number of tags recovered from Region 1S during the initial recovery period, reflecting the poor fit to the PSLS1S tag recoveries. The smaller number of longer-term recoveries from Region 1S is reasonably well estimated by the model (Figure 38). The model also provides a reasonable fit to the longer term tag recoveries from Region 1N.

The model provides a reasonable prediction of the tag recoveries from Region 3 from the LL3 fishery. A small number of tags were intermittently recovered in Region 2 by the LL2 fishery and the model predicts a correspondingly low number of tag recoveries from the region (Figure 38). The model has considerable flexibility in fitting to the LL2 and LL3 tag recoveries via the estimation of fishery specific reporting rates.

Not surprisingly, the model options with a higher (lambda 1.0) weighting to the tagging data generally exhibit a better fit to these data relative to the reference model, particularly for the PSLS1S fishery. The model fitted the overall magnitude of the tag recoveries from the fishery and improved the fit to the older (11–14 quarters) age classes (at recovery) (Figure 39).

Paradoxically, the model option with the lowest weighting of the tagging data (lambda 0.01) also resulted in a qualitative improvement in the fit to the tag recoveries from the PSLS1S fishery relative to the reference model, particularly with respect to the overall number of tags recovered (Figure 40). This appears to be due to the differences in model parameterization with respect to other key age-specific variables, especially selectivity, recruitment deviates, regional recruitment distribution and movement.

Overall, there was a good fit to the aggregated length frequency data for most of the main fisheries with comprehensive sampling (Figure 41). However, the model fits to the length composition data from the two Region 2 longline fisheries (LL2 and FL2) were poor. The model predicts considerably larger fish in the catch from these two fisheries than was observed from the fisheries.

An examination of the model residuals from the individual observations reveal a poor fit to the data from key fisheries during certain time periods. Most notably, the model does not account for the increase in the average length of fish in the sampled LL1N, LL1S, and LL3 longline catches from 2000 onwards (Figure 42 and Figure 43). The average length of fish from LL2 and FL2 is over-estimated by the model throughout the data period (Figure 42 and Figure 43).

There is a reasonable fit to the overall length composition of the fish sampled from the PSLS fisheries (Figure 41). However, there is a strong temporal pattern in the model residuals with the model over-estimating the length of fish in the catch from PSLS1S and PSLS1N from 2000 onwards (Figure 42 and Figure 43). Fish sampled during this period were generally smaller than from the preceding period.

The average length of fish sampled from the PSFS is highly variable (Figure 42) probably reflecting the proportion of the catch sampled from the smaller and larger modes of the combined length composition. The model prediction of average length represents the length of fish in the intermediate length range (80–100 cm) (Figure 41).

The model option with the higher weighting (lambda 1.0) of the tag data resulted in a considerably worse fit to the PSLS length composition data; i.e., the over-estimation of the fish caught from these fisheries was more pronounced than the reference model.

Conversely, the model option with the lower weighting (lambda 0.01) of the tag data resulted in a slightly better fit to the PSLS length composition data compared to the reference model, although the over-estimation of the length of fish sampled in recent years (from 2005) persisted. The over-estimation of the length composition of the Region 2 longline fisheries persisted in the low tag weighting model. This suggests a degree of conflict between these data and the CPUE indices and/or a miss-specification of other model structural assumptions related to selectivity and recruitment.

6.3 Derived quantities

For the reference model, the spawning biomass in Regions 1N&1S accounted for about 40% of the initial biomass. Region 2 accounted for 40% and Region 3 20% (Figure 44 and Figure 45). Relative trends in spawning biomass are estimated to be comparable between Region 1N&1S and Region 3, while the level of depletion was somewhat higher for Region 2.

For all model options, spawning biomass was estimated to decline in the late 2000s–early 2010s following a period of lower recruitment during 2004–2006 (Figure 47). Spawning biomass was estimated to reach the lowest level during 2009–2010 (Figure 44). Subsequent trends in spawning biomass follow recent recruitment patterns, with higher recruitment in 2007 resulting in an increase in spawning biomass in 2011–2012, following an increase in the CPUE indices in the equatorial regions (Regions 1N&1S and 2) (Figure 47). Total recruitment is estimated to have been lower in the more recent years and spawning biomass declined again during 2012–2015.

The overall magnitude of spawning biomass was strongly influenced by the relative weighting of the tagging data set (Figure 46). The level of spawning biomass for the model option with the lowest weighting of the tagging data (lambda 0.01) was approximately double the biomass level of the high tag data weighting option (lambda 1.0).

For the reference model, the estimates of fishing mortality were relatively low for the fisheries within Region 2 and Region 3 (Figure 48). Fishing mortality rates for the LL1N and LL1S fisheries were comparable to the longline fishing mortality rates in the two other regions. By comparison, fishing mortality rates for the PSLS fisheries were estimated to be relatively high in Region 1N and Region 1S from the mid-1990s (Figure 48).

The recent (2013–14) pattern of age specific fishing mortality reveals higher levels of fishing mortality for the age classes (4–9 quarters) vulnerable to the PSLS fisheries (Figure 49). For these younger age classes, the estimated level of fishing mortality was similar for the range of model options with different levels of weighting for the tagging data set. These younger cohorts were not included within the portion of the population included in the tag release/recovery programme.

Conversely, the estimated fishing mortality rates for the age classes vulnerable to the longline fishery (20+ quarters) were sensitive to the relative weighting of the tagging data set. The model option with the higher weighting of the tagging data estimated the highest levels of fishing mortality for the older age classes (Figure 49).

7 Stock status

MSY based estimates of stock status were determined for the reference model (lambda 0.1) with alternative values of SRR steepness (0.70, 0.90) and the alternative (higher) level of natural mortality. Stock status was also determined for the model option with the higher, default (native) weighting of the tag likelihood components (lambda 1.0), incorporating the three levels of steepness, and the low tag weighting model option (lambda 0.01) (Table 6).

MSY based reference points were derived for the model options based on the average F-at-age matrix for 2013–14 (Figure 49). The period was considered representative of the recent average pattern of exploitation from the fishery. However, it is important to note that recent fishery catches from the fishery have been quite variable; variation in the proportion of catches between the main fishing gears (LL and PSLS) are likely to influence the F-at-age matrix and, hence, *MSY* based indicators.

The data included in the stock assessment models are uninformative regarding SRR steepness. There was a small improvement in the total model likelihoods with different the higher levels of steepness, although the improvement was not considered significant.

For the selected model options, point estimates of *MSY* ranged from 86,000 mt to145,000 mt (Table 6) compared to most recent annual catches of about 92,000 mt (Table 2). The lower range of *MSY* values (86–92 k t) were derived from model options with the native weighting of the tagging data (lambda 1.0), while the low tag weighting yielded an *MSY* estimate of 145 k t. Lower levels of SRR steepness yielded comparatively lower estimates of *MSY*, while the higher natural mortality option yielded an *MSY* estimate that was 18% higher than the base level.

The levels of stock depletion and current fishing mortality were also sensitive to the relative weighting of the tagging data; i.e. depletion levels and fishing mortality were considerably higher for the model options with the higher weighting of the tag data set (lambda = 1.0) (Table 6). Clearly, the levels of stock depletion and current fishing mortality derived from the model options with the intermediate level of weighting of the tagging data set will have been influenced by the arbitrary selection of the relative weighting of these data (lambda =0.1). A lower or higher weighting of these data would have resulted in more or less optimistic estimates of stock status and yield.

Nonetheless, the general conclusions regarding current stock status relative to the *MSY* based benchmarks are not fundamentally different for the range of model options, although the proximity to the *MSY* benchmarks is sensitive to the different of tag data weight assumptions. For all model options, current (2015) spawning biomass was estimated to be above the SB_{MSY} level ($SB_{2015}/SB_{MSY} > 1.0$). Fishing mortality was estimated to be below the F_{MSY} level ($F_{2015}/F_{MSY} < 1.0$) for all options, with the exception of the model option with the high tag weighting and lower value of steepness. Nonetheless, there is considerable uncertainty associated with the estimates of current fishing mortality and there was a small probability that the F_{MSY} level had been breached for the high tag weighting, steepness 0.8 option also (Table 6).

Annual estimates of stock status (dynamic *MSY*) were determined for 1980–2015 for the intermediate and high tag weighting model options (Figure 50). For the *reference* model (lambda 0.1), fishing mortality rates were estimated to have increased during the 1990s although have remained well below the F_{MSY}. Biomass was estimated to have declined considerably from the late 1990s before stabilizing at a level well above the SB_{MSY} level in recent years (Figure 50).

For the higher tag weighting (lambda 1.0), the stock trajectory was broadly comparable to the *reference* model, although the *MSY* reference levels are shifted and the stock trends are more pronounced (Figure 50). Fishing mortality rates were estimated to have been at about the F_{MSY} level during 1999–2009 and below the F_{MSY} level in subsequent years. Stock biomass declined considerably, approaching the SB_{MSY} in 2010–2015 (Figure 50).

8 Stock Projections

Stock projections were conducted for the *reference* model (lambda 0.1) and the higher tag weighting model (lambda 1.0) assuming the intermediate value of SRR steepness (0.80). The projections were conducted for a 10 year period (2016–2025) at a constant level of catch set as a multiple of the fishery catches in 2015. Three levels of catch were investigated representing 80% (74,200 mt), 100% (92,700 mt) and 120% (111,300 mt) of the 2015 catch level. Recruitment during the projection period was at the equilibrium level. The uncertainty associated with the projected biomass was derived from the covariance matrix. For each stock scenario, the probability of the biomass being below the SB_{MSY} level was determined after 3 years (2018), 5 years (2020) and 10 years (2025).

The uncertainty associated with the projected biomass propagates rapidly, reflecting the uncertainty associated with the equilibrium recruitment level. For the *reference* model, current (2015) levels of catch are lower than the equilibrium surplus production and the stock biomass is projected to increase slightly over the 10 year period, with a relatively high probability of being maintained above the SB_{MSY} level (Table 7). Catches 20% higher than the 2015 level resulted in the biomass being maintained at approximately the SB_{2015} for the entire projection period.

For the higher weighting of the tagging data (Lambda 1.0), the current (2015) level of catch is very similar to the *MSY* level. Stock projections at this level of catch maintain the biomass at about the SB_{MSY} level throughout the projection period (Table 7). Projections at 80% of the 2015 catch maintain the stock biomass above the SB_{MSY} throughout the period, while catches at 120% of the 2015 level reduce the stock below SB_{MSY} within the first five years of the projection period.

9 Discussion

There are four major changes in the structure of the current assessment models compared to the previous assessments (Kolody *et al* 2010, Langley *et al* 2013b). Firstly, the current assessment model is spatially stratified (three regions), while the previous assessments were based on a spatially aggregated model.

Secondly, the current assessment incorporates the tag release/recovery data set. These changes in model structure are inherently linked – the spatial scale of the dispersal of tags from the main release site means that it is necessary to impose a regional stratification to reduce the potential for bias in the estimation of stock size due to the violation of the tag mixing assumptions. Examination of the tag recoveries reveals that tag dispersal is not adequate for stock assessment purposes at the basin scale.

Thirdly, the regional configuration of the assessment model necessitates the inclusion of region-specific longline CPUE indices. The CPUE indices incorporated in the current assessment are derived from combined logsheet data (from Japan, Korea and Taiwan), whereas previous assessments were based on longline CPUE indices from the Japanese fleet only. The combined CPUE indices reveal similar general declines in abundance amongst the three regions during 1980–2005, although trends in stock abundance differed amongst regions during the more recent period.

The fourth major difference was the time period of the assessment model. The previous assessment models commenced in 1952 and incorporated the entire time-series of longline CPUE indices (from the early 1950s), while the current assessment model commences in 1975 and incorporates CPUE indices from 1979 only. The earlier CPUE indices were excluded from the current assessment due to the disjunct in the series between the mid-1970s and early 1980s and the exceptionally high CPUE indices in the late 1970s. The overall change in the magnitude of the CPUE indices appears to be related to changes in the operation of the longline fishery which were not adequately accounted for in the CPUE standardisation.

The deviation in the trends in the CPUE indices from the equatorial regions during 2005–2015 were not adequately accounted for in the current stock assessment models. Movement of adult fish was estimated to be low between the two regions and, therefore, differential trends in stock biomass can only be accounted for differential patterns in exploitation, especially given that the regional distribution of recruitment was assumed to be temporally invariant. It was considered that there was insufficient information available to estimate recruitment at the regional scale. The low rates of movement between the two regions appear to be strongly informed by the tagging data set.

The different patterns in the recent CPUE indices appear to be correlated with the prevailing oceanographic conditions as indexed by the IODI. The higher CPUE in the western area during the late-2000s–early 2010s corresponds to positive IODI, while the recent decline in CPUE in the region corresponds to a shift to negative IODI in 2013–2014. Two basic mechanisms could account for these differential patterns in the CPUE indices: the change in prevailing oceanographic conditions may have influenced the relative distribution of the bigeye biomass between the two equatorial regions and/or the prevailing oceanographic conditions may have influenced the vertical distribution of bigeye, thereby, changing the relative catchability in the two regions. Both hypotheses were investigated within the framework of the stock assessment model; the first by incorporating the IODI as a covariate in the parameterisation of movement between the two regions and the second by incorporating the IODI as a covariate in the parameterisation of the longline catchability in each region. Neither model option adequately accounted for the different trends in the recent CPUE indices.

Recent trends in stock abundance in the western region also appear to be strongly influenced by the tag release/recovery data set. These data influence the estimates of recruitment strength during 2004–2007 and, thereby, estimates of the abundance of longline vulnerable biomass in the subsequent period (during 2009–2011). This is the period when the fit to the western (LL1) CPUE indices was very poor; i.e., the model predicted a considerably lower biomass than indicated by the CPUE indices. In addition to the tagging data, the estimates of stock size during the period were also informed by the lower PSLS catches during 2004–2008 and the temporal trends in the PSLS length composition data during the mid–late 2000s. Cumulatively, these data influenced the estimates of recruitment strength during 2004–2006 (low) and 2007 (high).

The tag release/recovery data are also highly influential in the estimation of the overall magnitude of the biomass of the stock, mediating the scale of the overall magnitude of the decline in the CPUE indices. The native weighting (lambda 1.0) of the tagging data set, based on the number of tag recoveries, assigns a relatively high weighting to the tagging data in the formulation of the total likelihood. The model estimated relatively high levels of fishing mortality during the tag recovery phase, thereby, influencing the estimation of the total population scale parameter (*R0*).

Reducing the influence of the tagging data yielded considerably higher estimates of stock size and, correspondingly, more optimistic estimates of current stock status and yield. The model has considerable flexibility to accommodate the contrasting emphasis of these two data sets via the estimation of the total population scale parameter, recruitment deviates, the regional distribution of recruitment and age-specific movement. There are very limited data incorporated within the model to allow the reliable estimation of recruitment deviates. Bias in the estimation of the recruitment deviates will influence the estimation of key stock parameters (*R0*) and will influence estimates of stock status and yield. The reliability of the assessment models would be improved substantially by the incorporation of data that was more informative regarding recruitment patterns, such as CPUE indices from the small fish PSLS fishery and/or reliable length/age catch composition data, especially from the longline fisheries.

Reweighting of the main data sets (CPUE and tag) can not resolve the inherent conflict between the two main data sets. This can only be addressed by improved understanding of the potential biases in the tag release/recovery and CPUE data sets that are not adequately accounted for in the structure of the current assessment models. Such biases may relate to the tag mixing assumptions, the assumption of proportionality between the longline CPUE indices and stock abundance, and the assumption of invariant selectivity for the key fisheries associated with the CPUE and tag data sets (longline and purse seine).

The bigeye tag recovery data set is predominantly based on the recovery of tagged fish by the PSLS fishery which catches bigeye over a relative narrow length (and age) range. The unbiased estimation of stock abundance from tagging data is dependent on either the homogeneous mixing of tagged fish within the population or the recovery of tags from catches that are taken in proportion to the abundance of the stock. The latter criterion is unlikely to have been met during the recovery phase, particularly given the appreciable shift in the operation of the purse seine fleet with limited fishing occurring within the Somali EEZ during the main recovery period (2007–2008) (Figure 17). During 2007–2009, approximately 6–8% of the bigeye purse seine catch was taken within the Somali EZZ compared to 12–16% during the preceding period (2000–2005) (source: IOTC Secretariat, Glaser *et al* 2015) (Figure 51).

The dispersal of tagged bigeye was examined to determine an appropriate mixing period for the assesment mdoelling. The aggregated nature of tag recoveries indicated that periods of less than six months were unlikely to allow for the dispersal of tags throughout the population within the western equatorial region. However, the age specific selectivity of bigeye tuna meant that tagged fish were vulnerable to the PSLS fishery for a relatively limited period. Consequently, assuming a mixing period of at least 18 months substantially reduced the number of tags available for inclusion in the recovery data set. On that basis, a mixing period of 12 months was adopted for the main model options.

Nonetheless, an examination of the tag recovery rates (post mixing) from within the western equatorial region revealed that tag densities declined with increasing distance from the primary release site. On that basis, the spatial structure of the assessment model was refined accommodate the heterogeneity in the distribution of tagged fish within the western equatorial region (after 12 months at liberty). The change in model structure reduced some of the conflict between the CPUE and tagging data sets; i.e., for the initial three region model structure there was a larger (250%) difference in the overall biomass level estimated from the model options with high and low weighting of the tagging data compared to the four region model structure (200%). However, clearly the additional stratification did not fully account for the conflict between the two data sets. Future assessments should further investigate the spatial scale of tag dispersal and refine the spatial stratification of the model, accordingly. Alternatively, the assessment modelling framework could incorporate a component to estimate tag dispersal that is independent of the parameterization of the stock movement dynamics.

Some preliminary modelling of the dispersal of tagged bigeye from the release location was previously conducted based on the assumption that the dispersal of small bigeye was strongly influenced by the prevailing oceanographic conditions (current flow). However, the modelling is not sufficiently advanced to reliably predict the pattern of dispersal and the extent of mixing of tagged fish amongst the total population. Further attention should be given to the development of such modelling approaches.

There was a general decline in the length of fish sampled from the PSLS fisheries from the mid–late 1990s. The decline in fish length is generally consistent with the increase in the catch (and fishing mortality) for the two fisheries (PSLS1N & 1S), although the model under-estimates the extent of the decline in the length of fish sampled, especially when the tagging data are assigned a high weighting. There have been recent changes in the spatial distribution of the operation of the purse seine fleet (e.g. related to access to the Somali EEZ) that may have resulted in subtle changes in the length composition of the catch from the fishery (Chassot *et al* 2010). Consequently, it is unknown whether the change in fish size is indicative of changes in fishing mortality or operational changes in the fishery (or a combination of the two factors). Model trials that down-weighted the PSLS length composition data (from ESS 10 to 1) resulted in a slightly more optimistic estimate of stock status and a larger estimate of *MSY* (*TagLambda01-dwtPSLF*). Conversely, model trials that down-weighted the PSLS length composition data (from ESS 10 to 50) resulted in a more pessimistic estimate of stock status and a lower estimate of *MSY* (*TagLambda01-upwtPSLF*), while the fit to the tagging data set deteriorated.

There is an interaction between the PSLS length composition data and the tag recovery data set via the estimation of the selectivity function for the PSLS fishery. There were a considerable number of tagged fish recovered from the PSLS fishery after 18 months at liberty. These fish were older (at recovery) than predicted by the assessment model based on the estimated selectivity function of the

PSLS fishery. However, the true age of the tagged fish is unknown and the age at release (and hence recovery) was assumed based on the established growth function. A range of alternative age assignments to the tag release groups were investigated during the assessment modelling. For example, all tagged fish were assigned to the age (6 quarters) that corresponded to the median length of fish at release (*TagLambda1-RelAge6*). These changes did not appreciably change the estimates of stock status from the base assessment models. Nonetheless, the estimation of the PSLS selectivity function is sensitive to the relative weighting of the tagging data and the length composition data, reiterating the need to further examine the factors that may be influencing the PSLS length data sets.

There is a need to improve the understanding of the reliability of the length composition data collected from the longline fisheries. The current study endeavoured to address some of the previous concerns regarding the reliability of the longline size data collected from the Taiwanese fleet and limited the length composition data sets to the core areas of the bigeye tuna fishery. However, the fundamental issues with these data sets remain with substantially larger fish sampled from the catch from 2000 onwards. Attempts to account for these changes by estimating different selectivity functions for the two periods were not successful and the large size of fish sampled in the latter period could not be accommodated in any assessment model option. This indicated that the length data are inconsistent with the growth parameters for the older age classes and/or the estimated levels of fishing mortality. Further work is required to validate the length composition data collected from the various longline fleets in the recent period.

The assessment model assumes that the age-based selectivity of the longline fishery has remained static over the model period. This is despite changes in fishing operation that are related to an increase in the degree of targeting of bigeye tuna (relative to yellowfin tuna). Changes in the distribution of fishing activity and the configuration of the longline gear could result in changes in the size of bigeye tuna caught by the longline fishery. Any changes in gear selectivity would interact with the estimation of stock abundance as the CPUE indices are mediated by the selectivity of the longline fishery.

There is a lack of reliable length frequency data from the "Other" fisheries (OT1 & 2). These fisheries accounted for approximately 5,000 mt of bigeye catch in recent years. In the absence of reliable length data, these fisheries were assumed to have a selectivity equivalent to the PSFS fishery, reflecting the bimodality in the length compositions from the two fisheries.

The stock assessment conclusions were also sensitive to the assumed level of natural mortality of the older age classes (*Mhigh* 0.10 and *Mlow* 0.0625 per quarter). As previously noted, there were a considerable number of tag recaptures after 7–8 years at liberty, indicating a considerable proportion of these fish reach an age of 8–10 years; 8 tags were recovered after 10 years at liberty and a few tags were recovered in the most recent year (2015) corresponding to an age at recovery of 11–12 years. On that basis, the lower level of natural mortality was considered to be the more plausible option, although it is recognised that there is considerable uncertainty associated with natural mortality. This source of uncertainty has not been incorporated amongst the suite of model options. Preliminary model options estimated the levels of natural mortality for both younger and older age classes. However, the estimates were sensitive to the structural assumptions of the model, including the spatial stratification of the model domain. This result indicated that model based estimates of natural mortality should be treated with caution due to potential biases in the tag recovery process (tag mixing) and potential confounding between the estimation of natural mortality and other parameters (especially related to the estimation of recruitment and selectivity and movement).

Catch history of the longline fishery may be influenced by the documented miss-reporting of the catches of bigeye tuna from the Atlantic longline fishery. IOTC to provide reference.

Comparison with Yellowfin tuna assessment results

The 2015 yellowfin tuna stock assessment (Langley 2015) shares a number of important similarities with the bigeye assessment, especially the regional stratification of the western equatorial region, the associated definition of the constituent fisheries with the region and the inclusion of the tag release/recovery data sets, with both bigeye and yellowfin recoveries dominated by recoveries from the purse seine fishery. A comparison of some key parameters between the two assessment models may provide additional insights into the performance of the two assessment models.

For the two assessment models, there are similarities in the recruitment time series estimated in the period immediately prior to the main tag recovery phase. For both species, recruitment was estimated to be relatively low during 2004–2006. The correspondence of these trends may indicate that the two assessments are influenced by similar dynamics in the species specific tag release/recovery data sets.

The tag reporting rate for bigeye tuna from the longline fishery in western equatorial region was estimated to be considerably higher (0.36) than the tag reporting rate for yellowfin tuna (0.079) from essentially the equivalent longline fleet (Langley 2015). It is unknown why the estimated reporting rates would differ substantially between the two species, but may be a function of the influence of the other data sets and/or structural assumptions of the two stock assessment models.

10 Acknowledgements

Adam Langley's participation of in the assessment was funded by IOTC. The staff of the IOTC, specifically Fabio Fiorellato and James Geehan provided support and assistance with the project. Simon Hoyle provided valuable comments on an earlier version of the report leading to further exploration of the tag release/recovery data set.

11 References

Chassot, E., Dewals, P., Floch, L., Lucas, V., Morales-Vargas, M., Kaplan, D. (2010). Analysis of the effects of Somali piracy on the European tuna purse seine fisheries of the Indian Ocean. IOTC Scientific Committee IOTC-2010-SC-09.

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.

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

Francis R. I. C. C. (in press). Revisiting data weighting in statistical fisheries stock assessment models. *Fisheries Research*.

Gaertner, D.; Hallier J.P. (2015). Tag Shedding by Tropical Tunas in the Indian Oceanand other factors affecting the shedding rate. Fisheries Research 163, 98-105.

Geehan, J.; Hoyle, S. (2013). Review of length frequency data of the Taiwan,China distant water longline fleet, IOTC-2013-WPTT15-41 Rev_1. Indian Ocean Tuna Commission Working Party on Tropical Tunas, San Sebastian, Spain, 23–28 October, 2013.

Glaser, S.M.; Roberts P.M.; Mazurek R.H.; Hurlburt K.J.; Kane-Hartnett L. (2015) Securing Somali Fisheries. Denver, CO: One Earth Future Foundation. DOI: 10.18289/OEF.2015.001

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.

Hillary, R.M., Million, J., Anganuzzi, A., Areso, J.J. 2008. Tag shedding and reporting rate estimates for Indian Ocean tuna using double-tagging and tag-seeding experiments. IOTC-2008-WPTDA-04.

Hoyle, S.D., Kim, D.N., Lee, S.I., Matsumoto, T., Satoh, K., Yeh, Y.M. (2016). Collaborative study of tropical tuna CPUE from multiple Indian Ocean longline fleets in 2016. IOTC-2016-WPTT18-14.

Hoyle SD, Leroy BM, Nicol SJ, Hampton J. 2015. 'Covariates of release mortality and tag loss in large-scale tuna tagging experiments'. Fisheries Research 163, 106-118.

Hoyle, S.D, 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.

Hoyle, S.D., Okamoto, H. (2015). Descriptive analyses of the Japanese Indian Ocean longline fishery, focusing on tropical areas. IOTC-2015-WPTT17-INF08.

Hoyle, S.D., Yeh, Y.M., Chang, S.T., Wu, R.F (2015). Descriptive analyses of the Taiwanese Indian Ocean longline fishery, focusing on tropical areas. IOTC-2015-WPTT17-INF09.

ICCAT 2015. Report of the 2015 ICCAT bigeye tuna stock assessment session. Madrid, Spain - July 13 to 17, 2015.

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

IOTC 2013. Report of the Fifthteenth Session of the IOTC Working Party on Tropical Tunas. IOTC–2013–WPTT15–R.

IOTC 2015. Review of the statistical data and fishery trends for tropical tunas. IOTC-2015-WPTT17-07 Rev_1.

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. 2015. Stock assessment of yellowfin tuna in the Indian Ocean using Stock Synthesis. IOTC-2015-WPTT17-30.

Langley, A.; Herrera, M.; Sharma, R. 2013a. Stock assessment of bigeye tuna in the Indian Ocean for 2012. IOTC-2013-WPTT15-30.

Langley, A.; Herrera, M.; Sharma, R. 2013b. Stock assessment of bigeye tuna in the Indian Ocean for 2012. IOTC-2013-WPTT15-30-Rev_1.

Matsumoto, T.; Satoh, K.; Okamoto, H. 2013. Japanese longline CPUE for bigeye tuna in the Indian Ocean standardized by GLM. IOTC-2013-WPTT15-25.

Methot, R.D. (2015). User manual for Stock Synthesis, model version 3.24s. Updated February 11, 2015

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

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., 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.

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.

Table 1. Definitions of the individual model fisheries.

Code	Method	Region	Flag	Notes
FL2	Longline, fresh tuna fleets	2	All	
LL1N	Longline, distant water	1N	All	
LL1S	Longline, distant water	1S	All	
LL2	Longline, distant water	2	All	
LL3	Longline, distant water	3	All	
PSFS1N	Purse seine, free school	1N	All	
PSFS1S	Purse seine, free school	1S	All	
PSFS2	Purse seine, free school	2	All	
PSLS1N	Purse seine, associated sets	1N	All	
PSLS1S	Purse seine, associated sets	1S	All	
PSLS2	Purse seine, associated sets	2	All	
BB1	Baitboat and small scale encircling gears (PSS, RN)	1N	All	Primarily catch from the Maldives baitboat fishery. Predominantly catching small bigeye.
LINE2	Mixed gears (hand-line, gillnet/longline combination)	2	All	Gears grouped on the basis that primarily catch large bigeye.
OT1	Other (trolling, gillnet, unclassified)	1N	All	
OT2	Other (trolling, gillnet, unclassified)	2	All	

Table 2: Recent bigeye tuna catches (mt) by fishery included in the stock assessment model. The annual catches are presented for 2014 and 2015 and the average annual catch is presented for 2012-14.

Fishery	Time period		
	2012-14	2014	2015
FL2	10,147	13,383	16,153
LL1N	10,416	7,552	5,764
LL1S	16,486	15,868	16,316
LL2	6,394	7,568	4,704
LL3	3,824	4,833	4,886
PSFS1N	1,073	888	2,013
PSFS1S	3,097	3,957	6,753
PSFS2	0	0	200
PSLS1N	5,418	6,958	7,233
PSLS1S	6,359	7,915	8,627
PSLS2	293	159	0
BB1	5,083	6,188	5,717
LINE2	6,395	9,001	8,132
OT1	1,931	2,468	2,492
OT2	3,553	4,313	4,050
Total	80,469	91,052	93,040

Table 3. The number of tag recoveries by fishery (SS adjusted data set). The mixing period was the four quarters following the release.

Fishery	Recovery period	
	Mix	Post Mix
FL2	0	0
LL1N	0	14
LL1S	24	199
LL2	0	7
LL3	4	57
PSFS1N	15	27
PSFS1S	190	184
PSFS2	0	0
OT1	110	20
OT2	0	0
PSLS1N	1007	205
PSLS1S	3536	975
PSLS2	4	0
BB1	0	0
LINE2	0	0
Total	4890	1688

Table 4. Main structural assumptions of the bigeye tuna *reference* model and details of estimated parameters.

Category	Assumptions	Parameters
Recruitment	Occurs at the start of each quarter as 0 age fish. Recruitment is a function of Beverton-Holt stock-recruitment relationship (SRR). Regional apportionment of recruitment to R1N, R1S and R2 only. Temporal recruitment deviates from SRR, 1985–2014. No temporal deviates in the proportion of recruitment allocated to R1N, R1S and R2.	R_0 Norm(10,10)*;* *h* = 0.80 *PropR2* Norm(0,1.0) *SigmaR* = 0.6. 140 deviates.
Initial population	A function of the equilibrium recruitment in each region assuming population in an initial, exploited state in 1975. Initial fishing mortality estimated for LL1N,1S,2,3 fisheries.	Norm(0.10,99)
Age and growth	40 quarterly age-classes, with the last representing a plus group. Growth based on VonBert growth model with age-specific *k* to approximate the mean length at age determined by Eveson *et al* (2012). SD of length-at-age based on a constant coefficient of variation of average length-at-age. Mean weights (W_j) from the weight-length relationship $W = aL^b$.	*Linfinity* = 150.913 cm, *k* (base) = 0.332, *k* deviates for ages 1,8,9,10. CV =0.10 *a* = 3.661e-05, *b* = 2.901
Natural mortality	Age-specific, fixed. Ramp function Age 0-12, initial 0.2 at age 0. Constant age 12-40 at 0.0625	
Maturity	Length specific logistic function from Shono *et al* (2009). Mature population includes both male and female fish (single sex model).	Mat50_Fem 110.888 cm Mat_slope_Fem -0.25
Movement	Age-dependent with two blocks; age classes 3-8 and 15-40. Ramp function Age 8-15. No movement prior to age class 3. Constant among quarters.	12 movement coefs. Norm(0,4).
Selectivity	Age specific, constant over time. Principal longline fisheries (LL1N,1S,2,3) share logistic selectivity parameters. PSLS fisheries. Separate selectivity for PSLS1N, common selectivity PSLS1S and PSLS2 Common selectivity for all PSFS fisheries. LF2 fishery logistic selectivity. LINE2 share principal LL selectivity. BB fishery: double normal selectivity.	Logistic *p1* Norm(20,10), *p2* Norm(1,10) Double Normal Five node cubic spline

	OT 1 & 2 share PSFS selectivity. CPUE indices share principal LL selectivity.	
Catchability	Temporally invariant. Shared regional catchability coefficient. No seasonal variation in catchability for LL CPUE. LL2,3 CPUE indices have CV of 0.2; LL1N,1S CV 0.25.	Unconstrained parameter *LLq*
Fishing mortality	Hybrid approach (method 3, see Methot & Wetzel 2013).	
Tag mixing	Tags assumed to be randomly mixed at the model region level four quarters following the quarter of release. Accumulation after 28 quarters	
Tag reporting	All (adjusted) reporting rates constant over time. Common tag reporting rate fixed for all PS fisheries. Non PS tag reporting rates uninformative priors.	 PS RR 1.0 Other fisheries Norm(-0.7,5)
Tag variation	Over dispersion parameters estimated for each tag release groups.	beta prior (mean 10, sd 3)
Length composition	Multinomial error structure. PSLS length samples assigned maximum ESS of 10. PSFS length samples assigned maximum ESS of 1.0. LL and Other fisheries length samples assigned ESS of maximum 1.0.	

Table 5. Details of objective function components for the key set of stock assessment models and sensitivities.

Model option	Total	Component					
		Survey	Length_comp	Tag_comp	Tag_negbin	Recruitment	Parm_priors
./TagLambda001	795.46	-466.22	1,207.11	25.85	21.78	-20.04	26.16
./TagLambda01	1,174.80	-452.04	1,185.51	218.81	200.07	-12.91	34.65
./TagLambda01-Mhigh	1,176.96	-454.96	1,187.70	219.92	207.58	-18.53	34.39
./TagLambda1	4,665.38	-431.36	1,255.92	2,016.42	1,709.25	12.45	102.20
./TagLambda1-RelAge6	4,723.84	-432.28	1,245.06	2,067.05	1,721.18	19.16	103.31

Table 6. Estimates of management quantities for the set of stock assessment models and selected model sensitivities. The 95% confidence intervals for the current stock status metrics are provided for the reference model.

Model	SB_0	SB_{MSY}	SB_{MSY}/SB_0	SB_{2015}	SB_{2015}/SB_0	SB_{2015}/SB_{MSY}	F_{2015}/F_{MSY}	MSY
./TagLambda01	2,102,400	614,278	0.292	859,204	0.41	1.40 (1.27-1.52)	0.59 (0.45-0.74)	115,786
./TagLambda01-steep70	2,228,900	718,078	0.322	919,937	0.41	1.28	0.69	108,273
./TagLambda01-steep90	2,008,970	520,338	0.259	815,496	0.41	1.57	0.52	122,732
./TagLambda01-Mhigh	1,285,950	360,771	0.281	524,146	0.41	1.45	0.48	137,001
./TagLambda01-dwtPSLF	2,349,420	683,831	0.291	968,950	0.41	1.42	0.53	135,390
./TagLambda01-upwtPSLF	1,566,650	459,515	0.293	583,845	0.37	1.27	0.70	92,179
./TagLambda1	1,486,900	438,431	0.295	510,399	0.34	1.16 (1.06-1.27)	0.88 (0.71-1.05)	92,811
./TagLambda1-steep70	1,529,250	495,228	0.324	521,990	0.34	1.05	1.05	86,248
./TagLambda1-steep90	1,385,020	364,320	0.263	467,022	0.34	1.28	0.82	98,758
./TagLambda1-RelAge6	1,450,980	416,561	0.287	500,726	0.35	1.20	0.90	90,418
./TagLambda001	2,591,050	595,521	0.230	1,176,930	0.45	1.98	0.39	145,454

Table 7. Projected stock status relative to SB_{MSY} and the probability of being below SB_{MSY} in 3-, 5- and 10 years for three alternative levels of catch (relative to 2015) for the base model options with different levels of natural mortality. A value of zero for SB/SB_{MSY} indicates that catches exceeded the stock biomass and the stock crashed.

Model option	Catch	3 years (2018)		5 year (2020)		10 year (2025)	
		SB/SB_{MSY}	$Pr(SB<SB_{MSY})$	SB/SB_{MSY}	$Pr(SB<SB_{MSY})$	SB/SB_{MSY}	$Pr(SB<SB_{MSY})$
Lambda 0.1	80%	1.48	0.00	1.57	0.00	1.90	0.00
(Reference)	100%	1.42	0.00	1.44	0.01	1.64	0.00
	120%	1.36	0.00	1.32	0.04	1.39	0.00
Lambda 1.0	80%	1.13	0.12	1.15	0.16	1.42	0.06
	100%	1.05	0.31	0.99	0.53	1.05	0.40
	120%	0.97	0.59	0.83	0.87	0.67	0.88

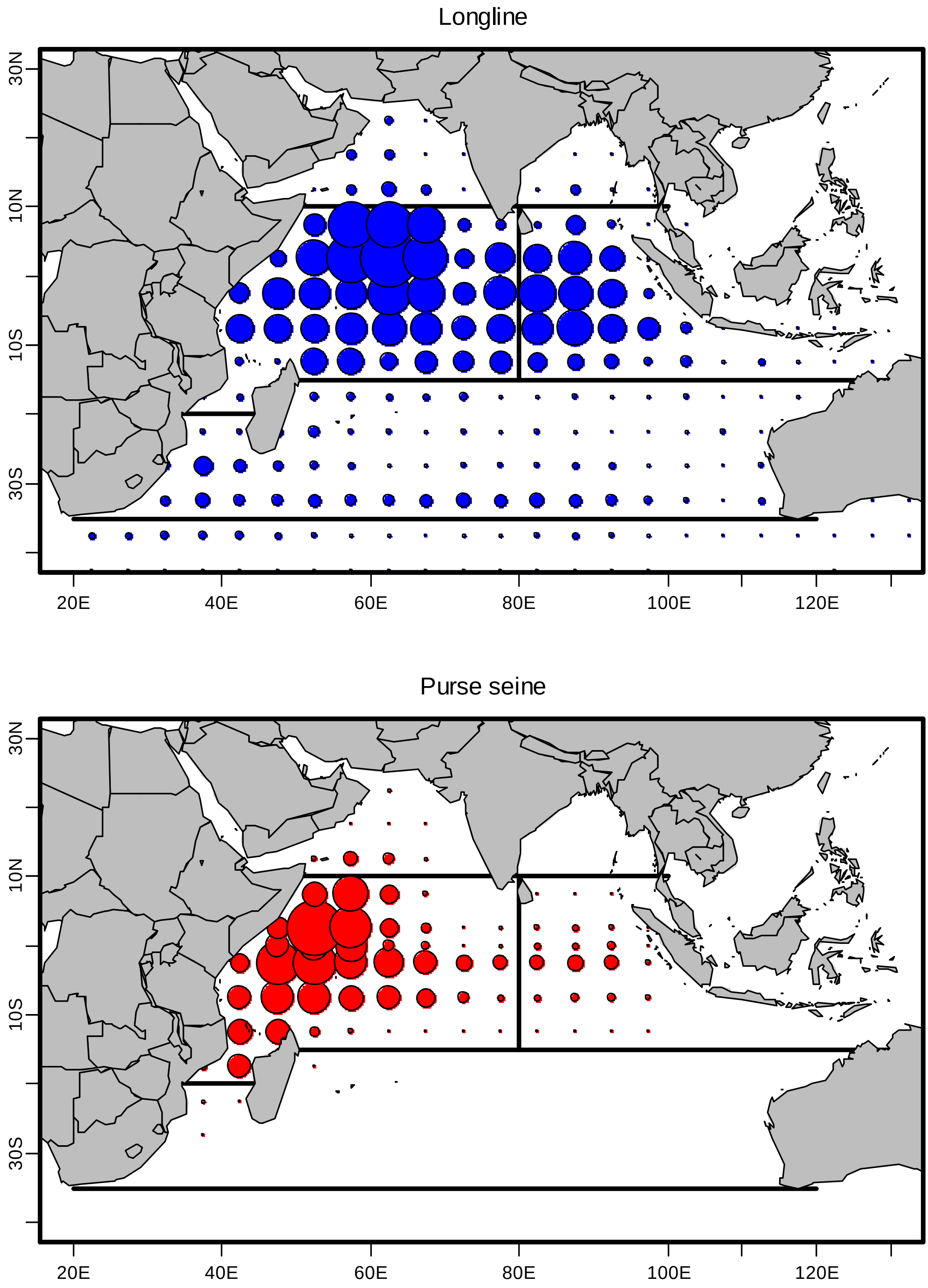

Figure 1. LL and PS catch (max bet catch in 5 deg cell aggregated over time 70159.62 mt)

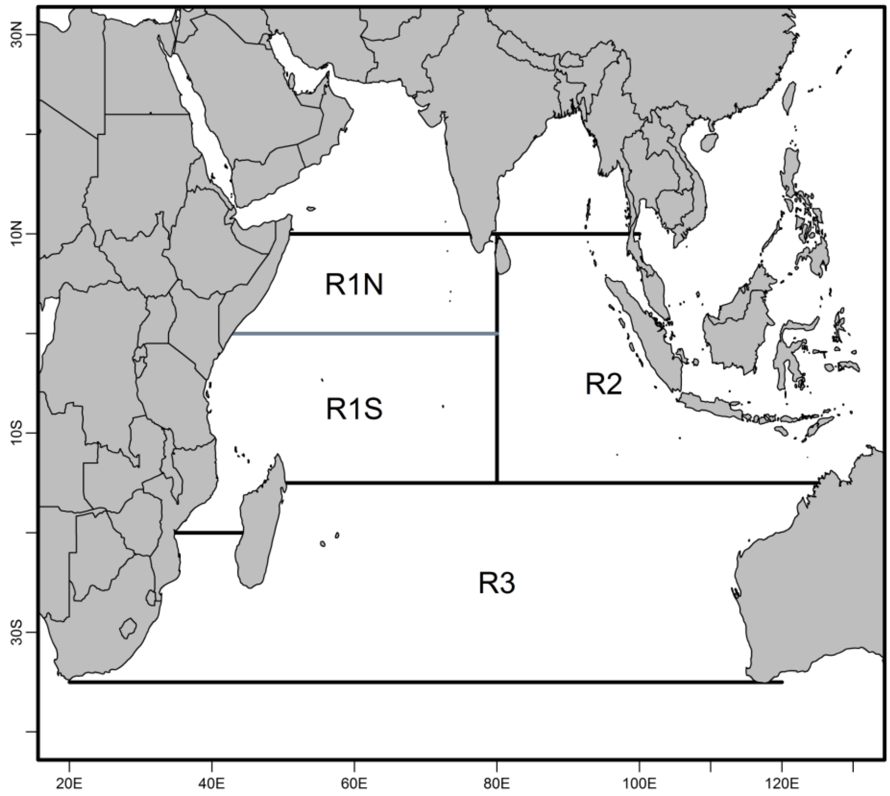

Figure 2. The spatial structure applied to define the model fisheries and configure the assessment models.

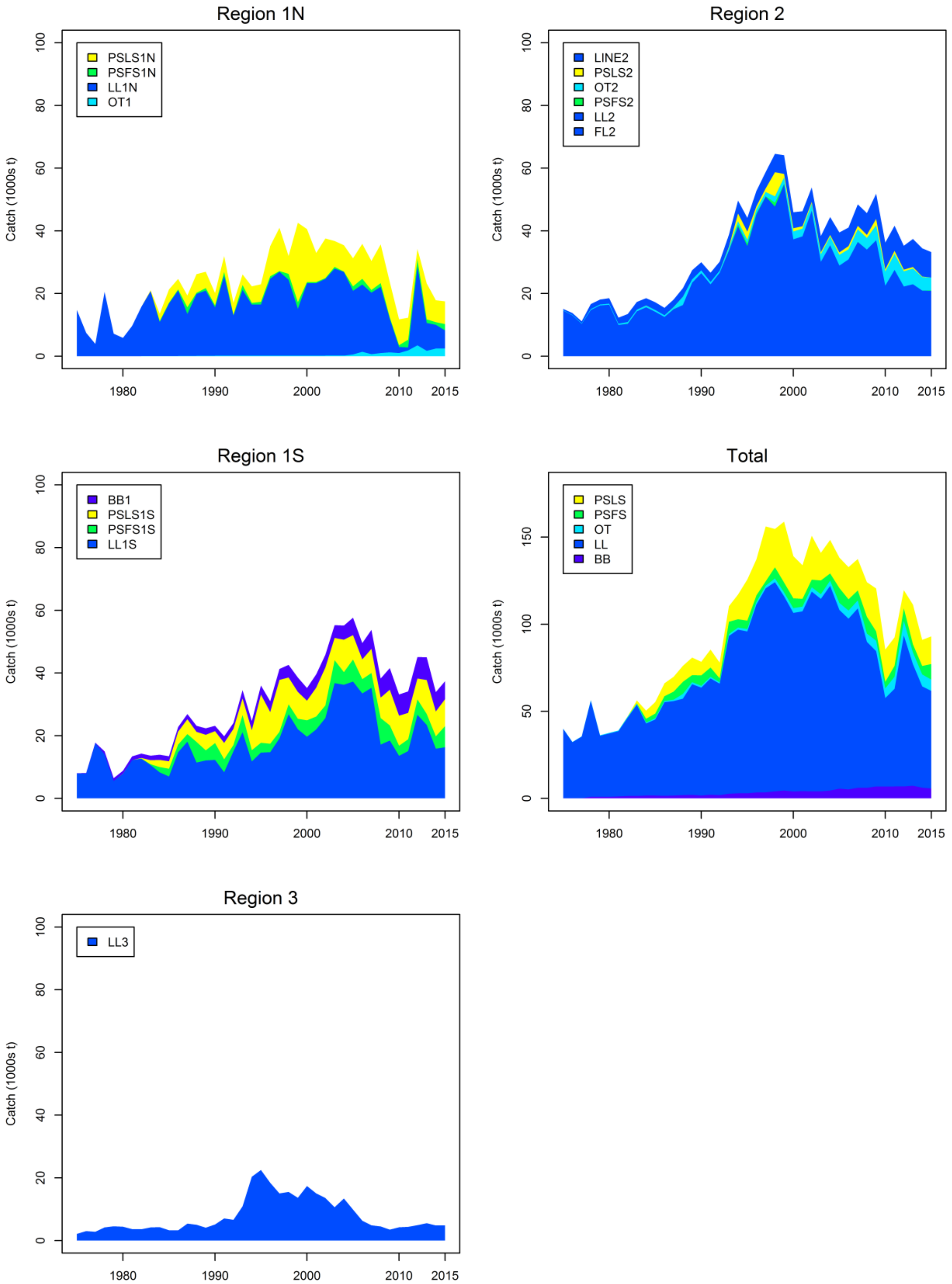

Figure 3. Annual catches by fishery and region 1975-2015.

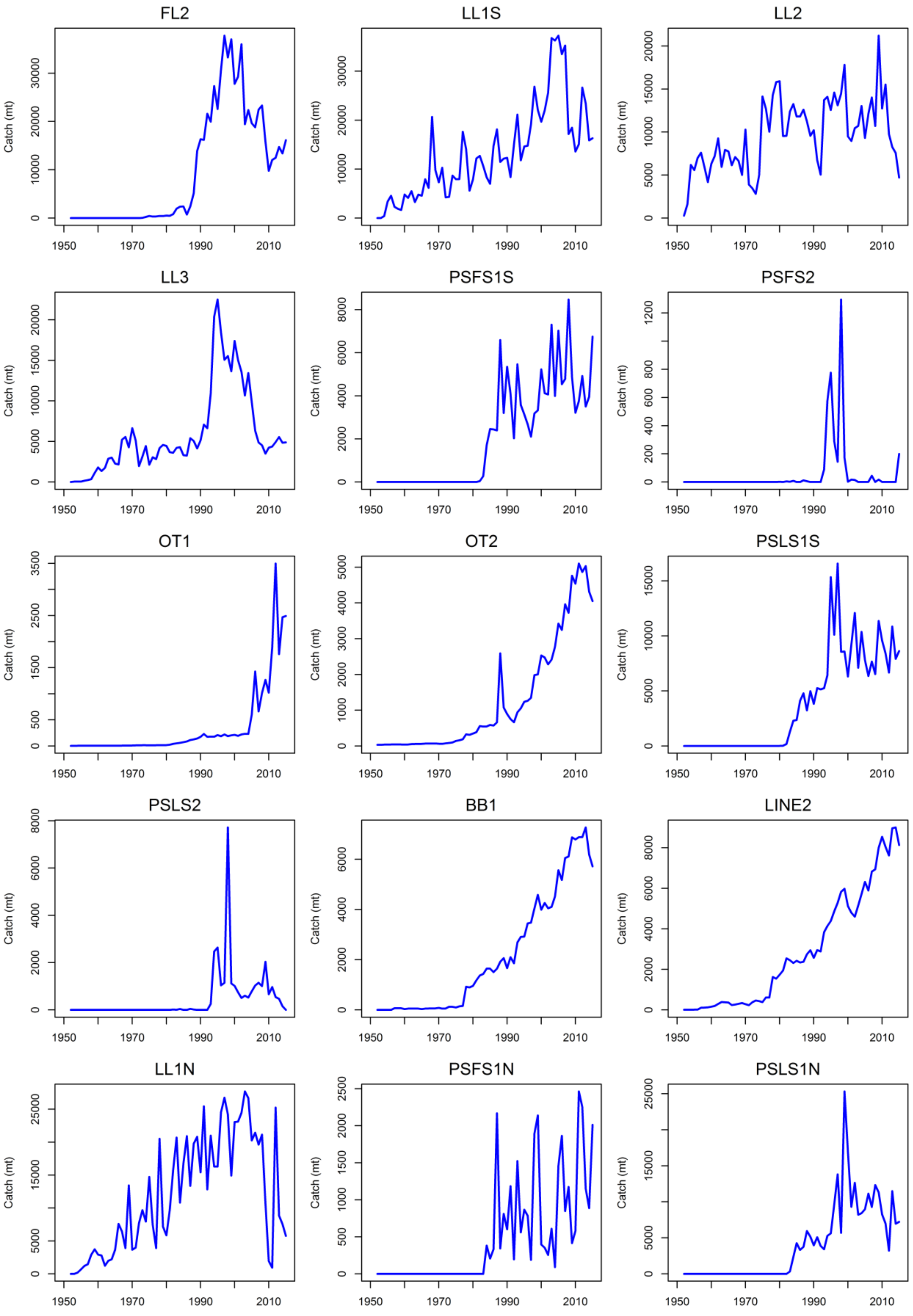

Figure 4. Annualised catches by fishery, 1950-2015.

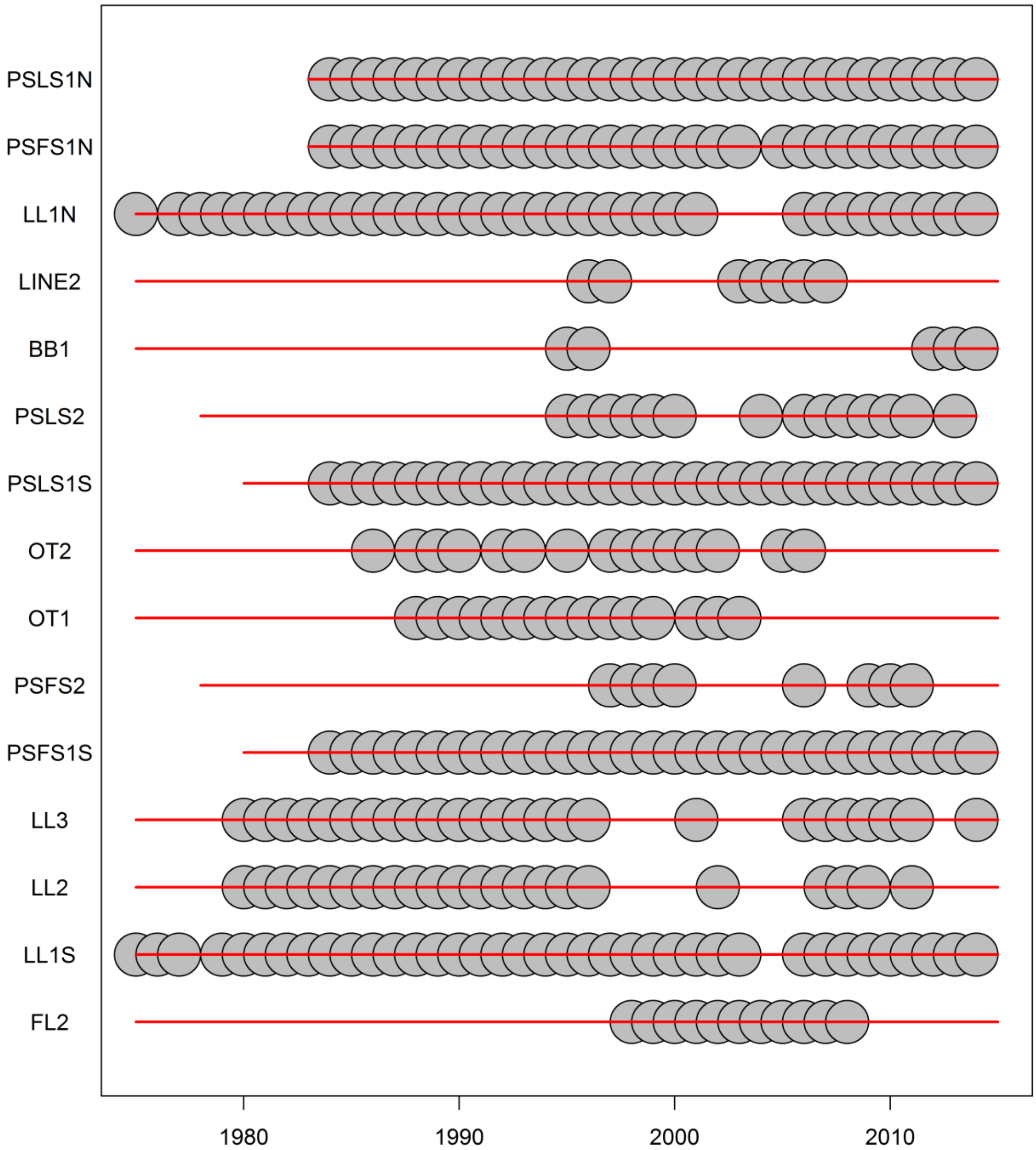

Figure 5. The annual occurrence of length samples (circles) from each fishery in the SS data set. The red line spans the time period of catch from the individual fishery.

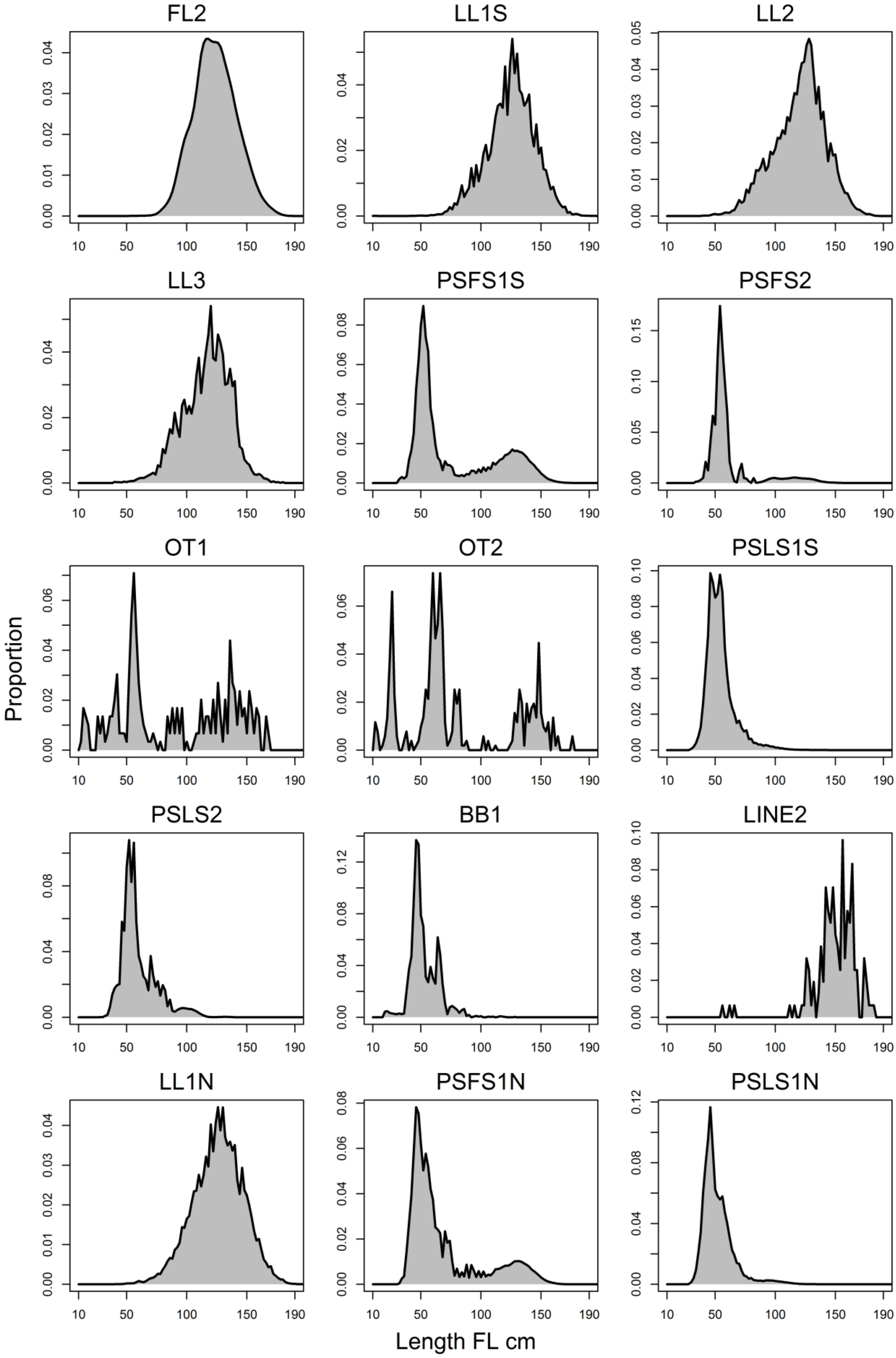

Figure 6. Length compositions of bigeye tuna samples aggregated by fishery.

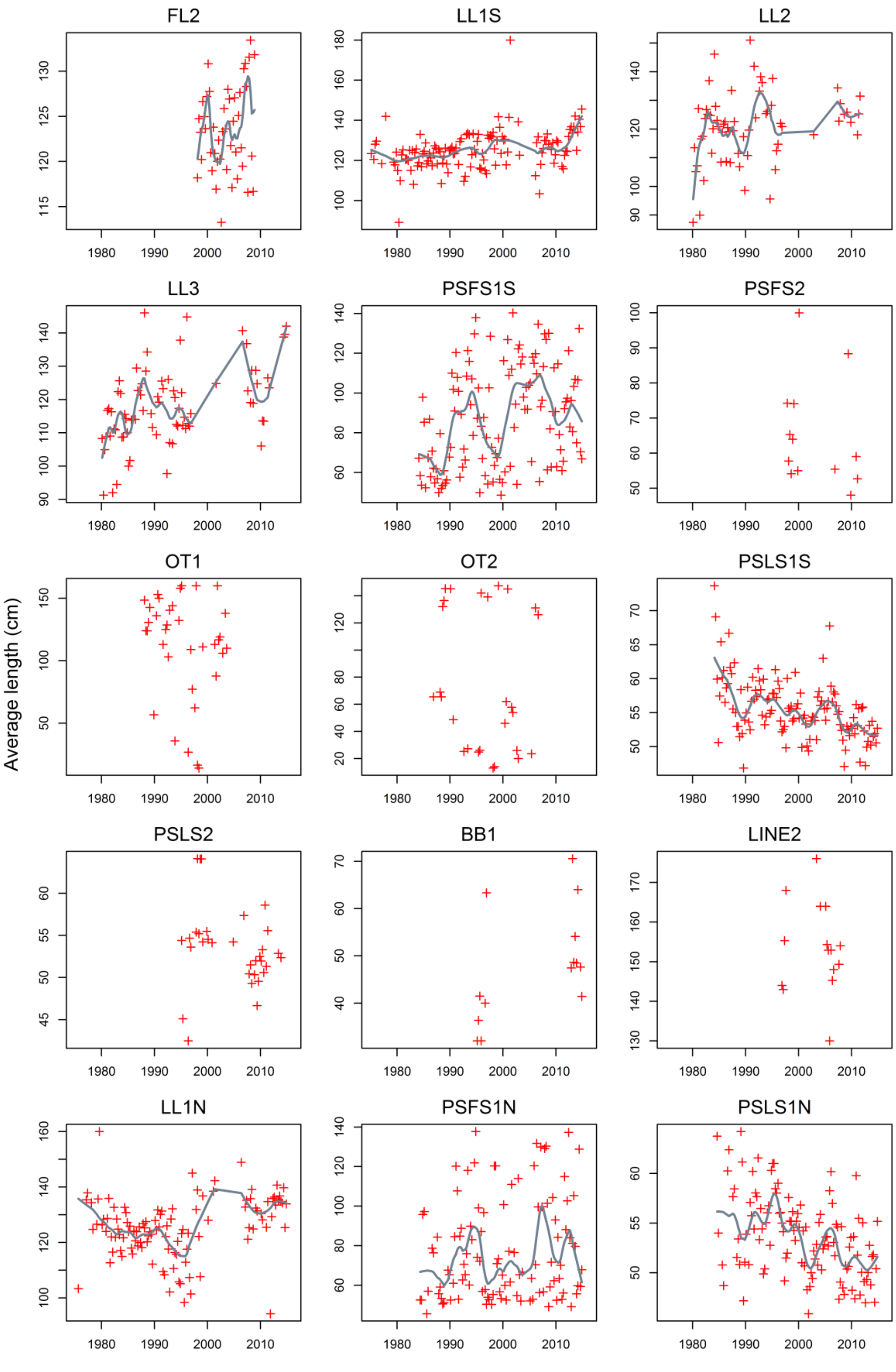

Figure 7. The average length (fork length, cm) of bigeye in the individual samples from each fishery. The grey line represents a lowess smoothed trend. The y-axis differs among the individual plots.

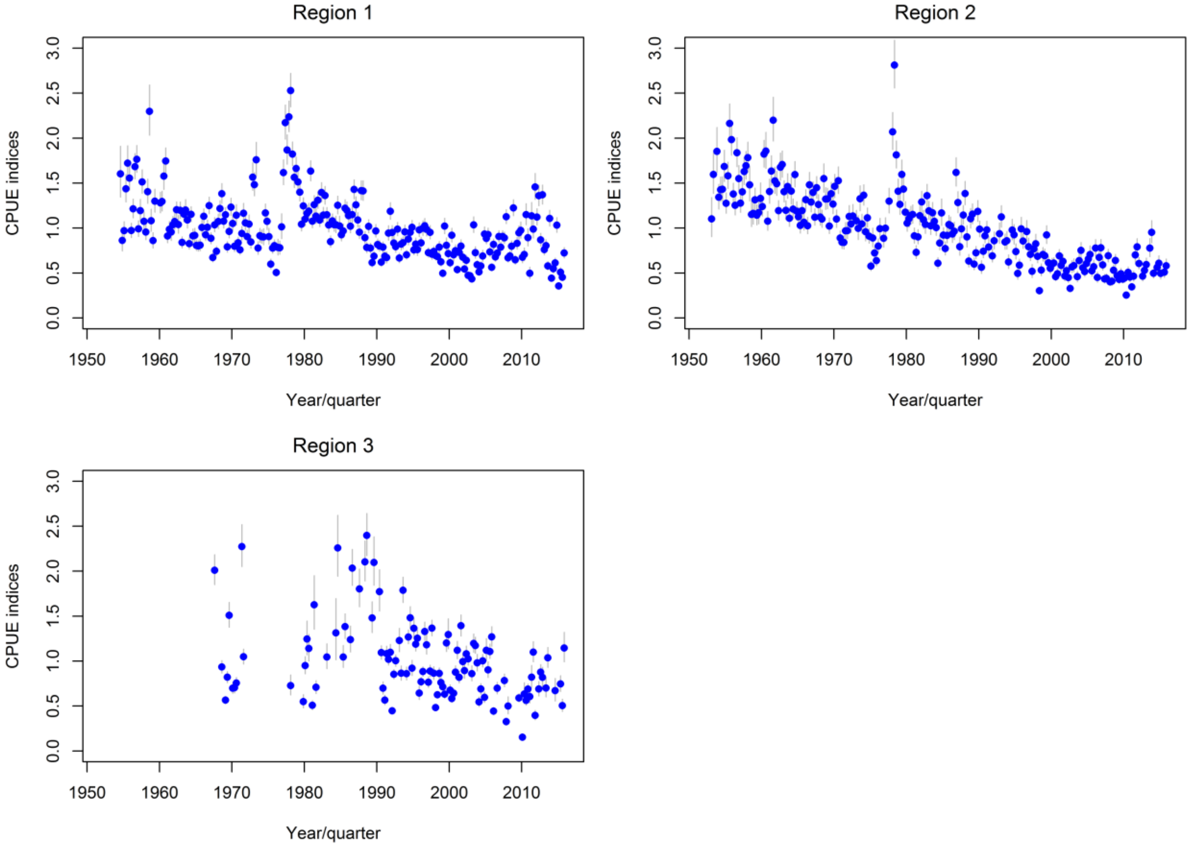

Figure 8. The entire time series of longline CPUE indices for each region from Hoyle et al (2016).

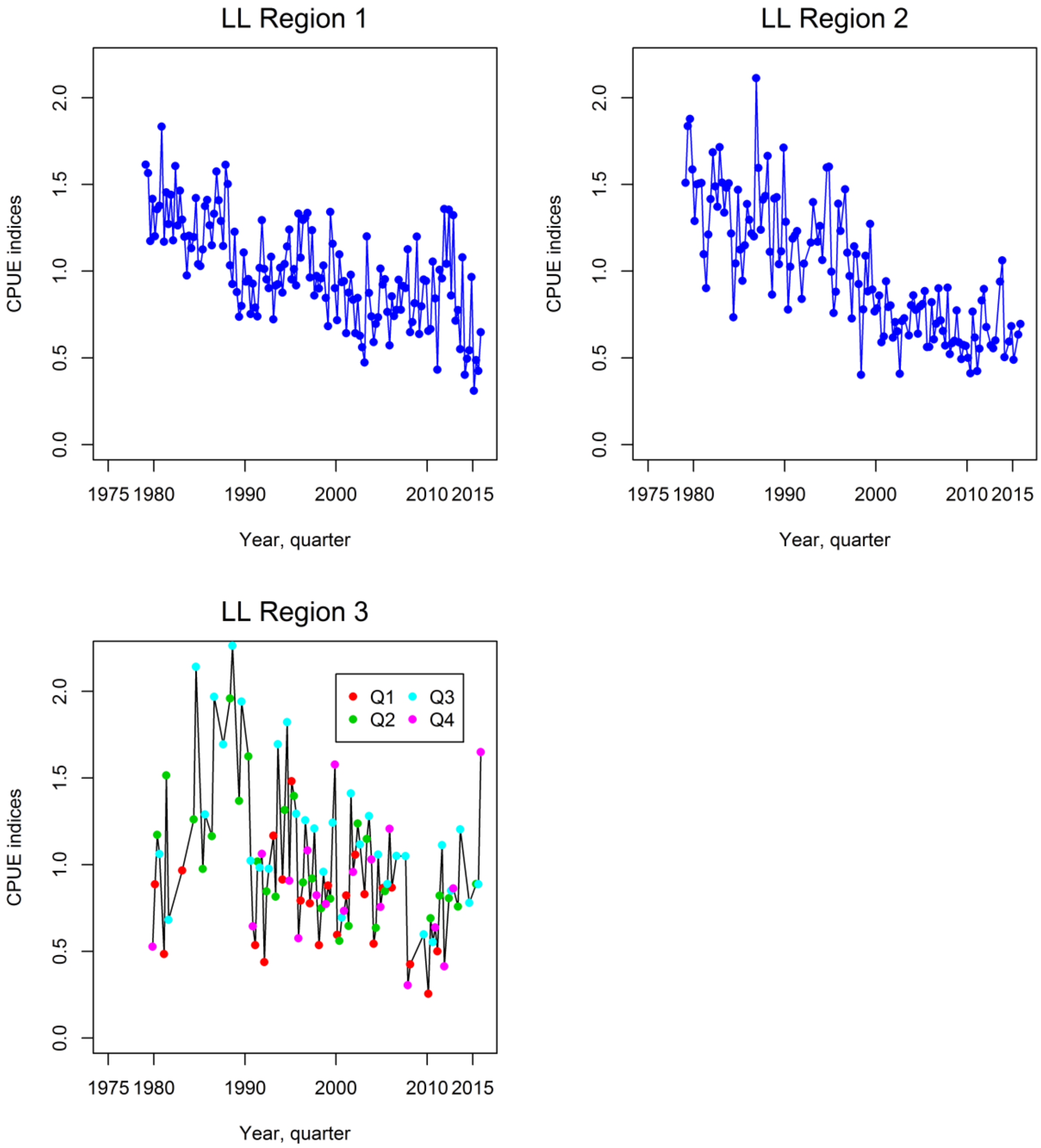

Figure 9. Standardised quarterly longline CPUE indices by region included in the stock assessment models.

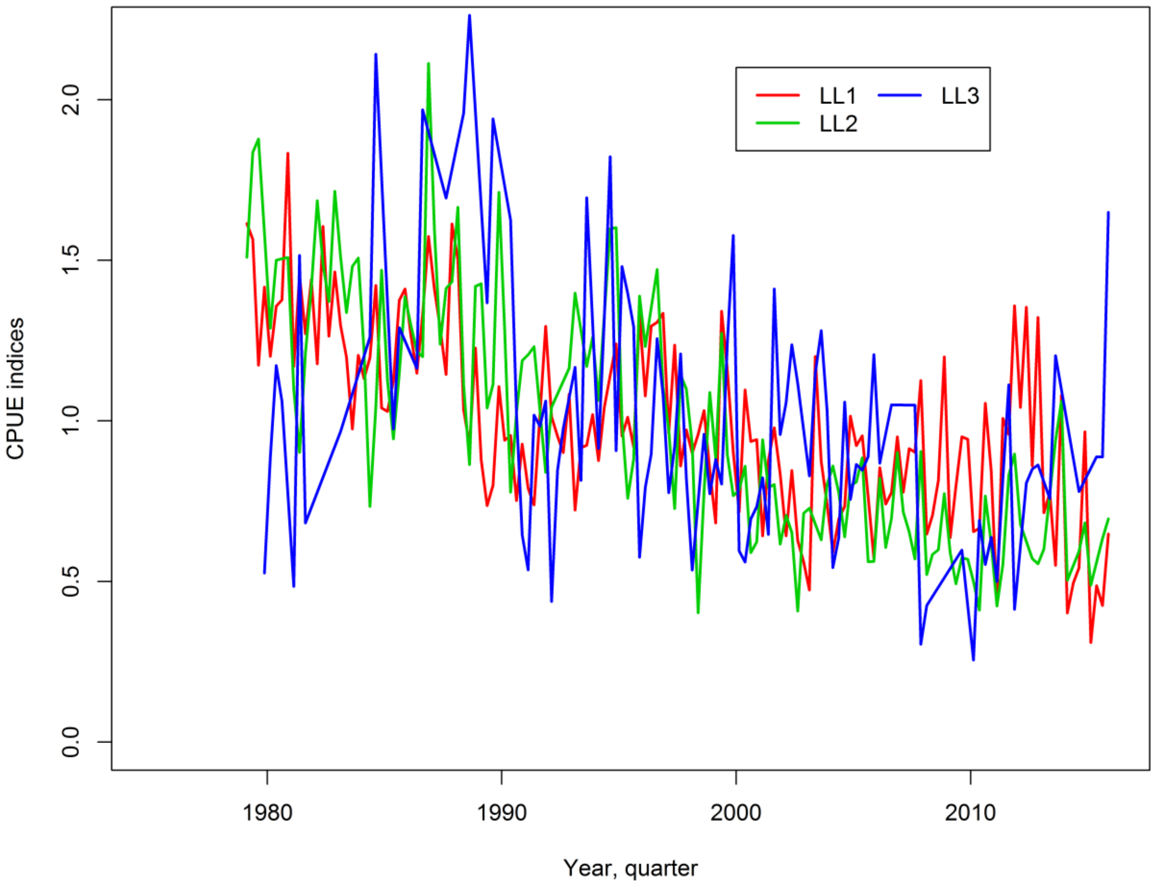

Figure 10. A comparison of the quarterly longline CPUE indices by region included in the stock assessment models.

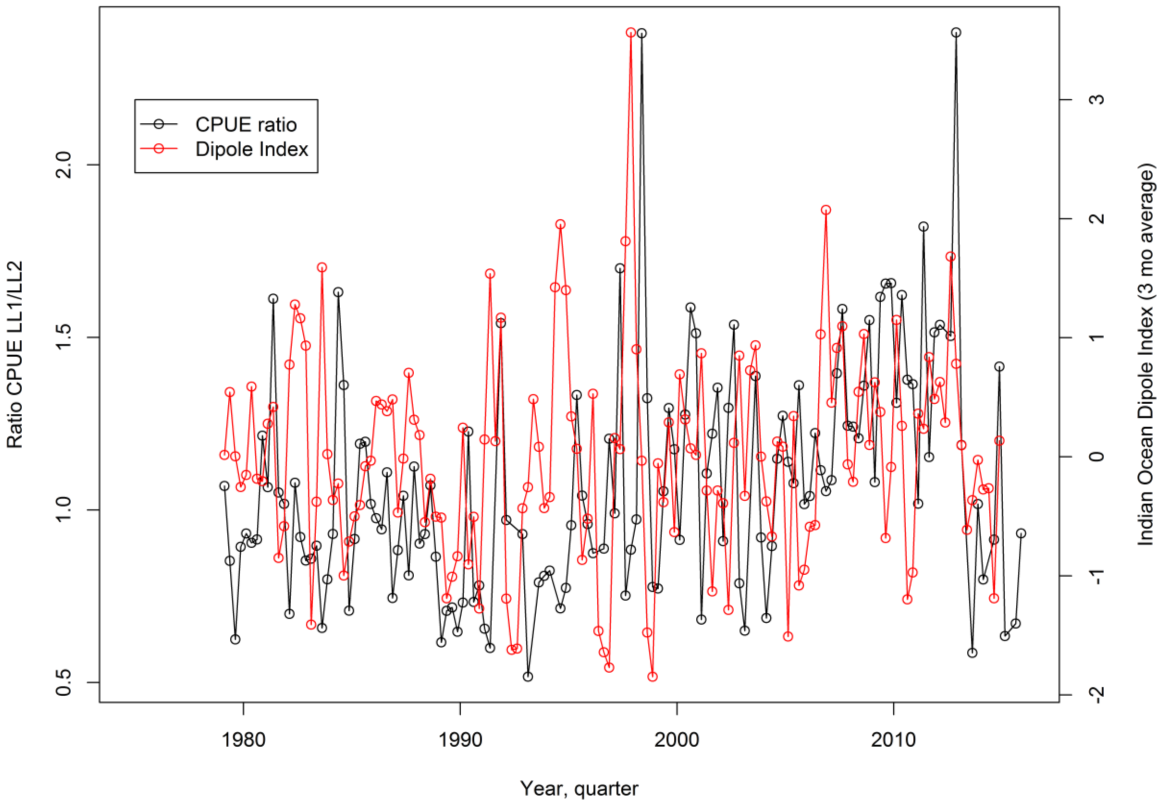

Figure 11. A comparison of the relative trends in the longline CPUE indices from regions 1 and region 2 (expressed as the ratio R1/R2) and the quarterly average of the Indian Ocean Dipole Index (IODI) from 1979 to 2014 (source: http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.ERSST/.version4/.IOD/.C1979-2014/.iod/datatables.html).

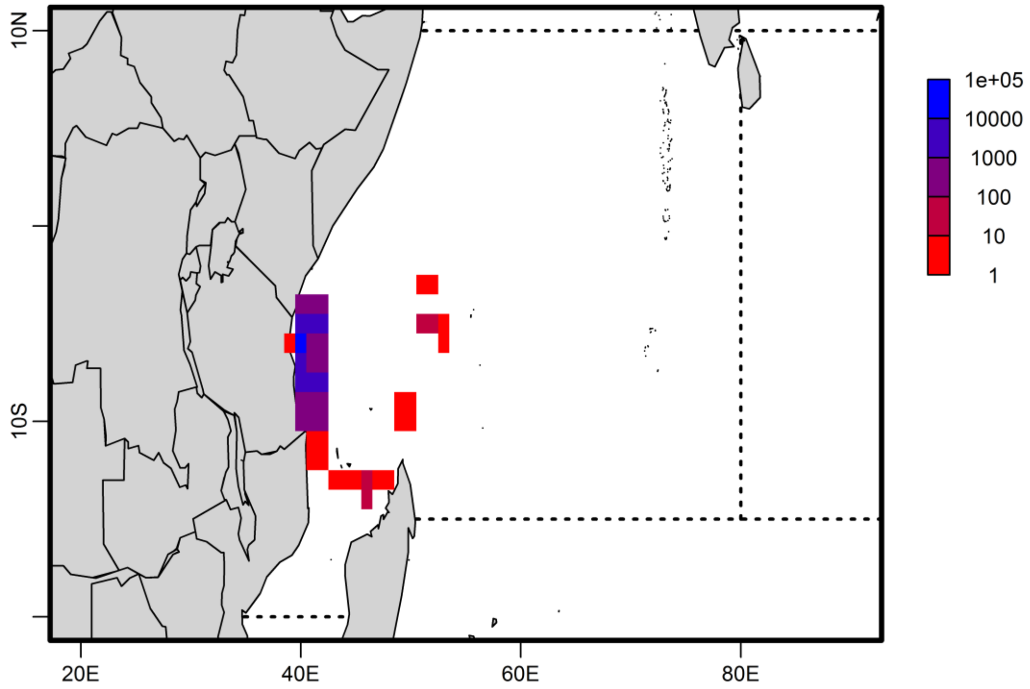

Figure 12. The total number of bigeye tuna tag releases during the RTTP by one degree of latitude/longitude.

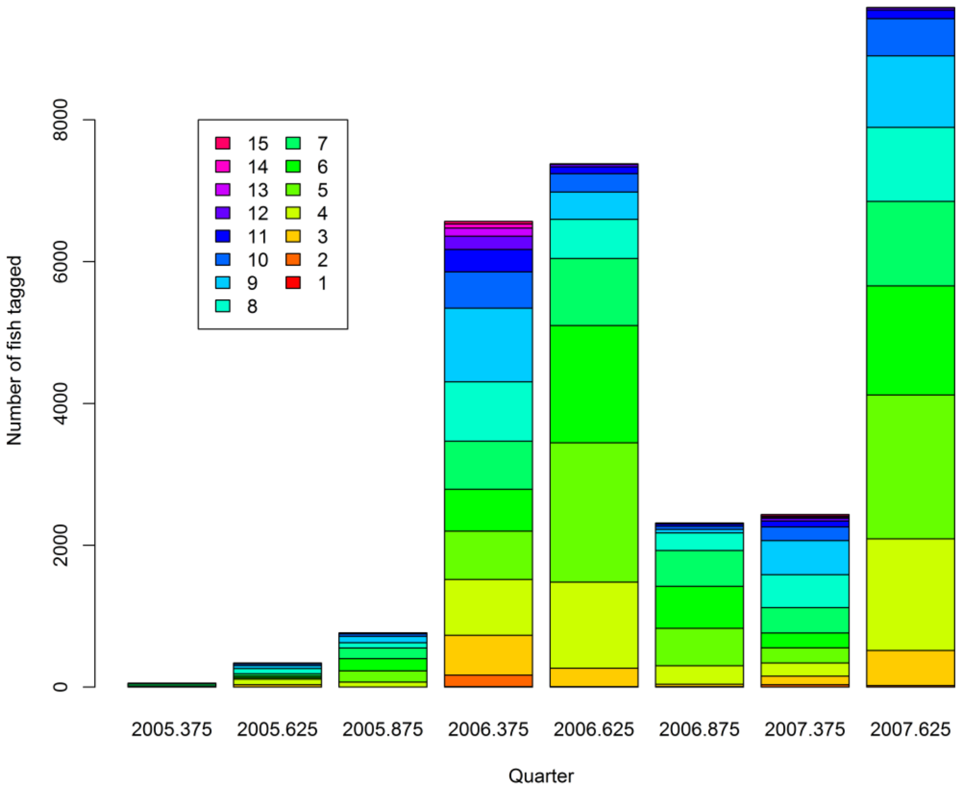

Figure 13. The total number of bigeye tuna tag releases by year/quarter and assumed age at release (adjusted release data set).

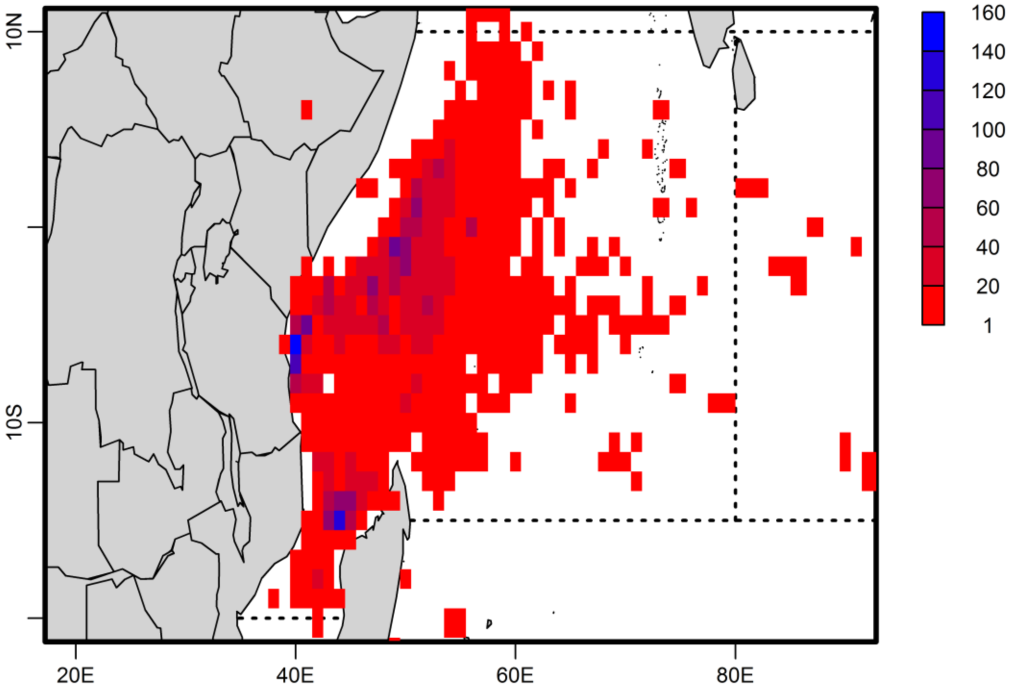

Figure 14. Total number of bigeye tag recoveries by location (one degree latitude/longitude cell).

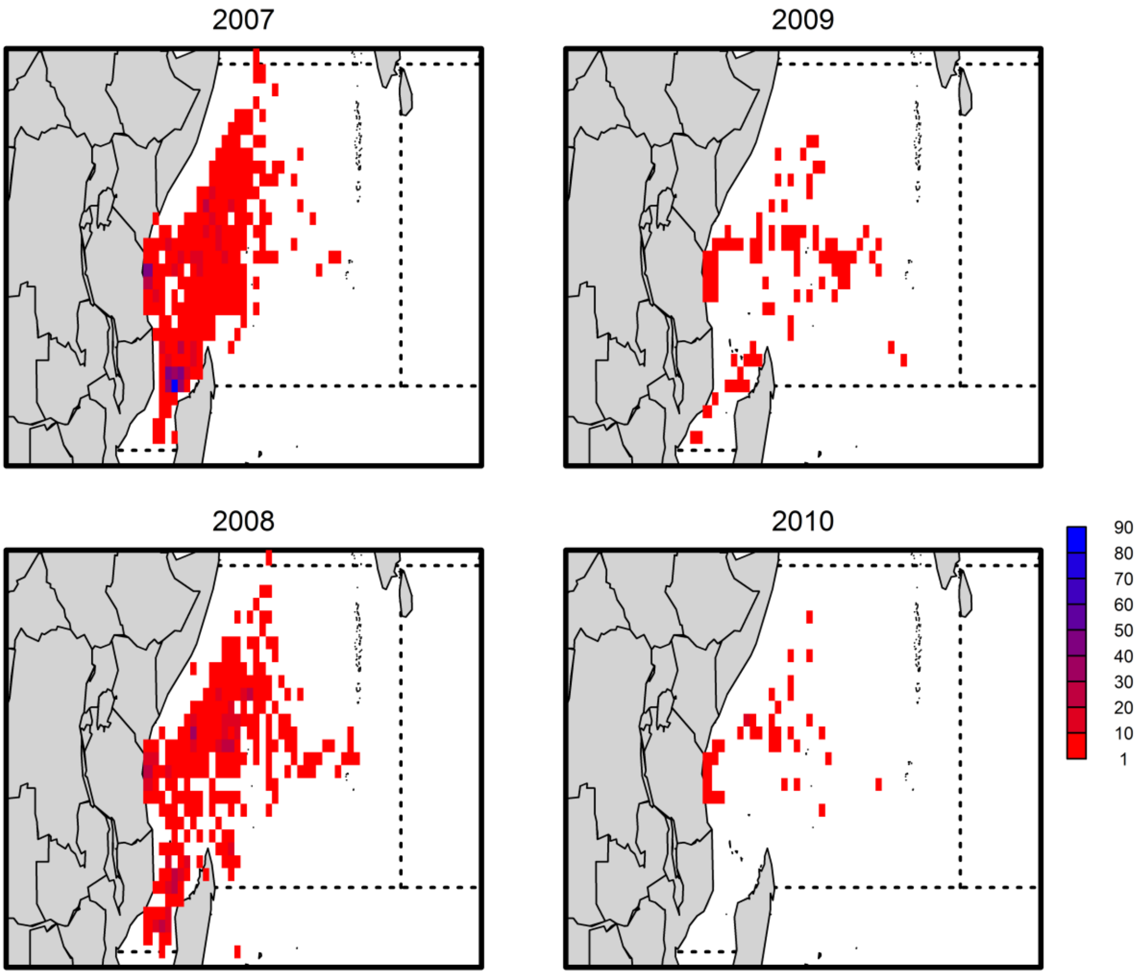

Figure 15. Number of bigeye tuna RTTP tag recoveries of fish at liberty for at least 6 months from the purse seine fishery by location (1 degree cell) and year (main recovery period).

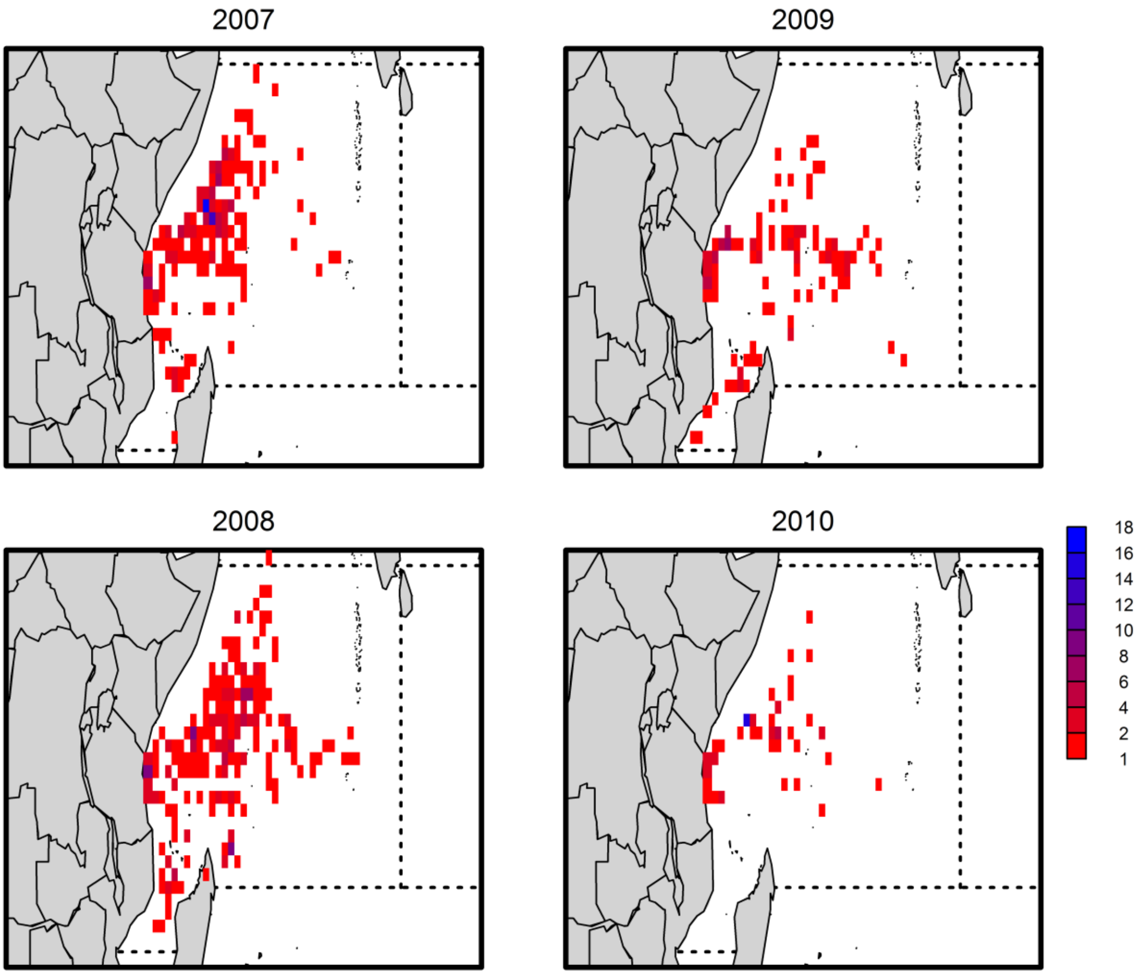

Figure 16. Number of bigeye tuna RTTP tag recoveries of fish at liberty for at least 12 months from the purse seine fishery by location (1 degree cell) and year (main recovery period).

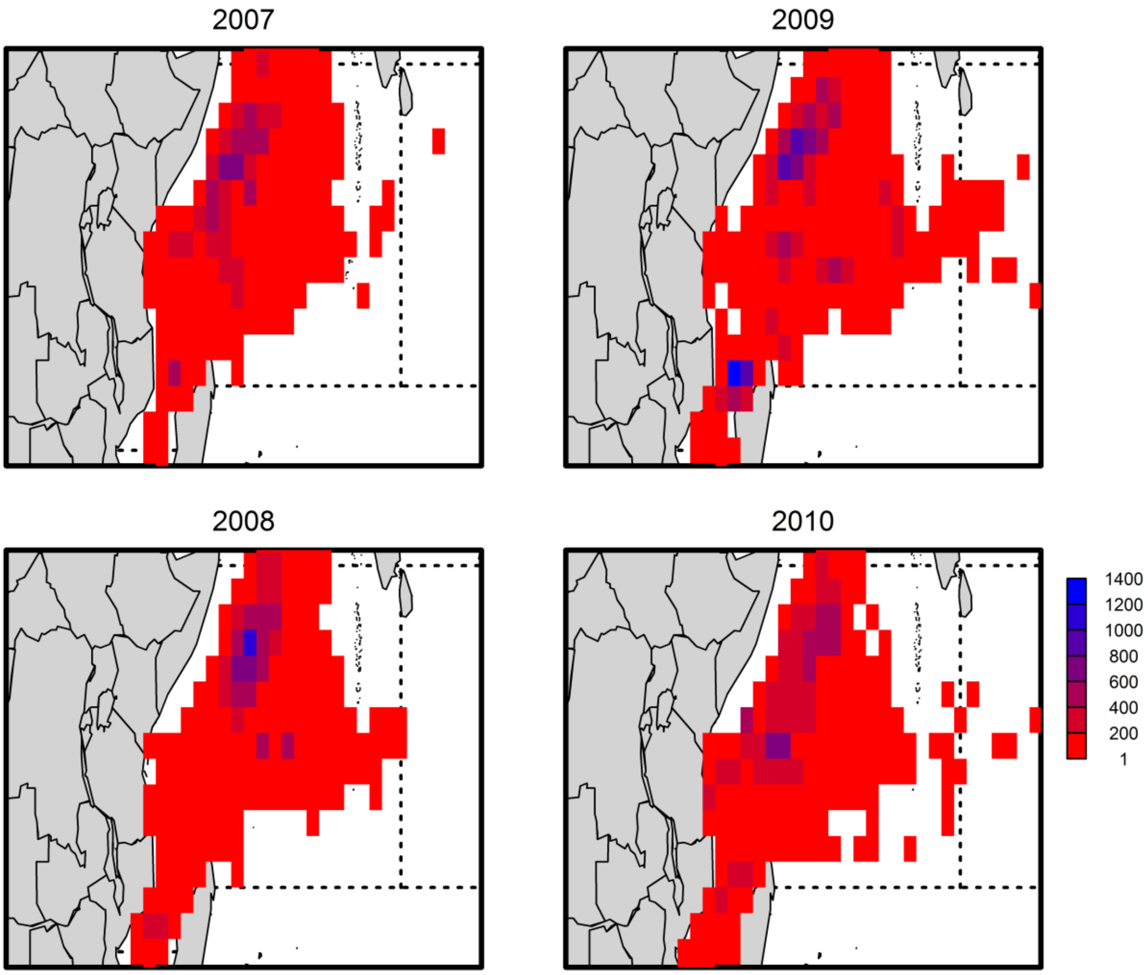

Figure 17. Bigeye tuna catch (mt) from the FAD purse seine fishery by 2 degree of latitude/longitude for the years that accounted for most of the RTTP tag recoveries.

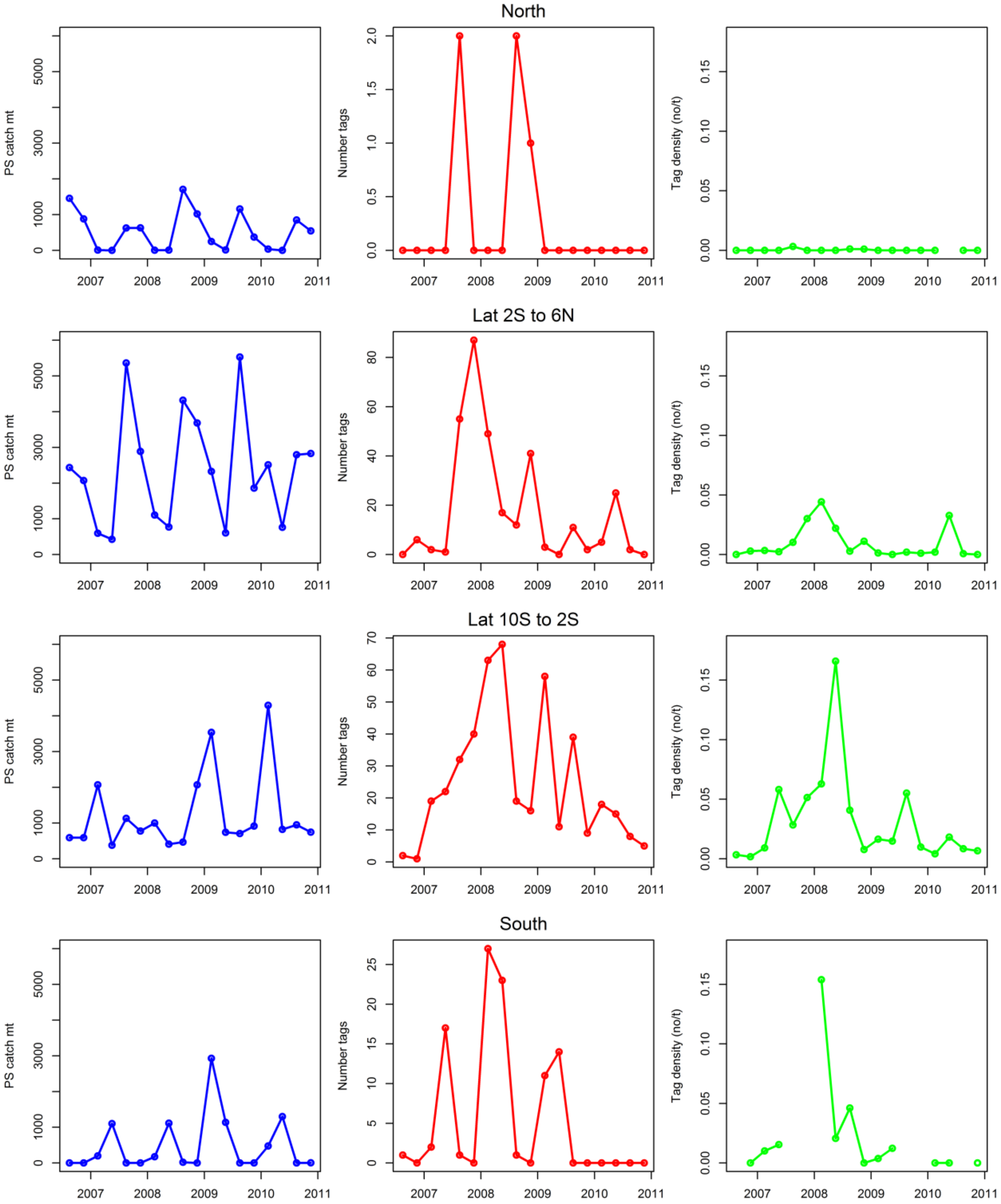

Figure 18. Quarterly bigeye tuna catch (t) and number of tags recovered by the purse seine fishery (region 1) by latitudinal band. Only tags at liberty for at least 12 month mixing period are included. The tag recovery density (tags/catch) is also presented for each latitudinal band.

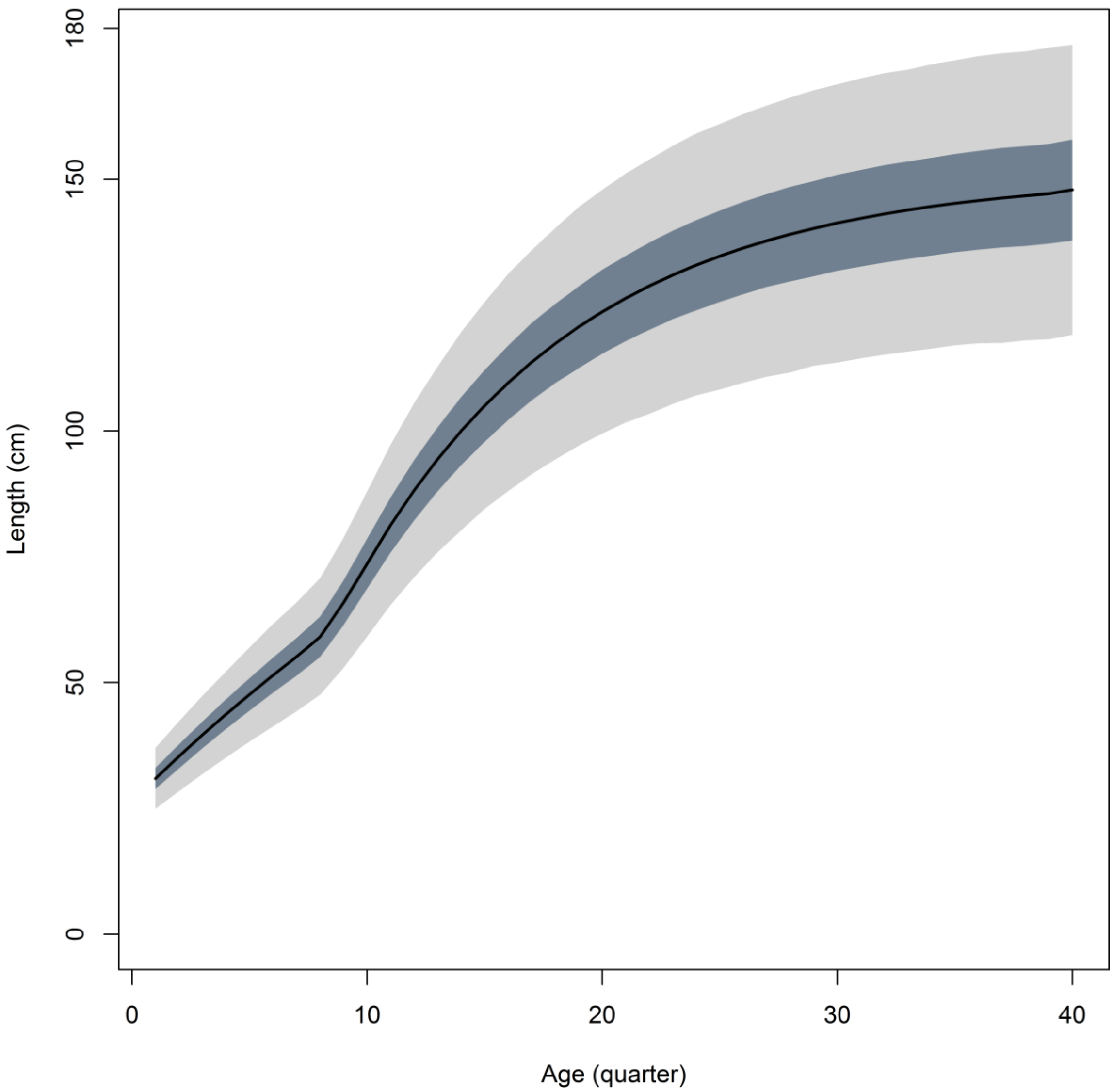

Figure 19. Growth of Indian Ocean bigeye tuna (following Everson et al 2012). The dark grey region represents the quartile range of the distribution of length-at-age and the light grey region represents the 95% confidence interval.

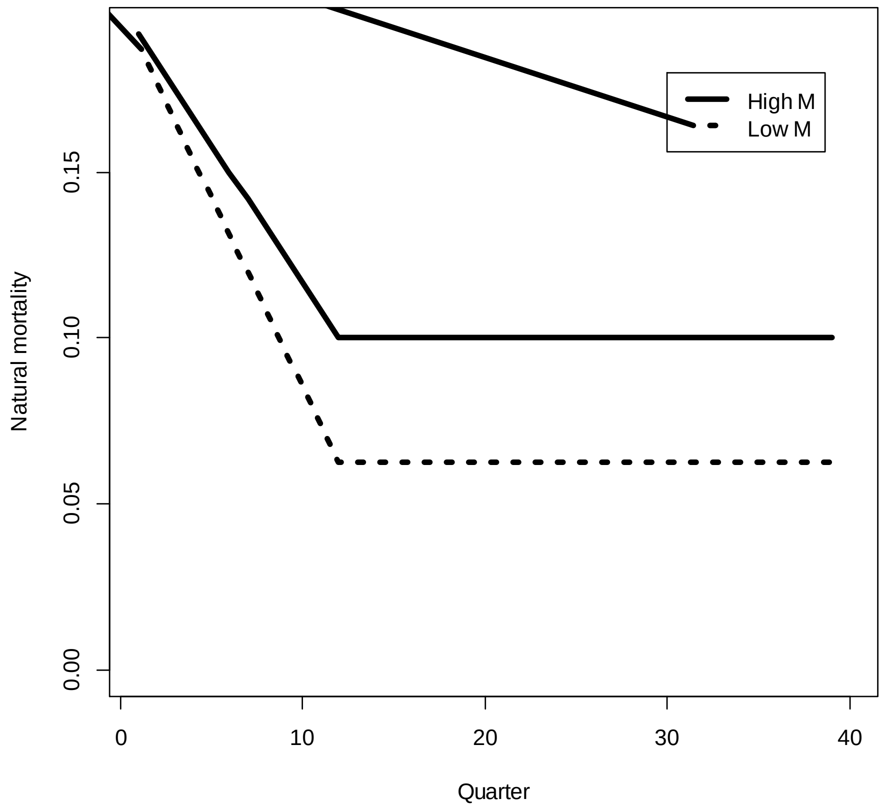

Figure 20. Age specific natural mortality (per quarter) patterns assumed for the highM and lowM assessment model options.

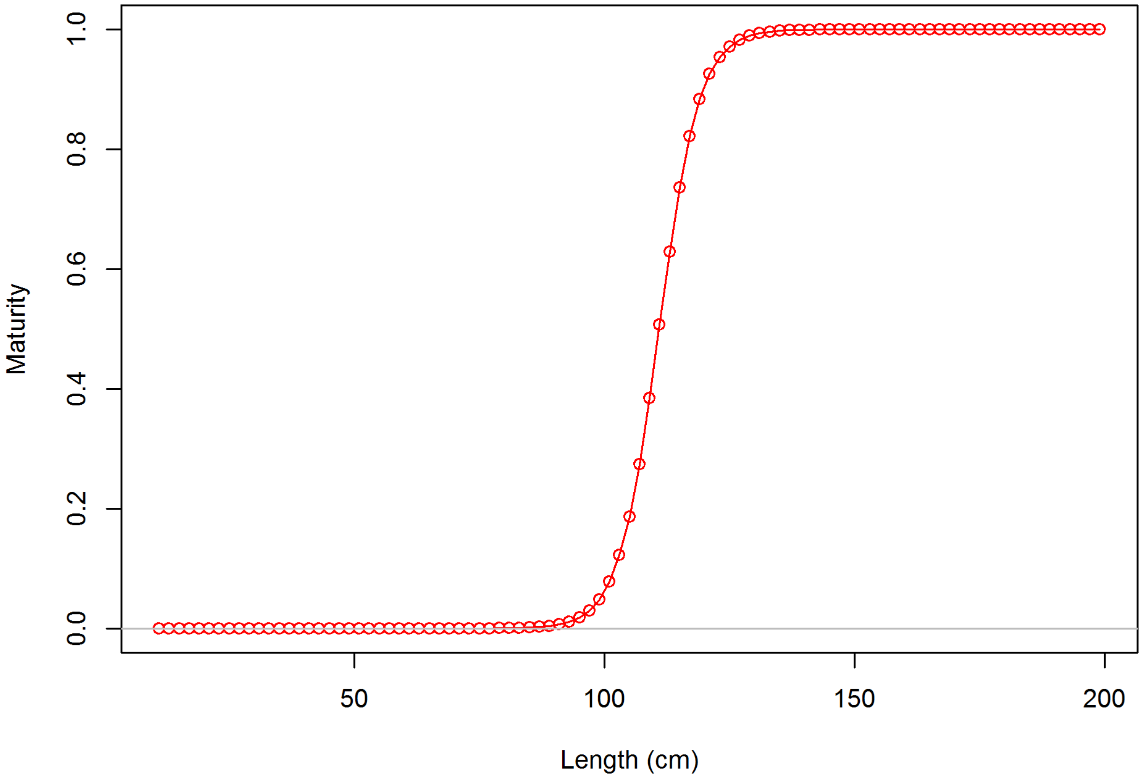

Figure 21. Proportion of fish mature by length class.

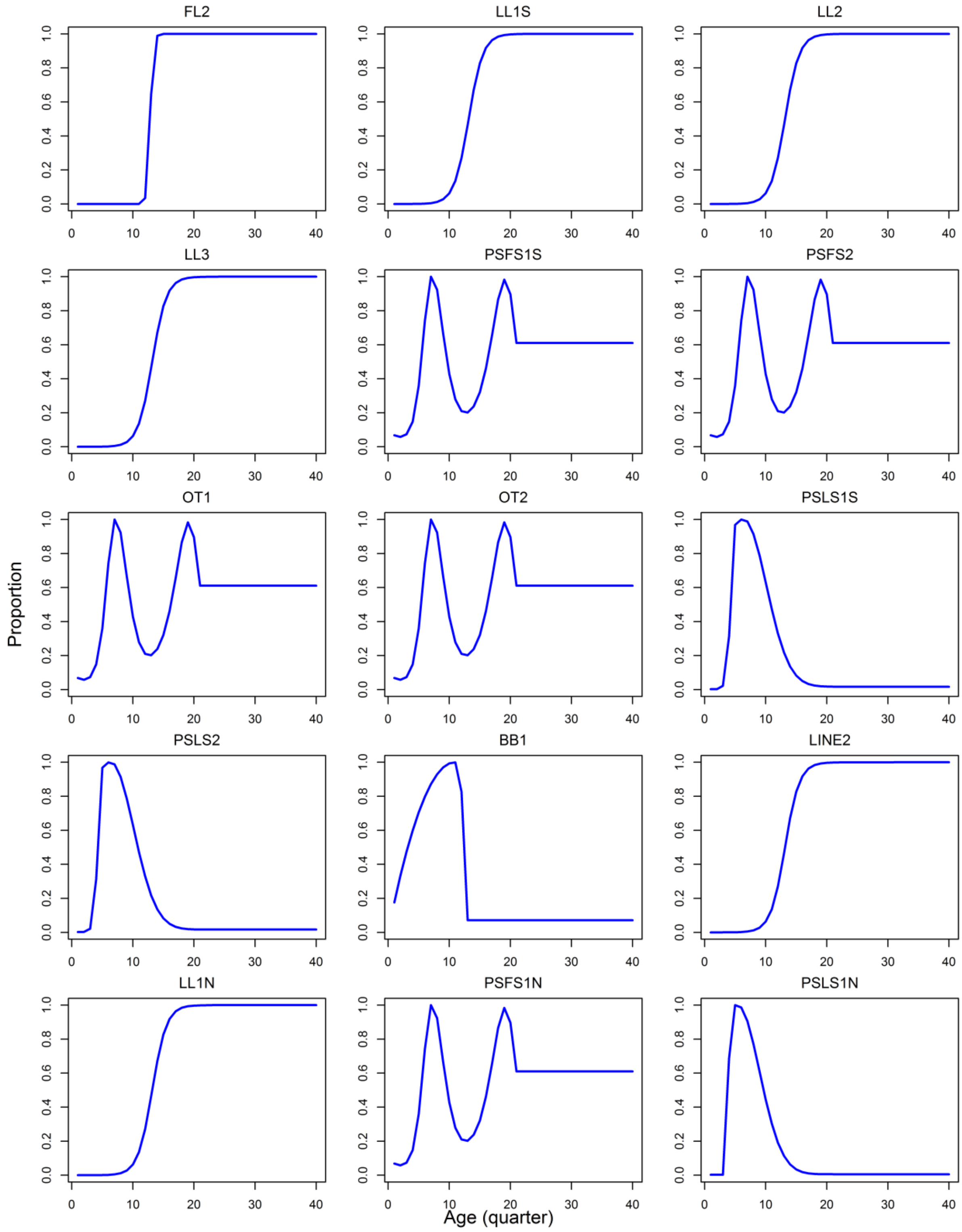

Figure 22. Selectivity for the individual fisheries for the reference model.

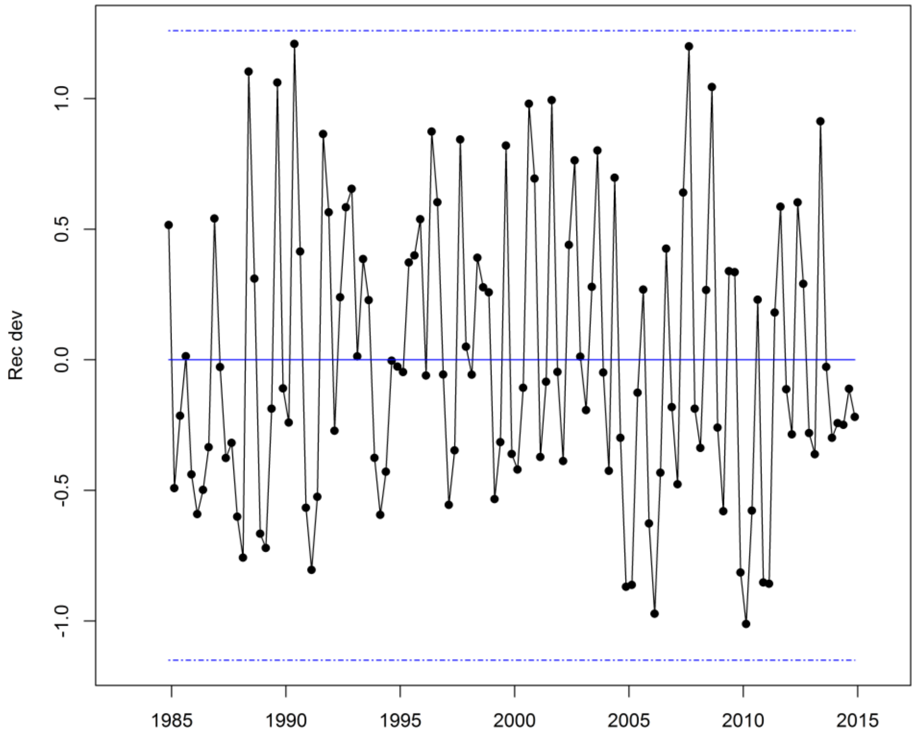

Figure 23. Quarterly recruitment deviates for the reference model option (steepness 0.8). The blue horizontal lines represent +/- 1.96 standard deviation (SigmaR =0.6).

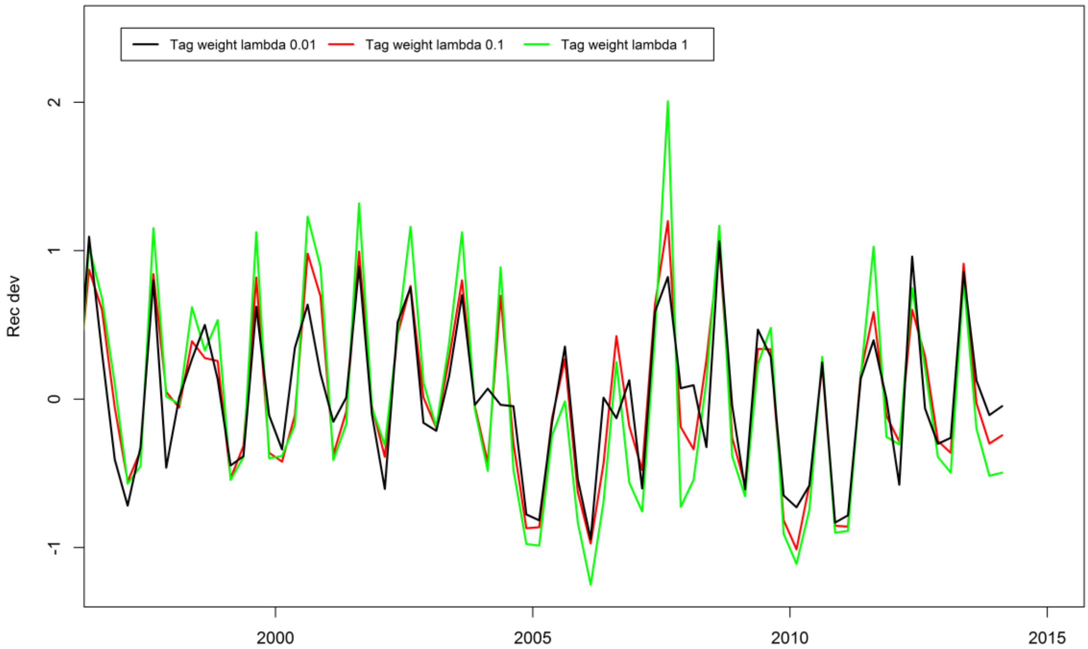

Figure 24. A comparison of the recruitment deviates estimated for the three model options with different weightings of the tag likelihood components.

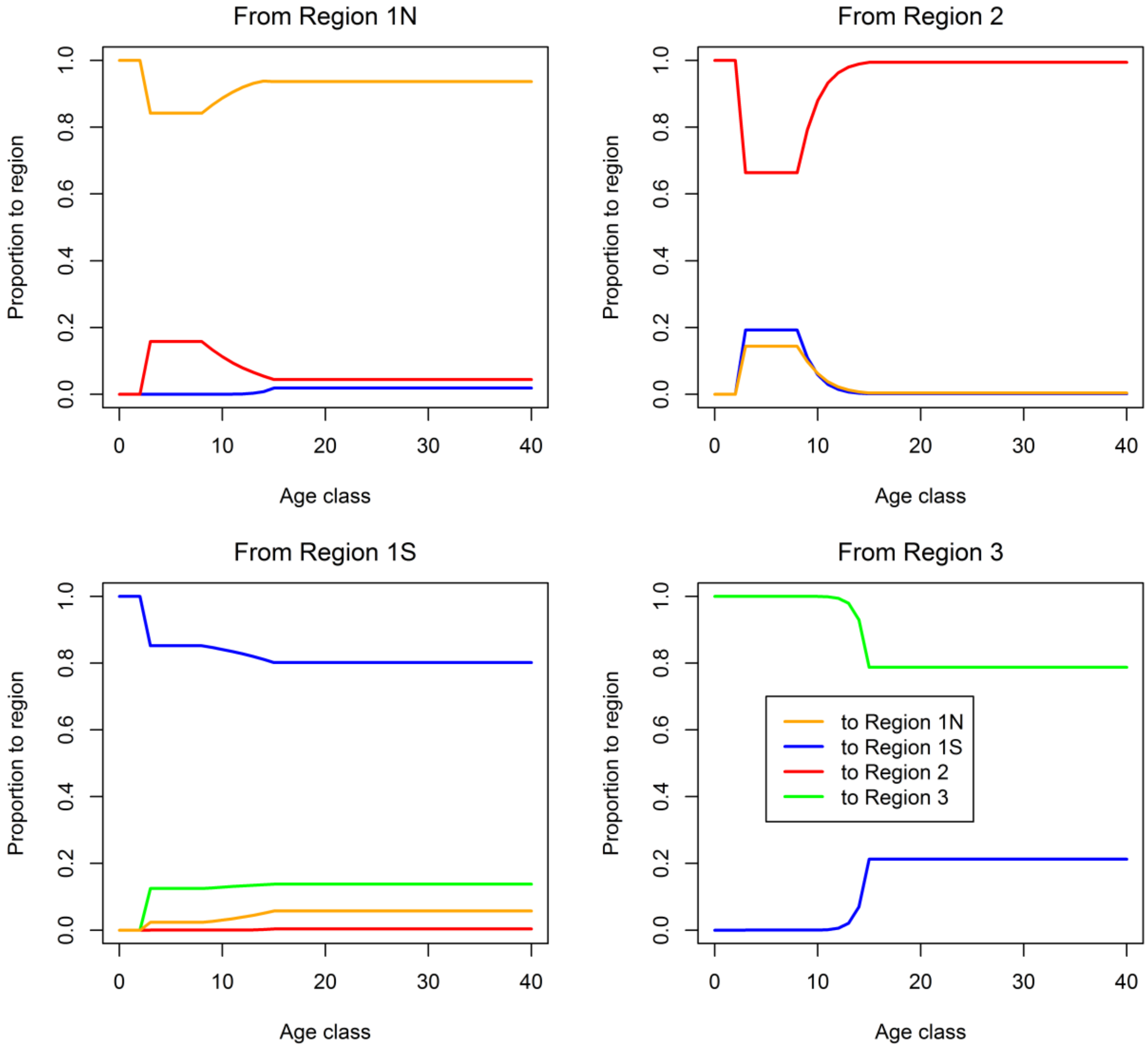

Figure 25. Estimated age specific movement parameters.

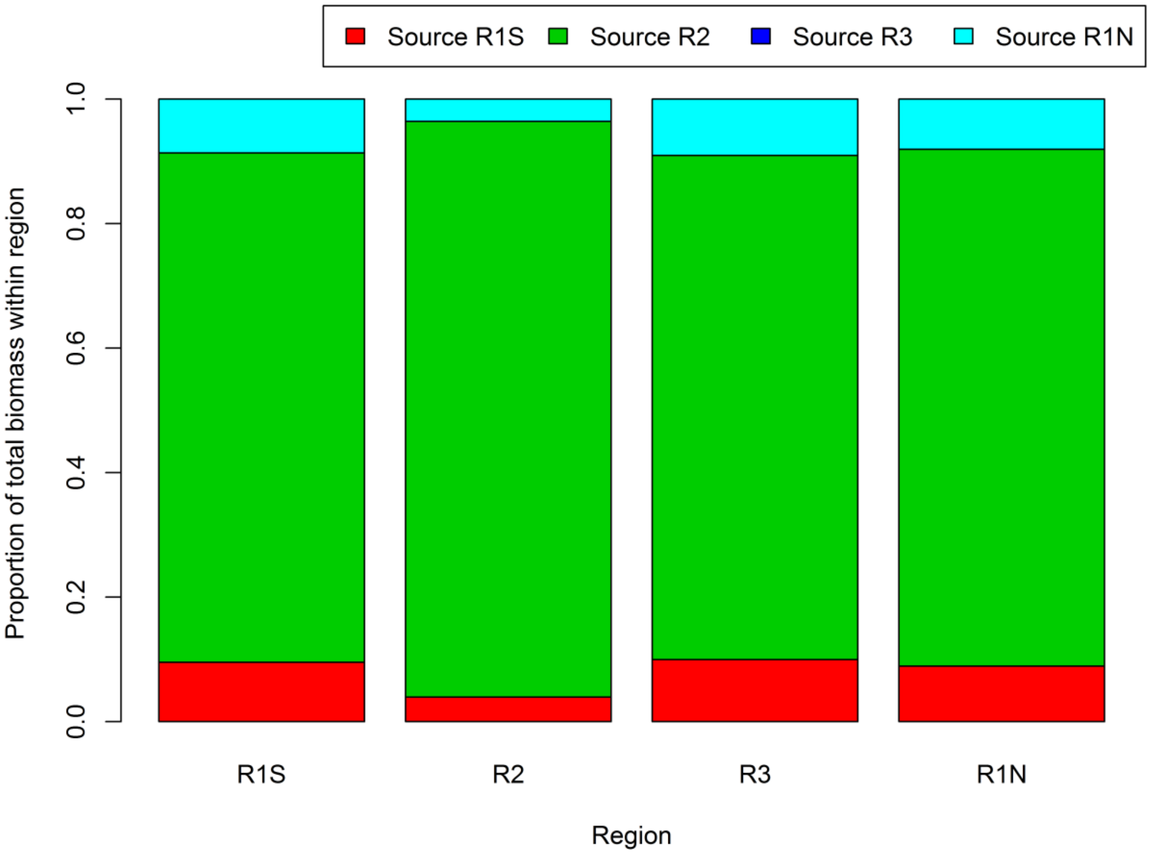

Figure 26. Proportional distribution of total biomass in each region by the source region (natal region).

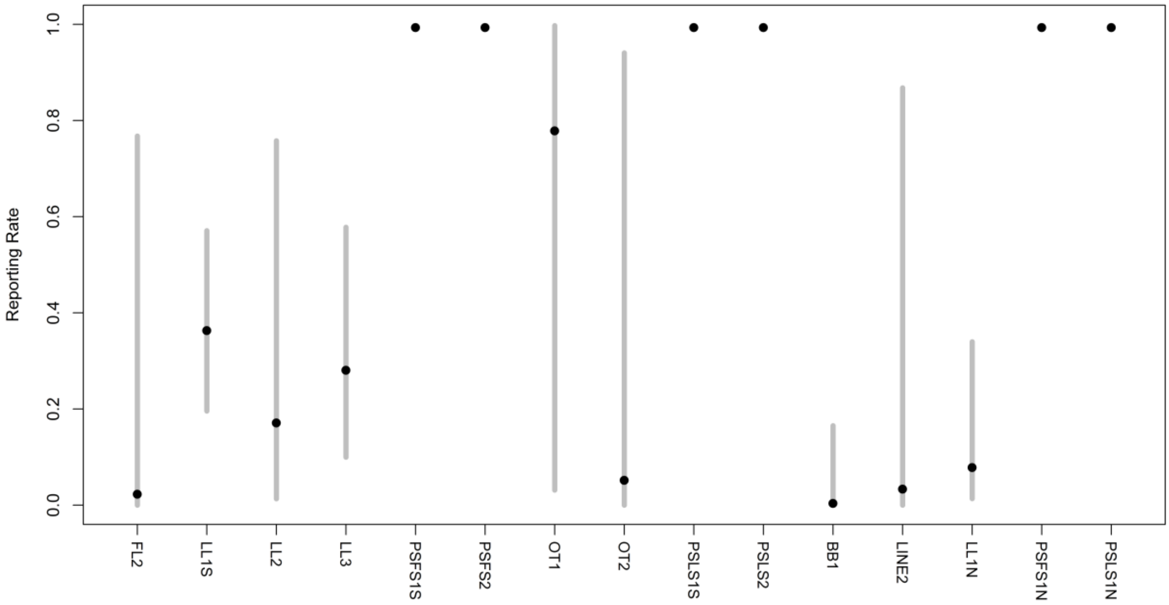

Figure 27. Tag reporting rates for each fishery. Purse seine reporting rates were fixed at a value of 1.0. Reporting rates for the other fisheries were estimated. The grey lines represent the 95% confidence interval for the estimated values.

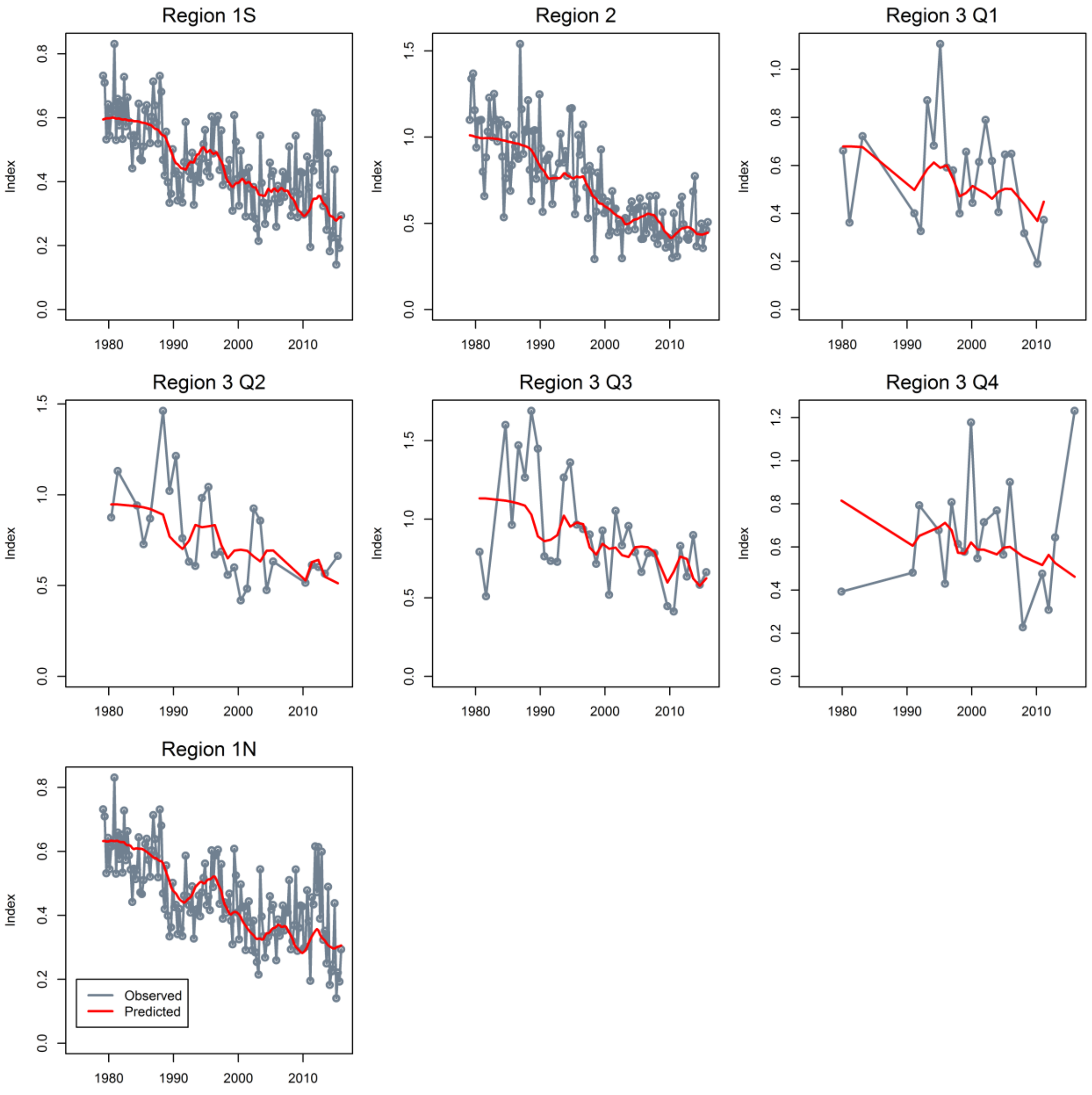

Figure 28. The fit to the region specific CPUE indices for the reference model. The CPUE indices for region 3 are fitted as separate seasonal sets of indices.

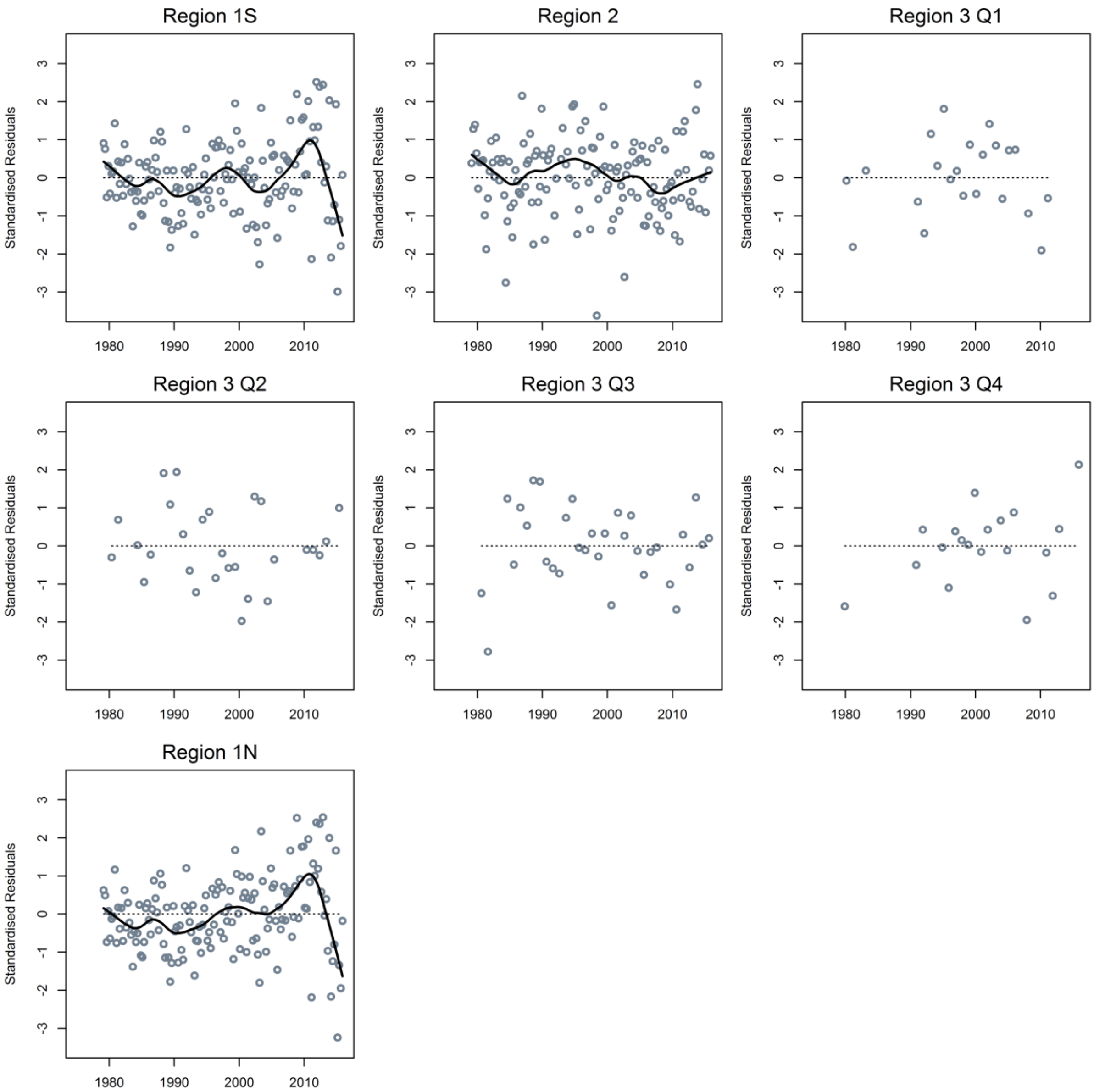

Figure 29. The residuals from the fit to the region specific CPUE indices for the reference model. The CPUE indices for region 3 are fitted as separate seasonal sets of indices. The black line represents a lowess smooth fit to the residuals (region 1N, 1S and 2 only).

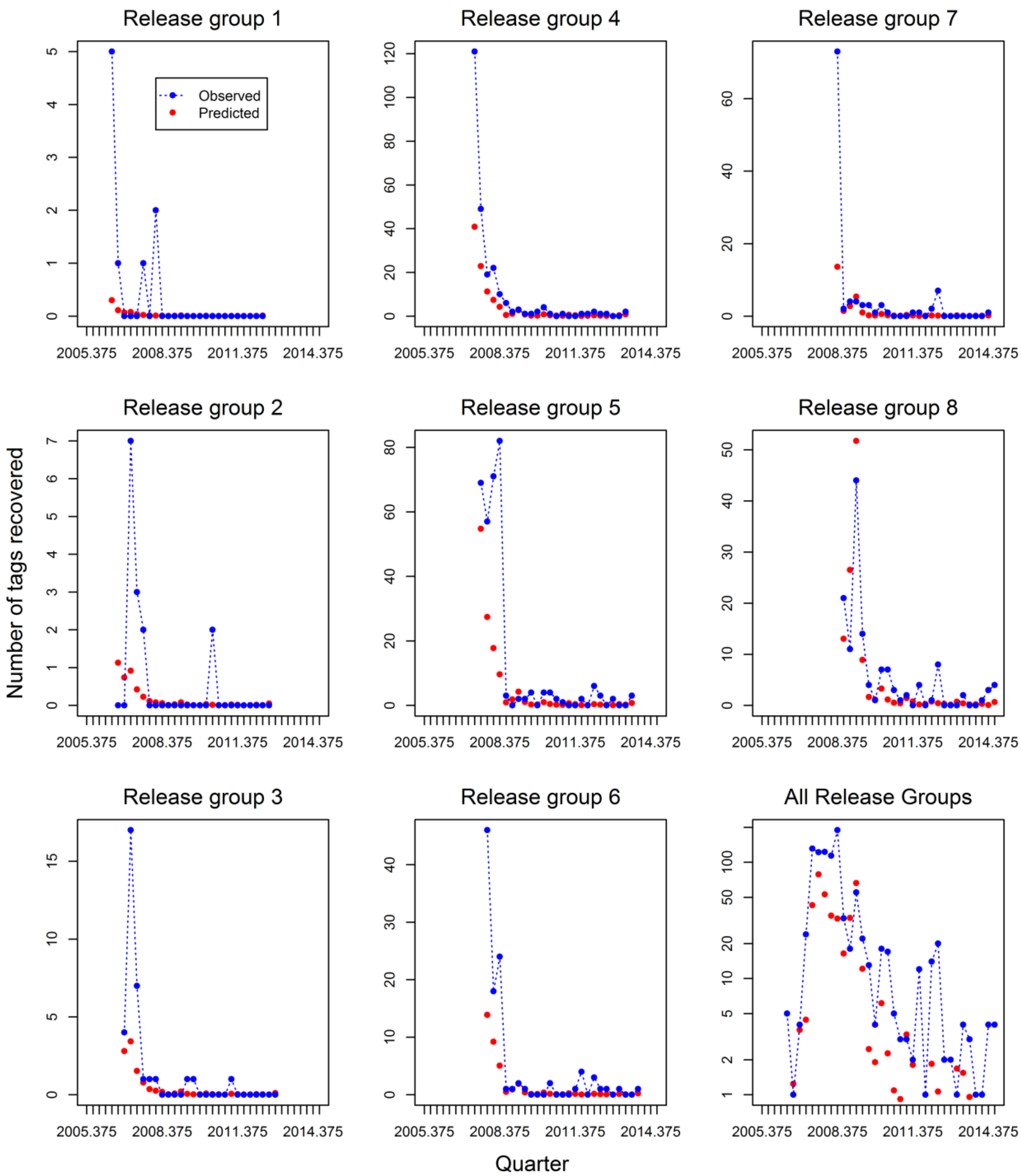

Figure 30. Observed and predicted the number of tag recoveries by the PSLS1S fishery by quarter following the mixing period (reference model). Tag release groups (1-8) represent the total releases in each quarter (aggregating the age groups that define individual SS release groups). The "All Release Groups" aggregates all recoveries by quarter (plotted on log scale).

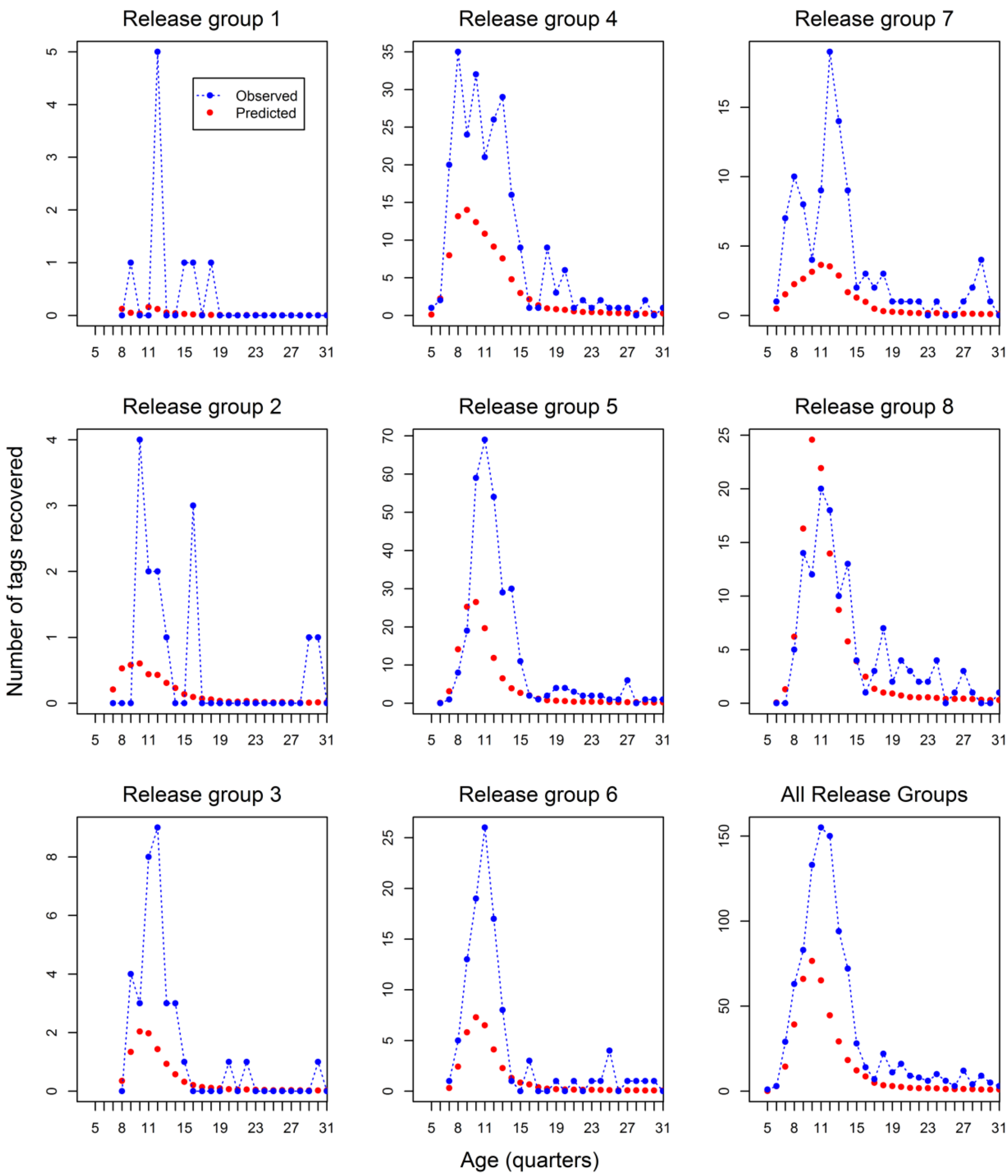

Figure 31. Observed and predicted the number of tag recoveries by the PSLS1S fishery by age (quarter) at recovery, excluding the recoveries during the mixing period (reference model). Tag release groups (1-8) represent the total releases in each quarter (combining the age groups that define individual SS release groups). The "All Release Groups" plot aggregates all recoveries by age class (at recovery).

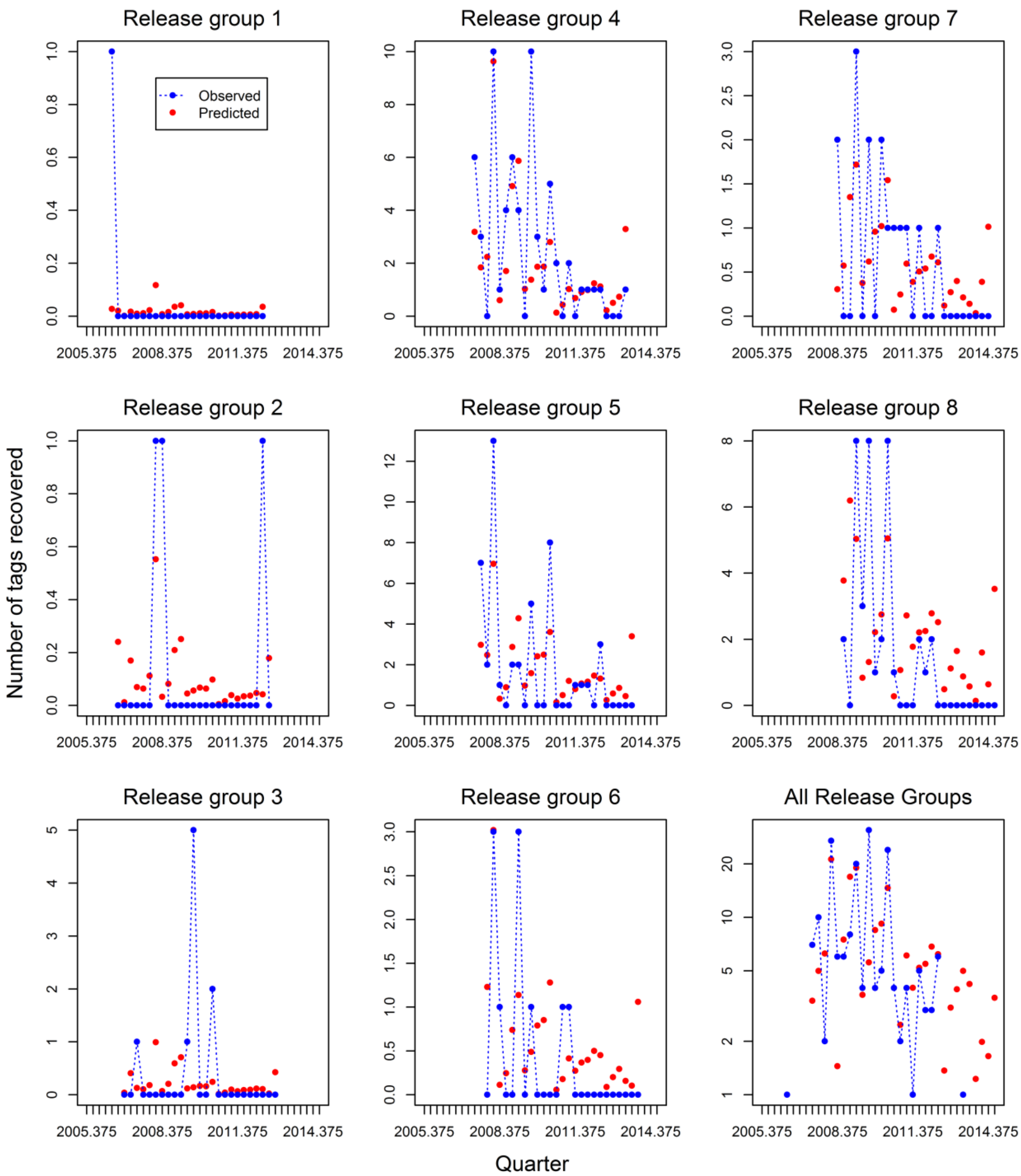

Figure 32. Observed and predicted the number of tag recoveries by the PSFS1S fishery by quarter following the mixing period (reference model). Tag release groups (1-8) represent the total releases in each quarter (aggregating the age groups that define individual SS release groups). The "All Release Groups" aggregates all recoveries by quarter (plotted on log scale).

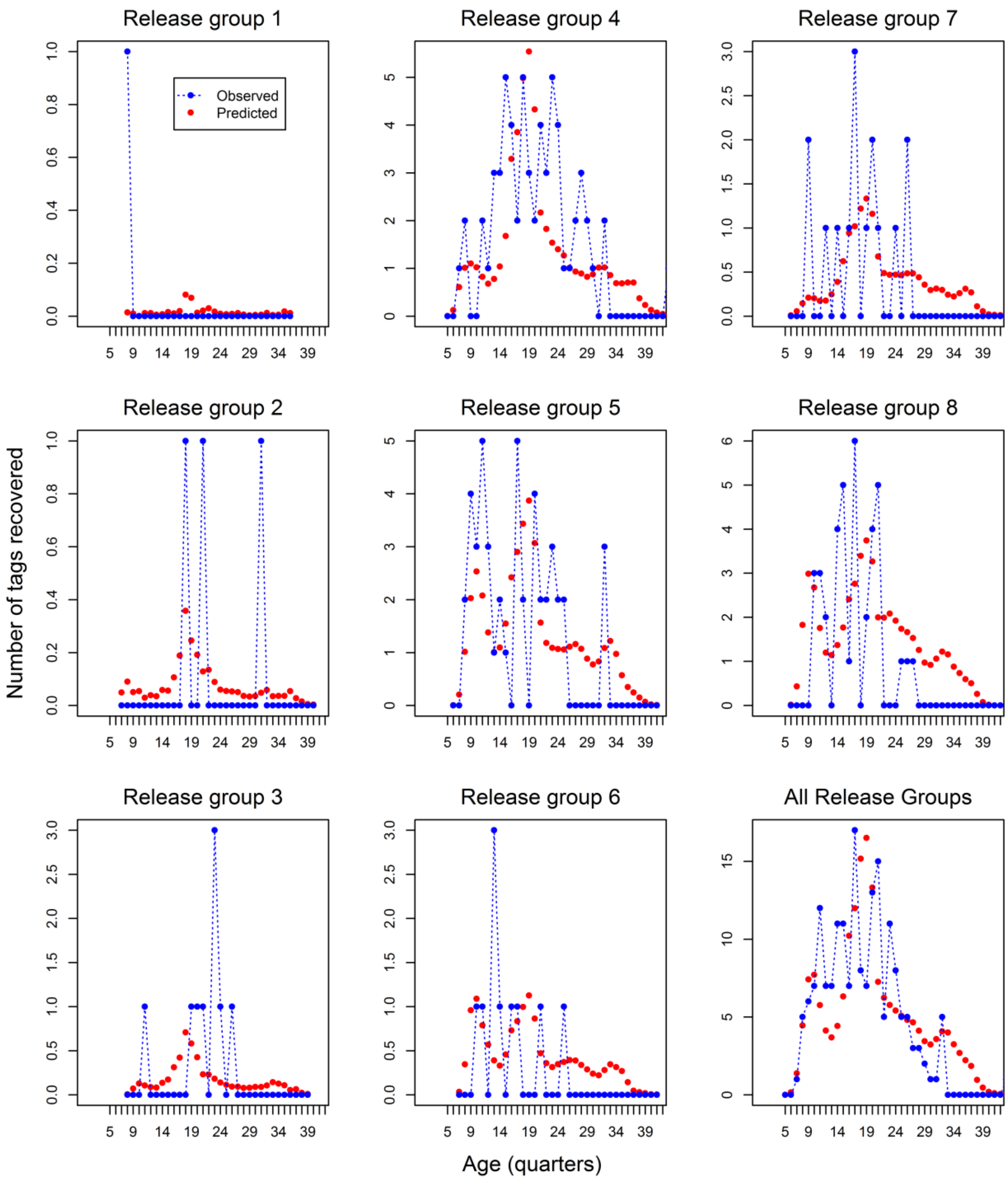

Figure 33. Observed and predicted the number of tag recoveries by the PSFS1S fishery by age (quarter) at recovery, excluding the recoveries during the mixing period (reference model). Tag release groups (1-8) represent the total releases in each quarter (combining the age groups that define individual SS release groups). The "All Release Groups" plot aggregates all recoveries by age class (at recovery).

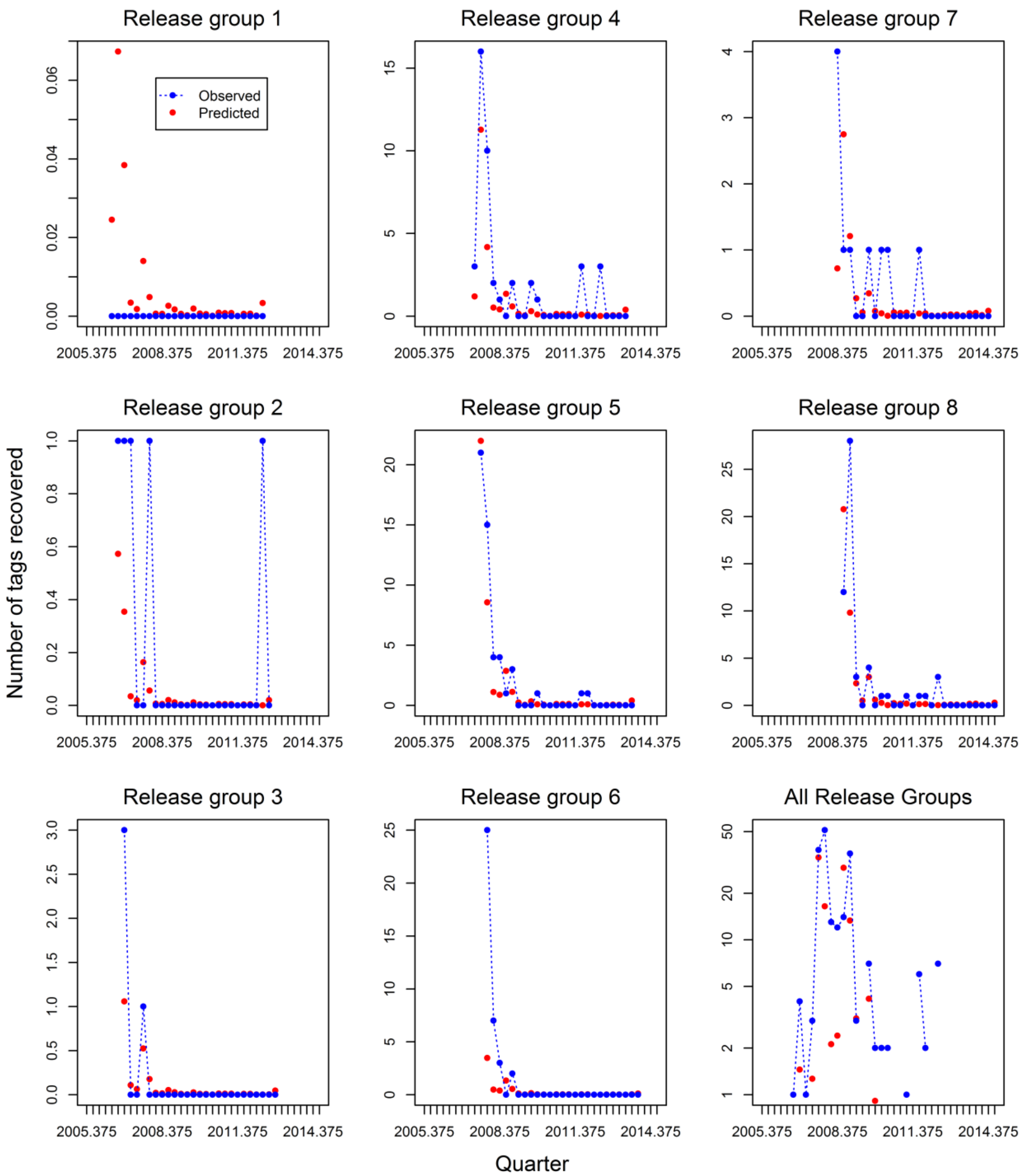

Figure 34. Observed and predicted the number of tag recoveries by the PSLS1N fishery by quarter following the mixing period (reference model). Tag release groups (1-8) represent the total releases in each quarter (aggregating the age groups that define individual SS release groups). The "All Release Groups" aggregates all recoveries by quarter (plotted on log scale).

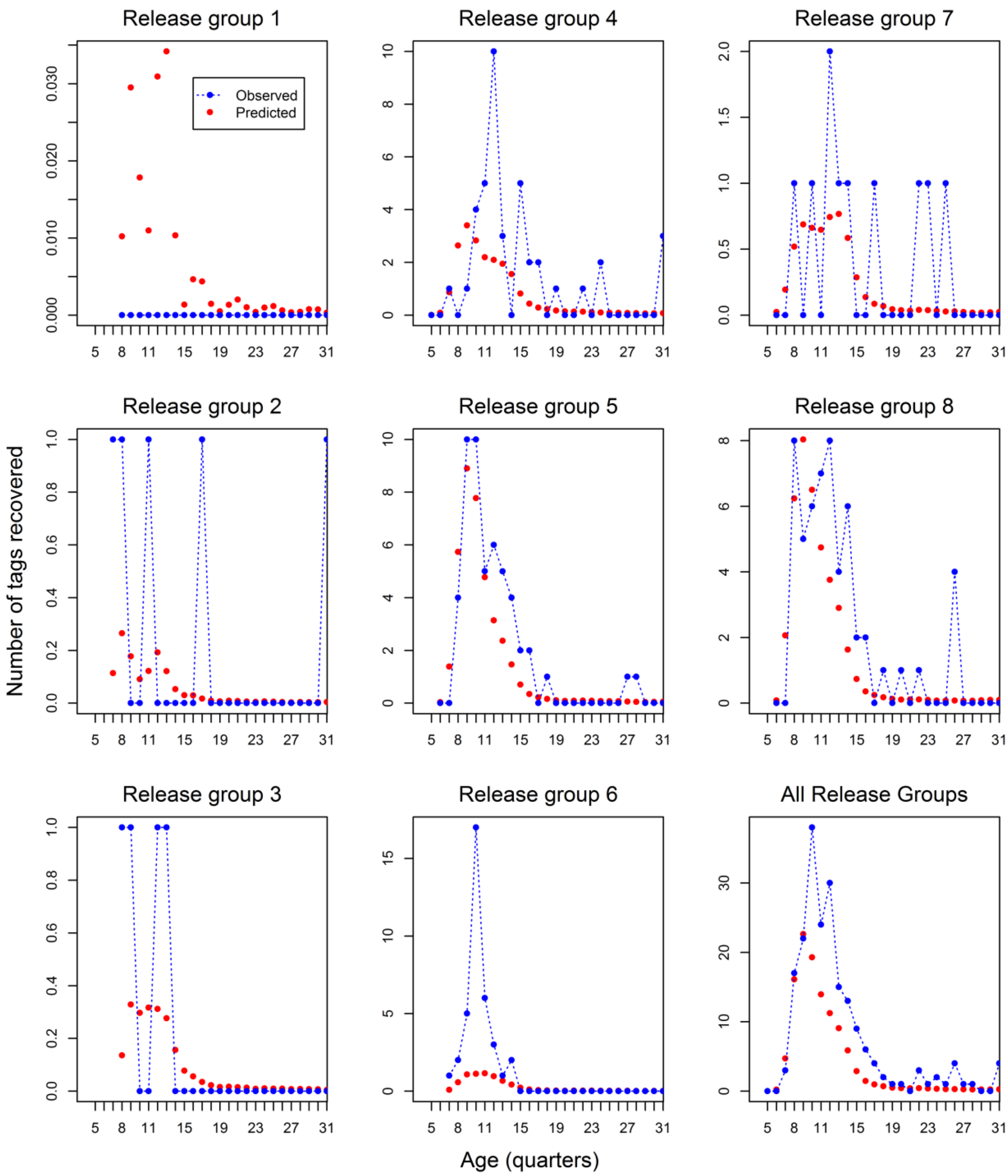

Figure 35. Observed and predicted the number of tag recoveries by the PSLS1N fishery by age (quarter) at recovery, excluding the recoveries during the mixing period (reference model). Tag release groups (1-8) represent the total releases in each quarter (combining the age groups that define individual SS release groups). The "All Release Groups" plot aggregates all recoveries by age class (at recovery).

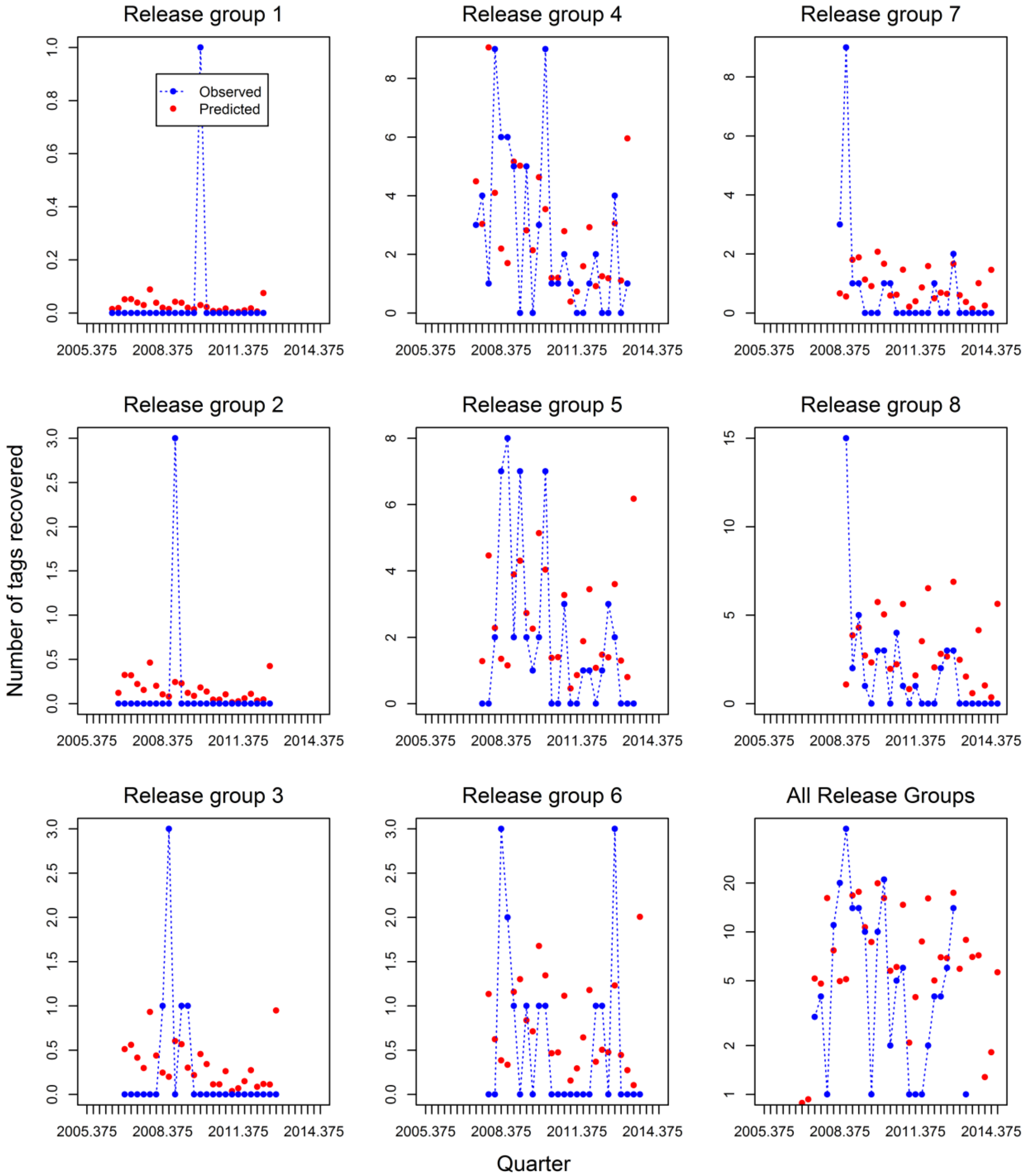

Figure 36. Observed and predicted the number of tag recoveries by the LL1S fishery by quarter following the mixing period (reference model). Tag release groups (1-8) represent the total releases in each quarter (aggregating the age groups that define individual SS release groups). The "All Release Groups" aggregates all recoveries by quarter (plotted on log scale).

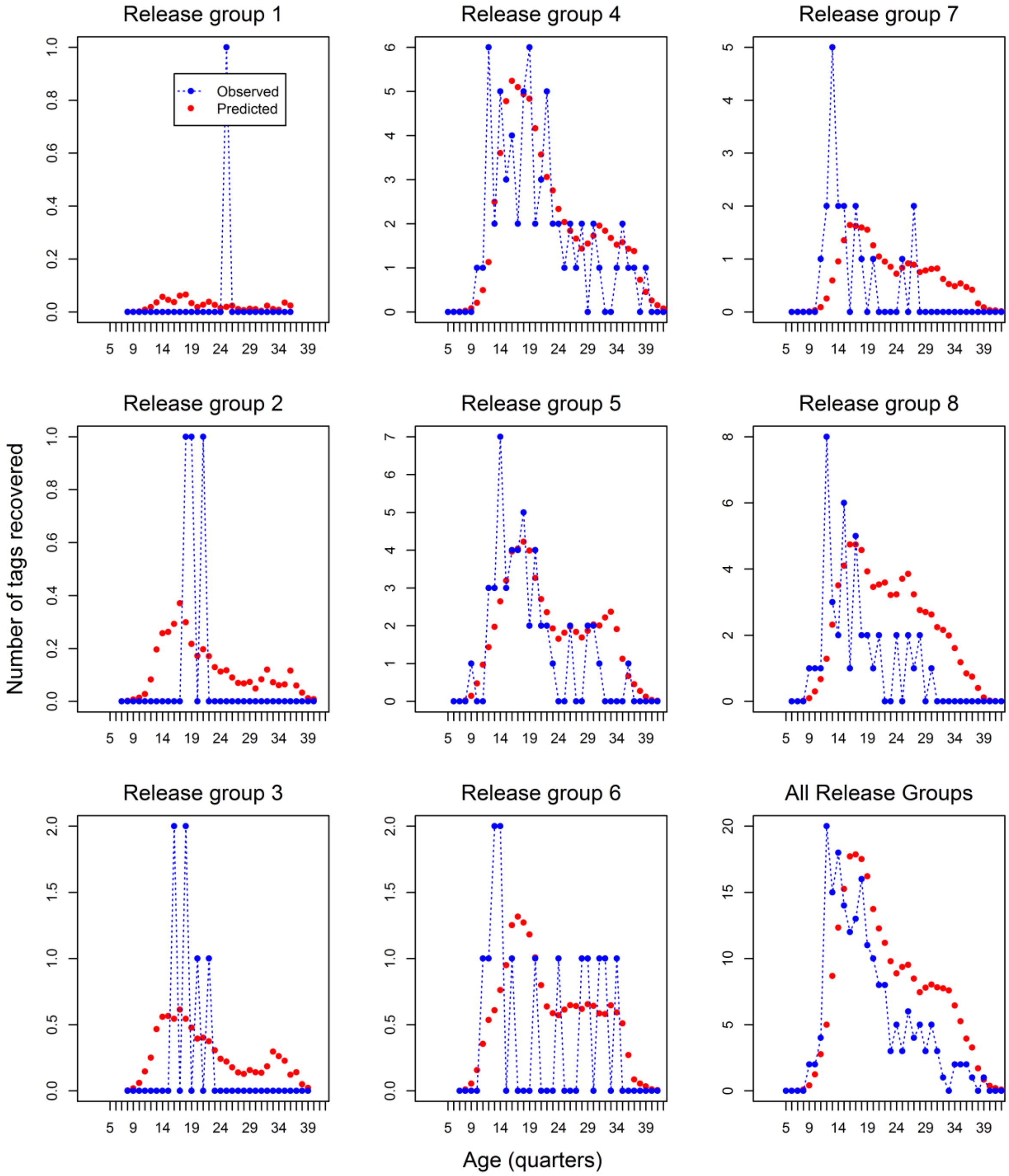

Figure 37. Observed and predicted the number of tag recoveries by the LL1S fishery by age (quarter) at recovery, excluding the recoveries during the mixing period (reference model). Tag release groups (1-8) represent the total releases in each quarter (combining the age groups that define individual SS release groups). The "All Release Groups" plot aggregates all recoveries by age class (at recovery).

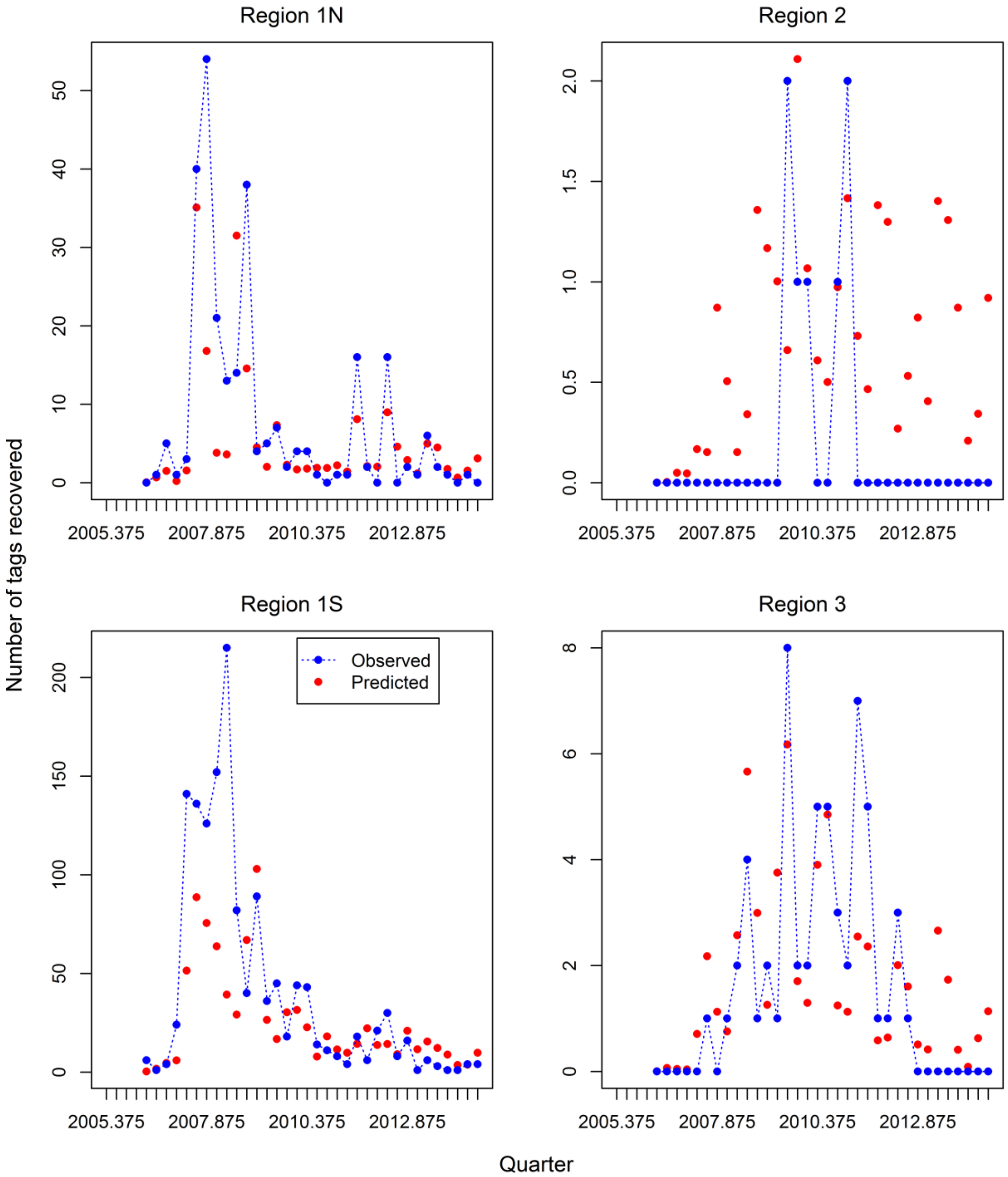

Figure 38. Total observed and predicted the number of tag recoveries by region (all regional fisheries combined) by quarter following the mixing period (reference model).

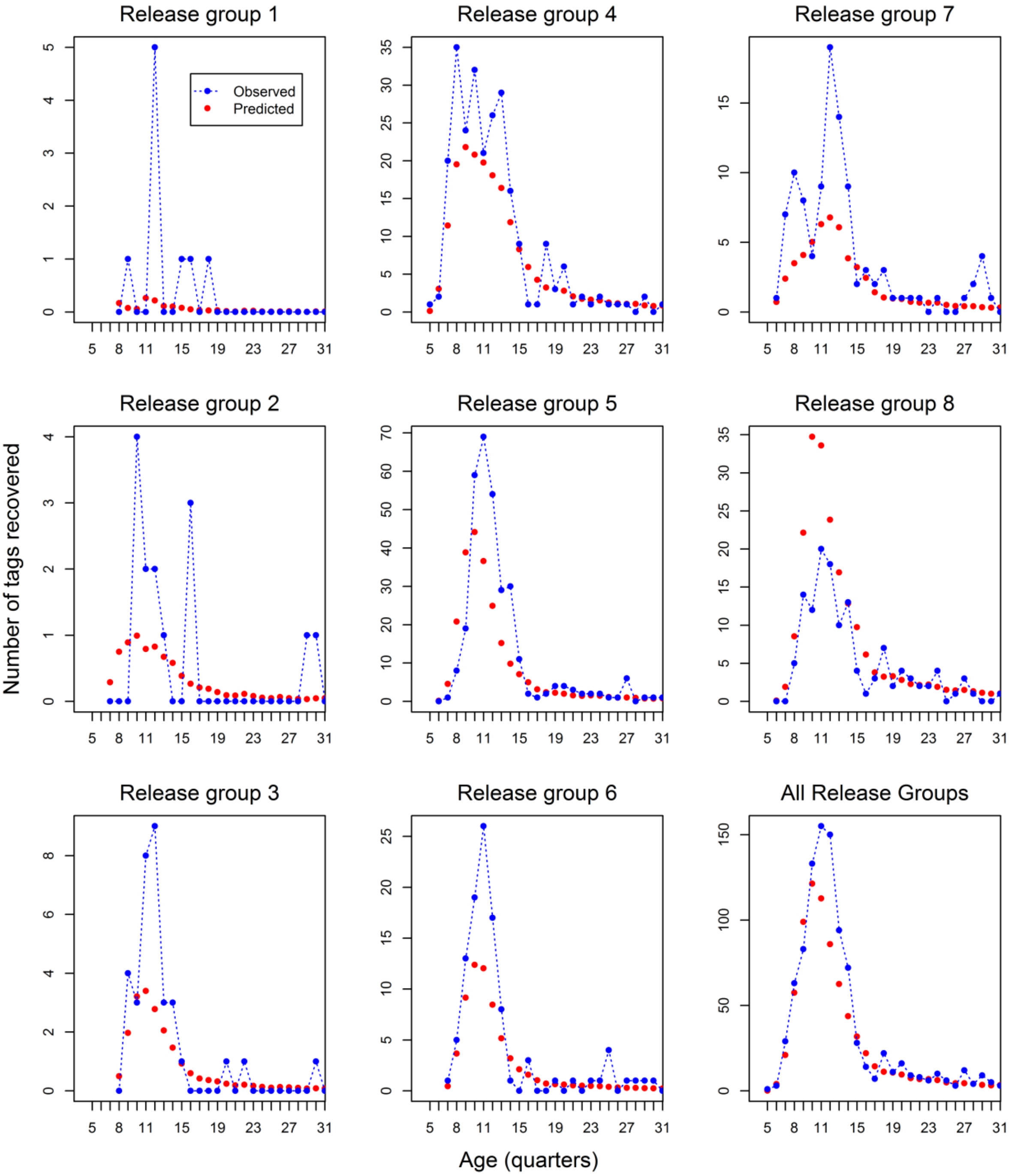

Figure 39. Observed and predicted the number of tag recoveries by the PSLS1N fishery by age (quarter) at recovery, excluding the recoveries during the mixing period for the model option with a <u>high</u> weighting for the tagging data (lambda 1.0). Tag release groups (1-8) represent the total releases in each quarter (combining the age groups that define individual SS release groups). The “All Release Groups” plot aggregates all recoveries by age class (at recovery).

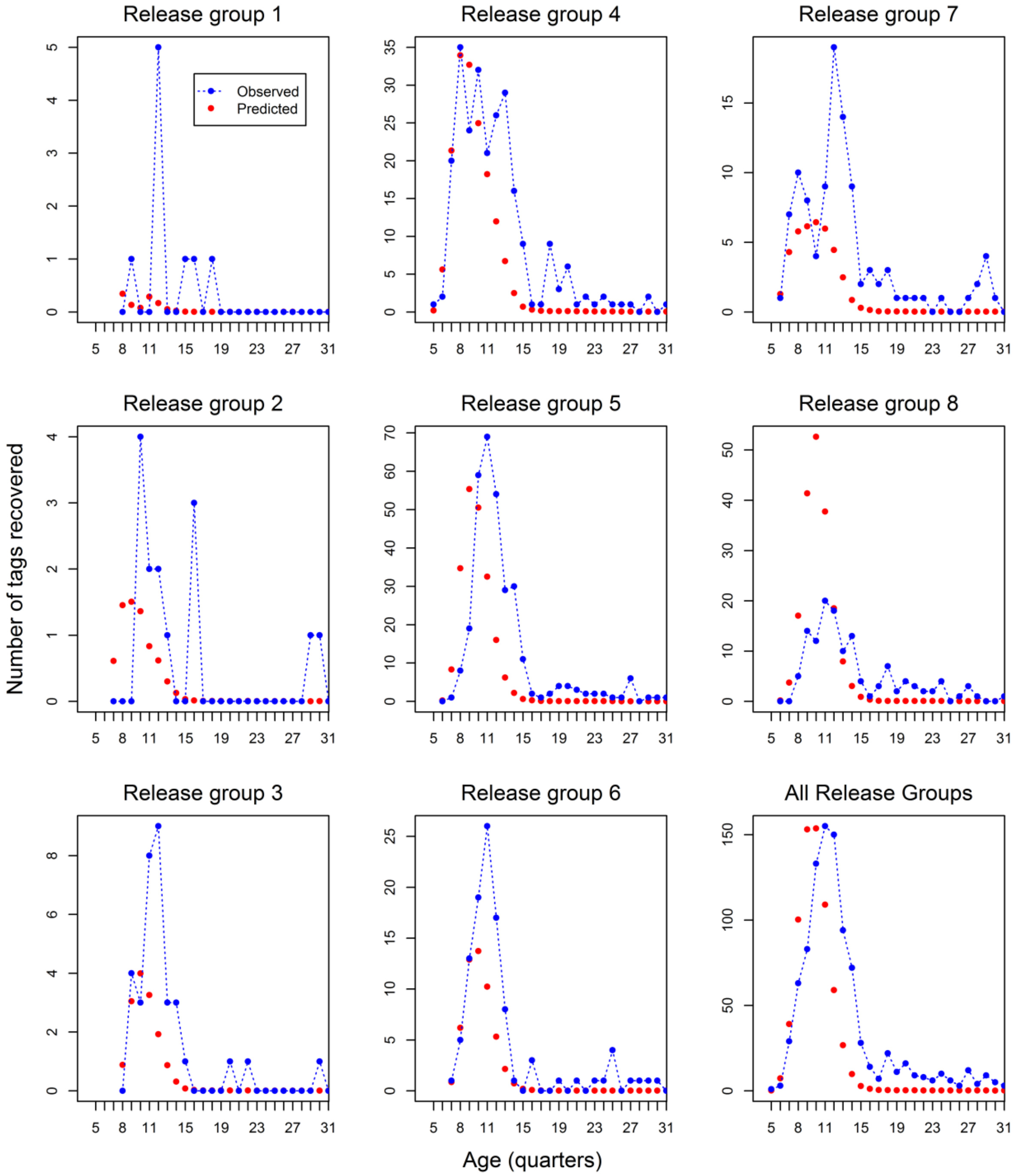

Figure 40. Observed and predicted the number of tag recoveries by the PSLS1N fishery by age (quarter) at recovery, excluding the recoveries during the mixing period for the model option with a <u>low</u> weighting for the tagging data (lambda 0.01). Tag release groups (1-8) represent the total releases in each quarter (combining the age groups that define individual SS release groups). The "All Release Groups" plot aggregates all recoveries by age class (at recovery).

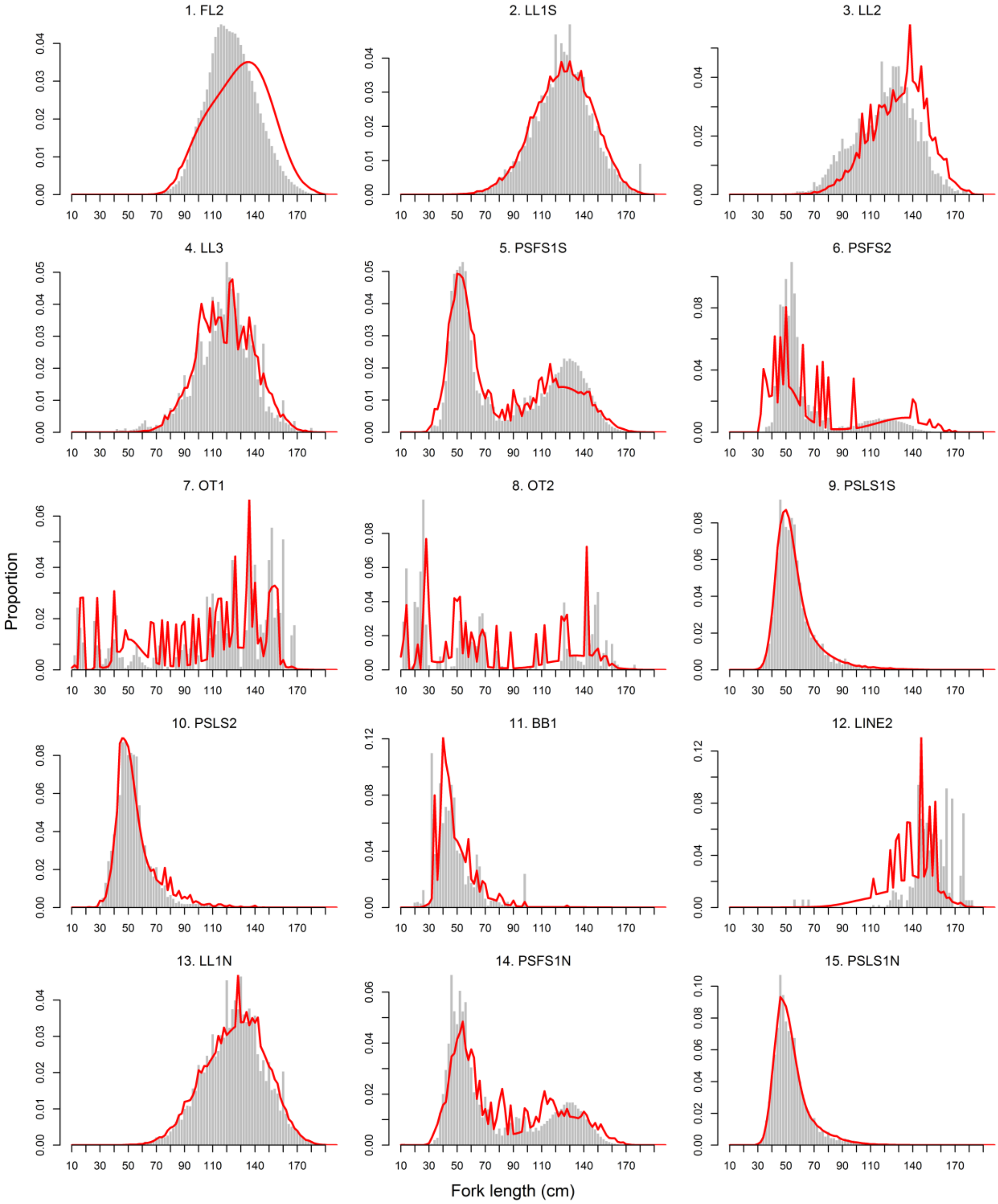

Figure 41. The observed (grey polygon) and predicted (red line) aggregated length compositions for the main fisheries with length frequency data from the reference model.

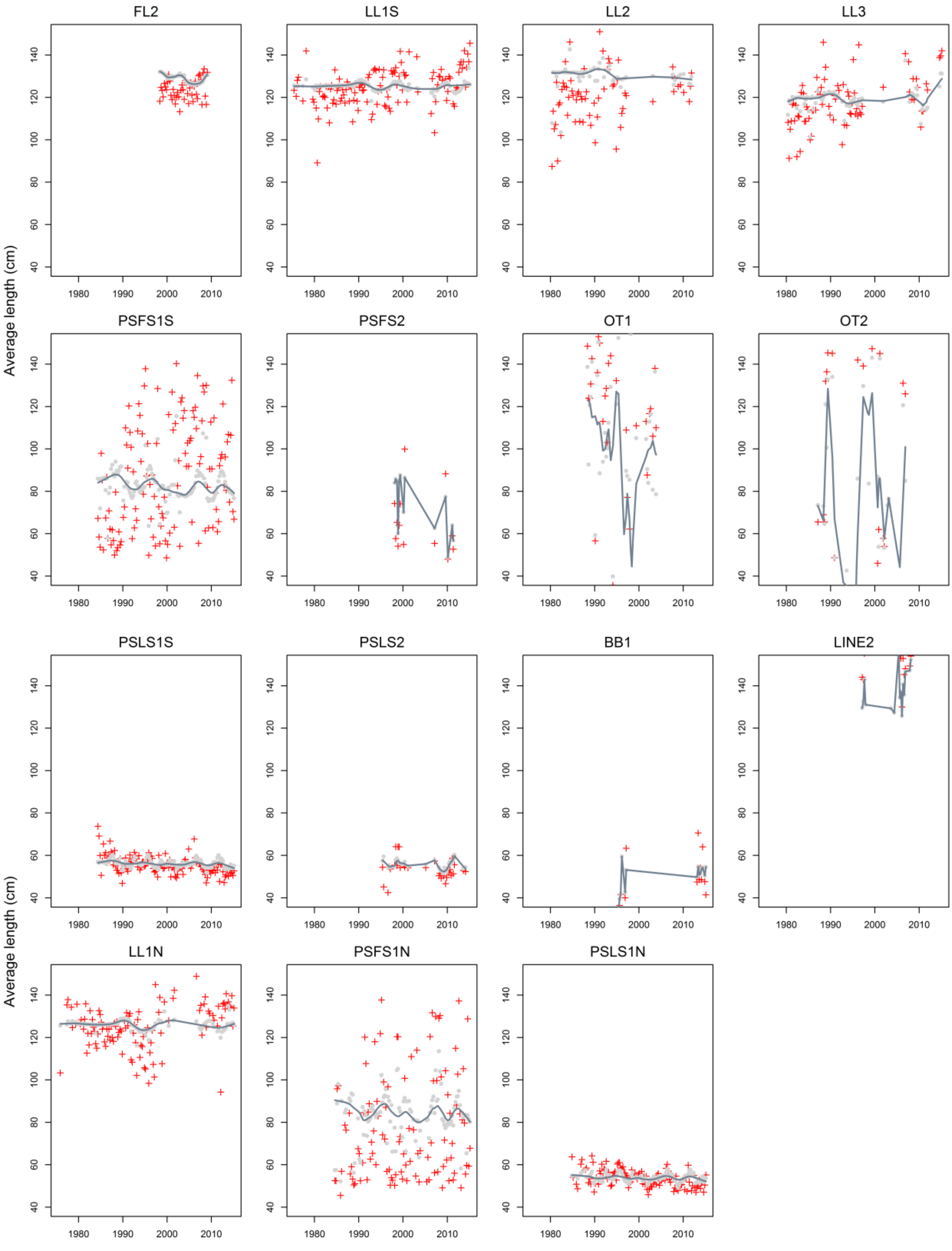

Figure 42. The observed (red points) and predicted (grey points) average fish length by quarter derived from the observed and predicted fishery length compositions from the reference model. The grey line represents the lowess smoothed fit to the predicted average.

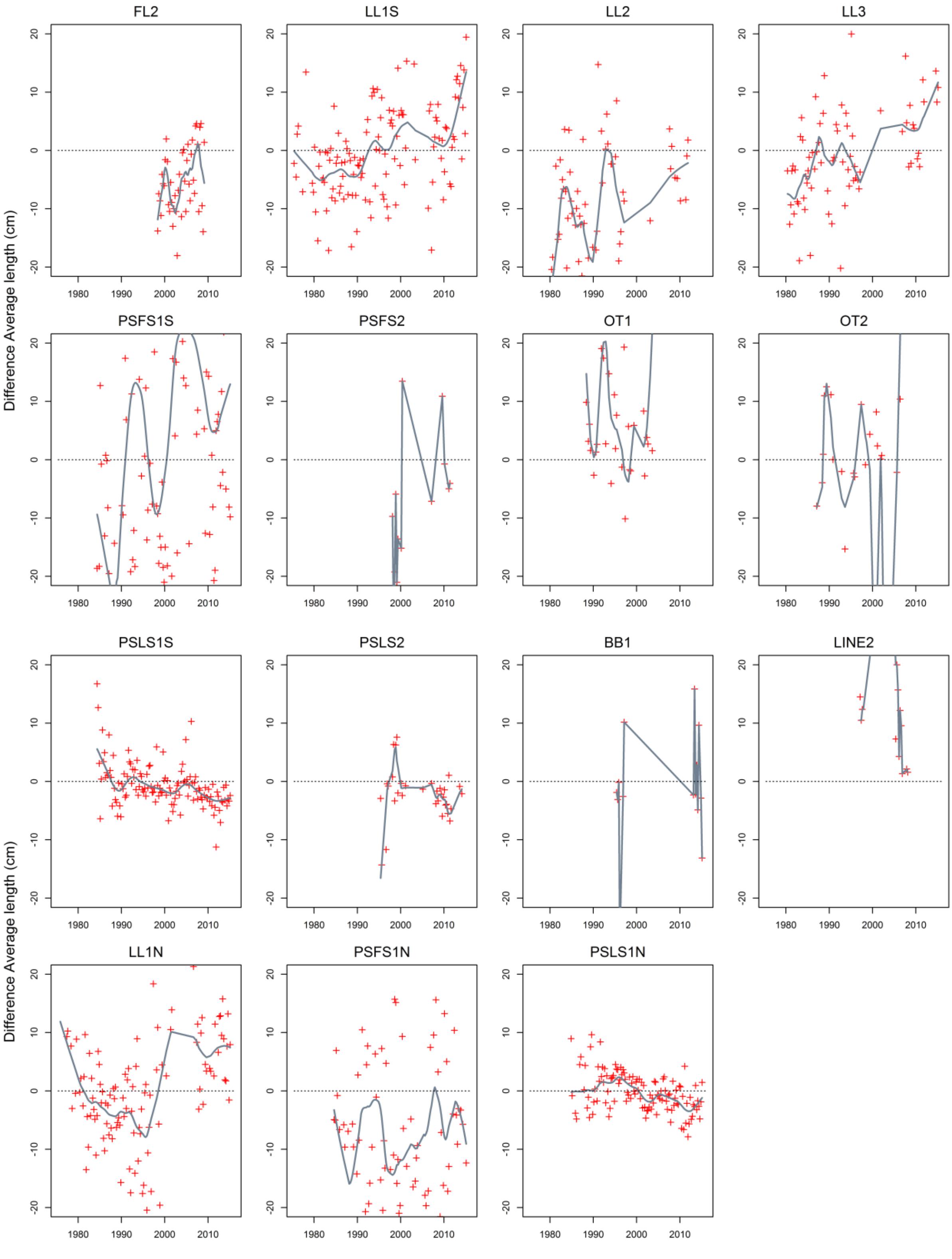

Figure 43. The difference between the observed and predicted average fish length for each fishery by quarter derived from the observed and predicted length compositions from the reference model. The grey line represents the lowess smoothed fit to the values.

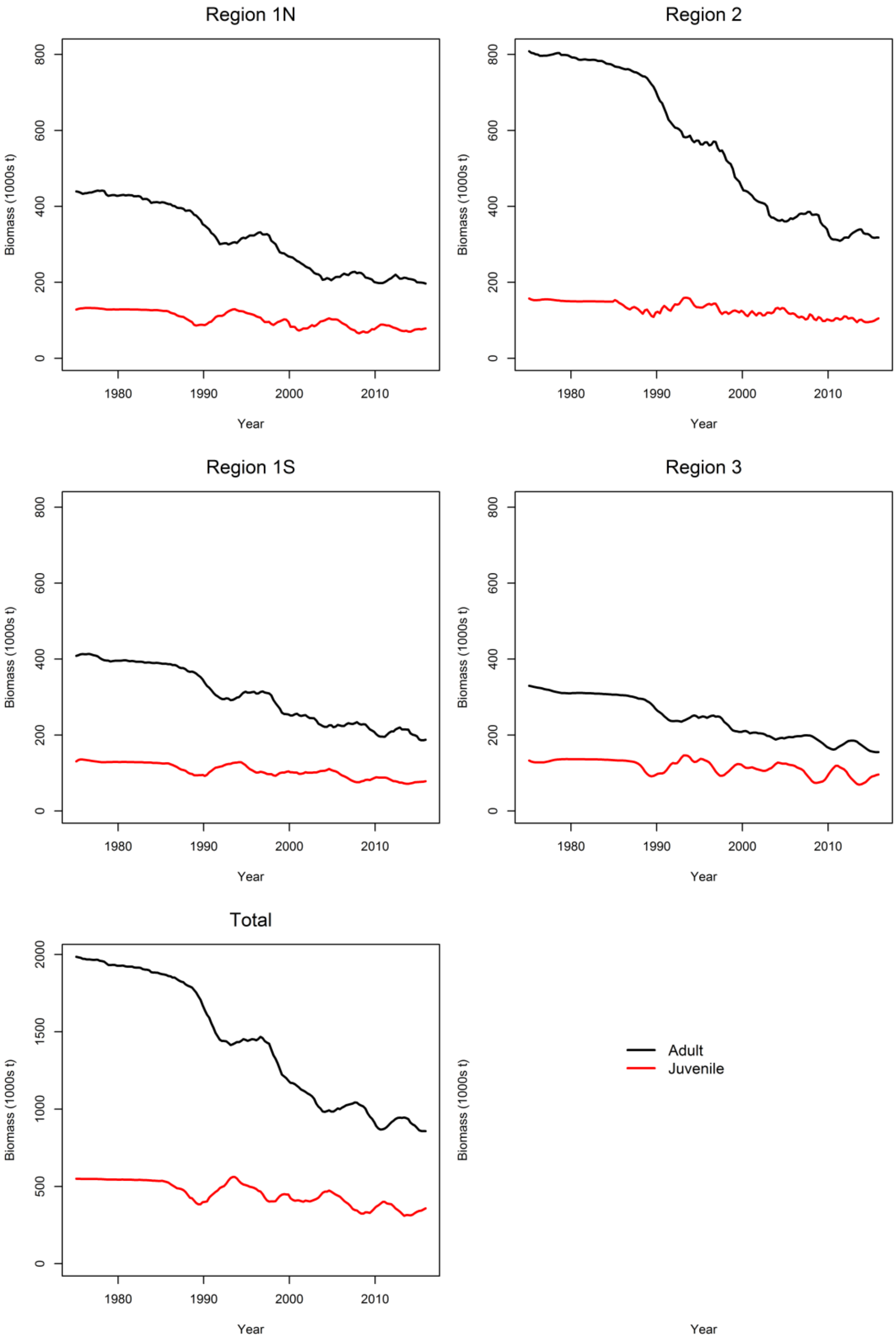

Figure 44. Trend in juvenile and adult biomass by region and total for the reference model.

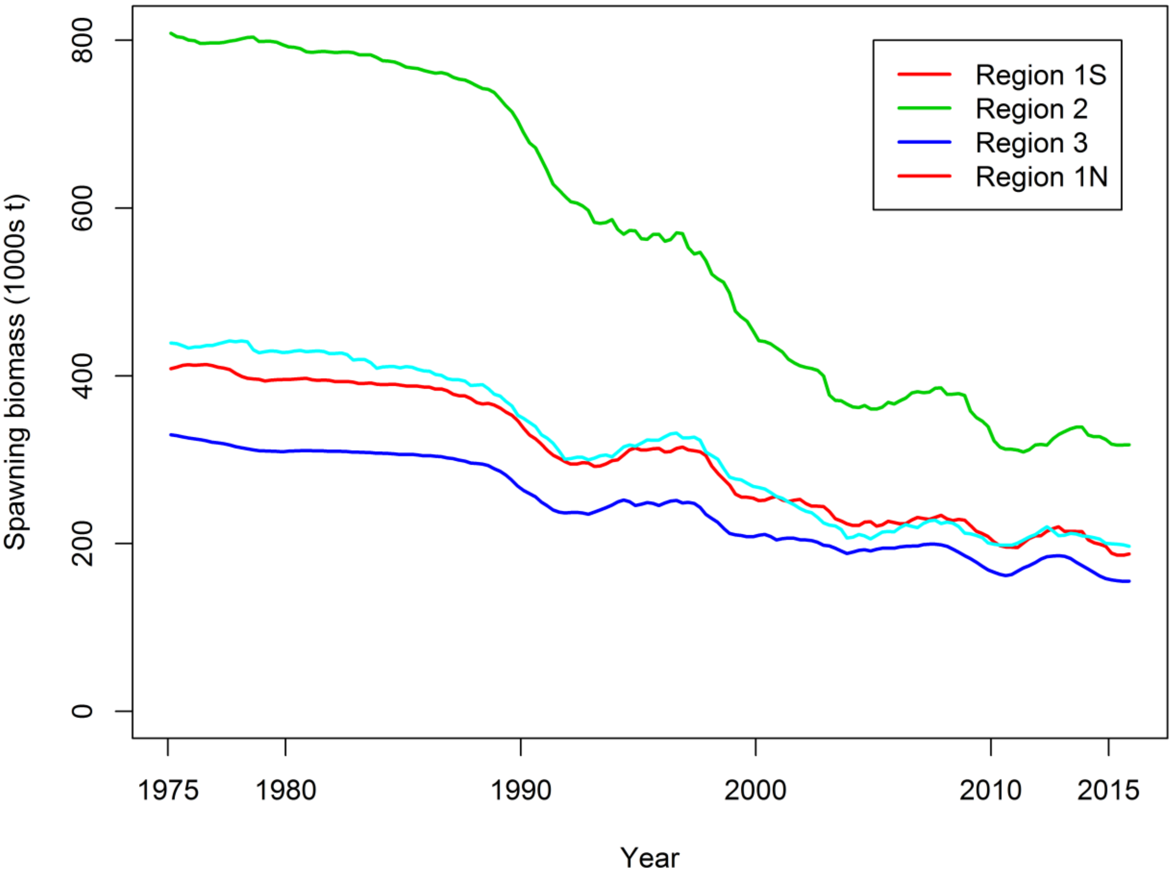

Figure 45. Comparison of trends in spawning biomass by region for the reference model.

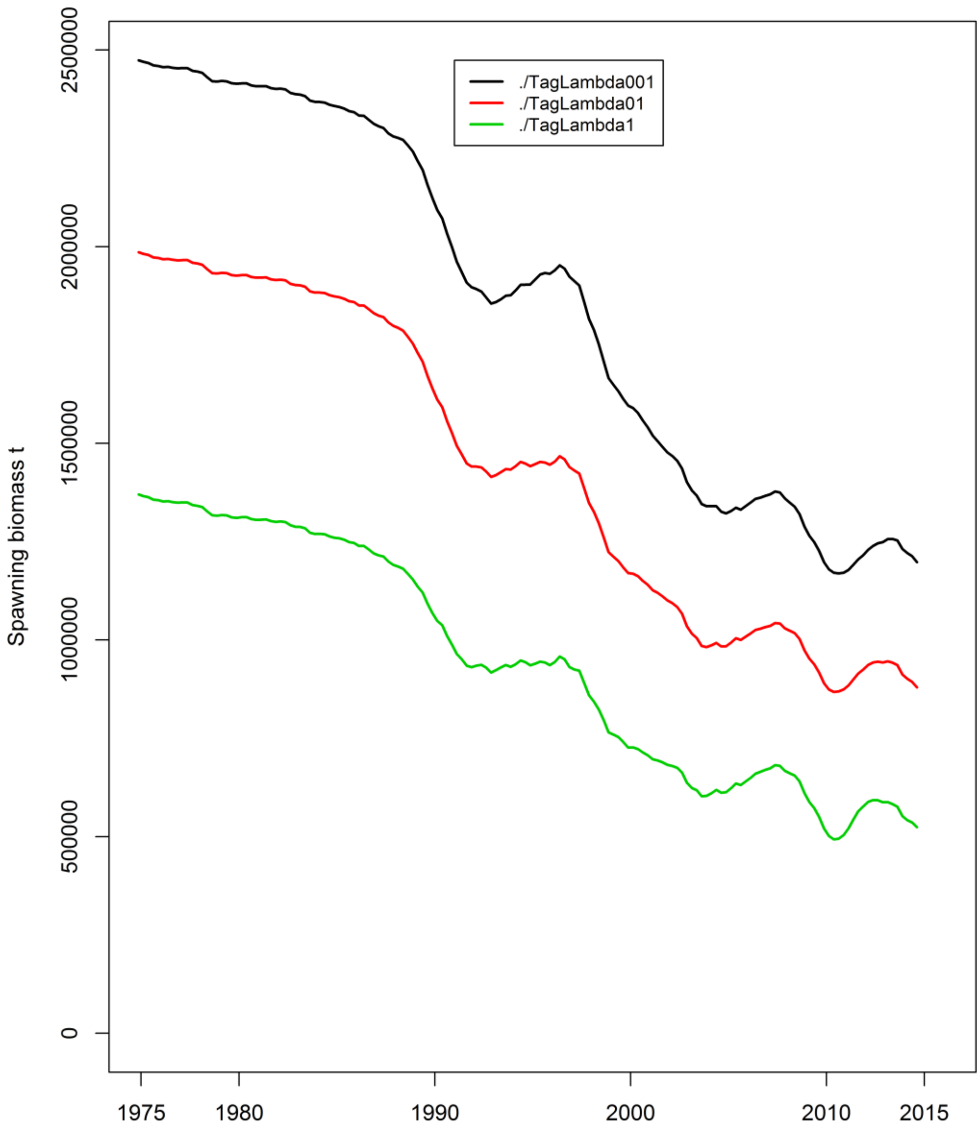

Figure 46. Comparison of trends in spawning biomass by region for the model options with different relative weighting of the tagging data set (lambda 0.1 reference model).

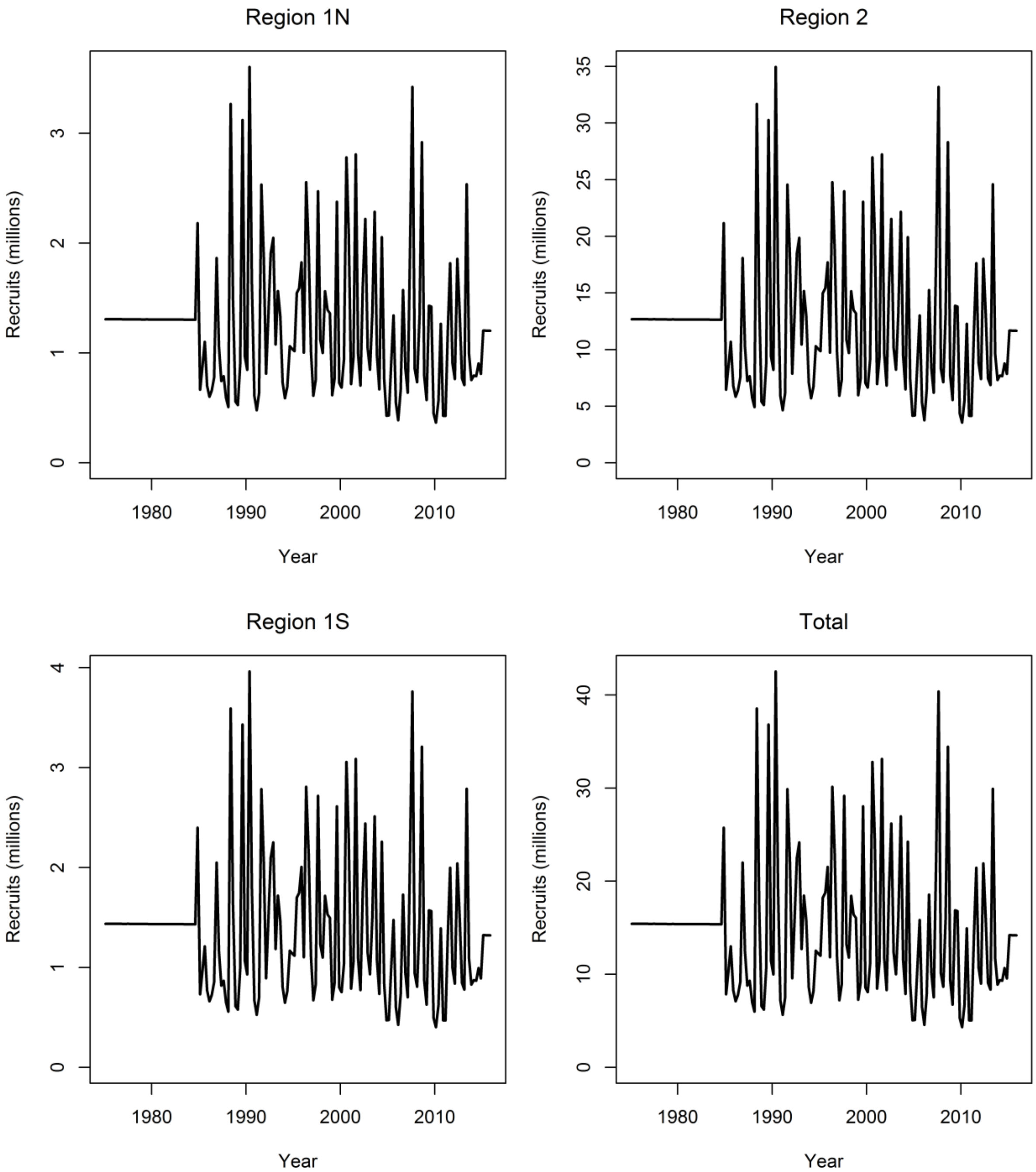

Figure 47. Quarterly recruitment by region and aggregated from the reference model.

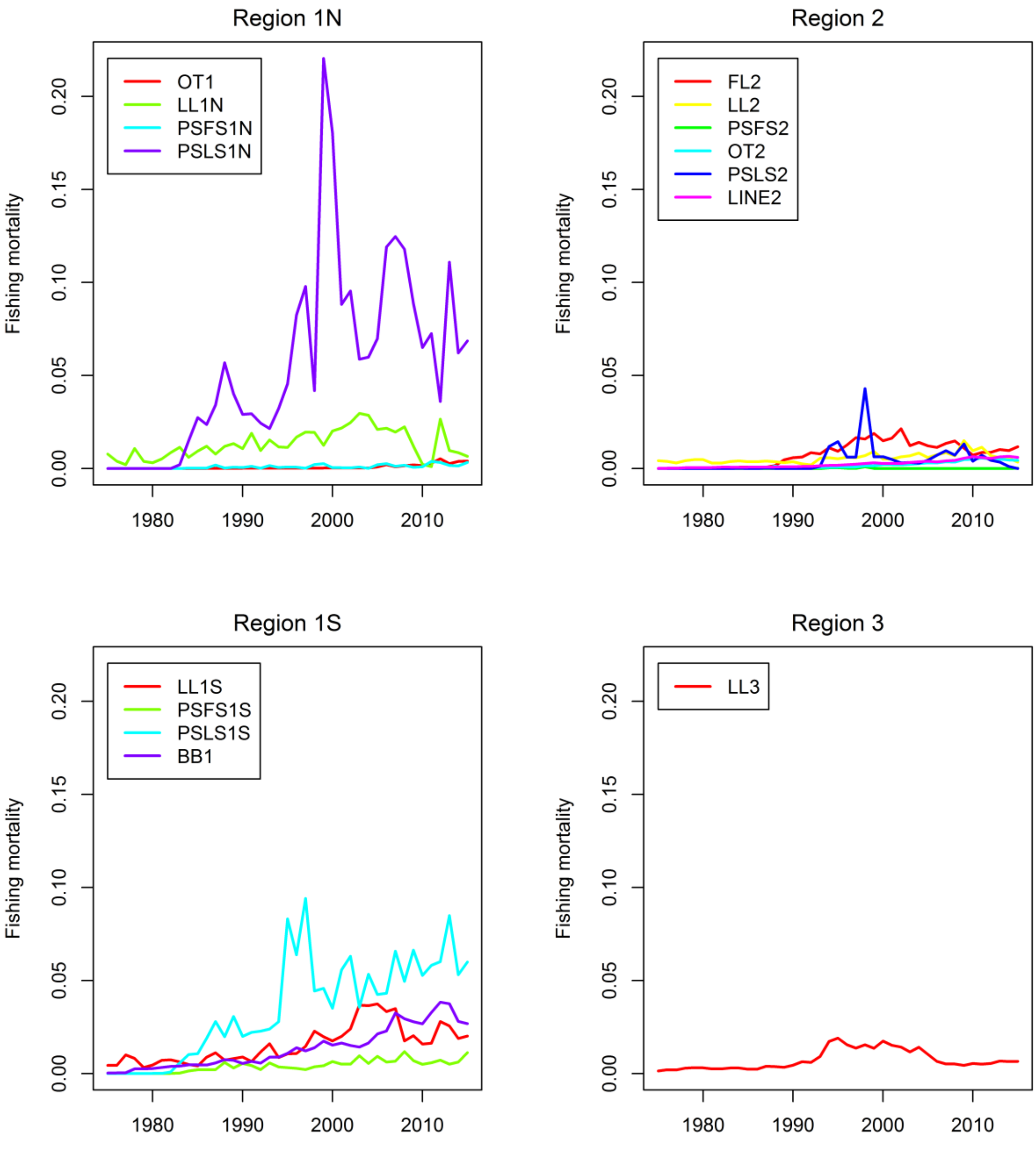

Figure 48. Estimates of fishery specific fishing mortality by fishery region from the reference model.

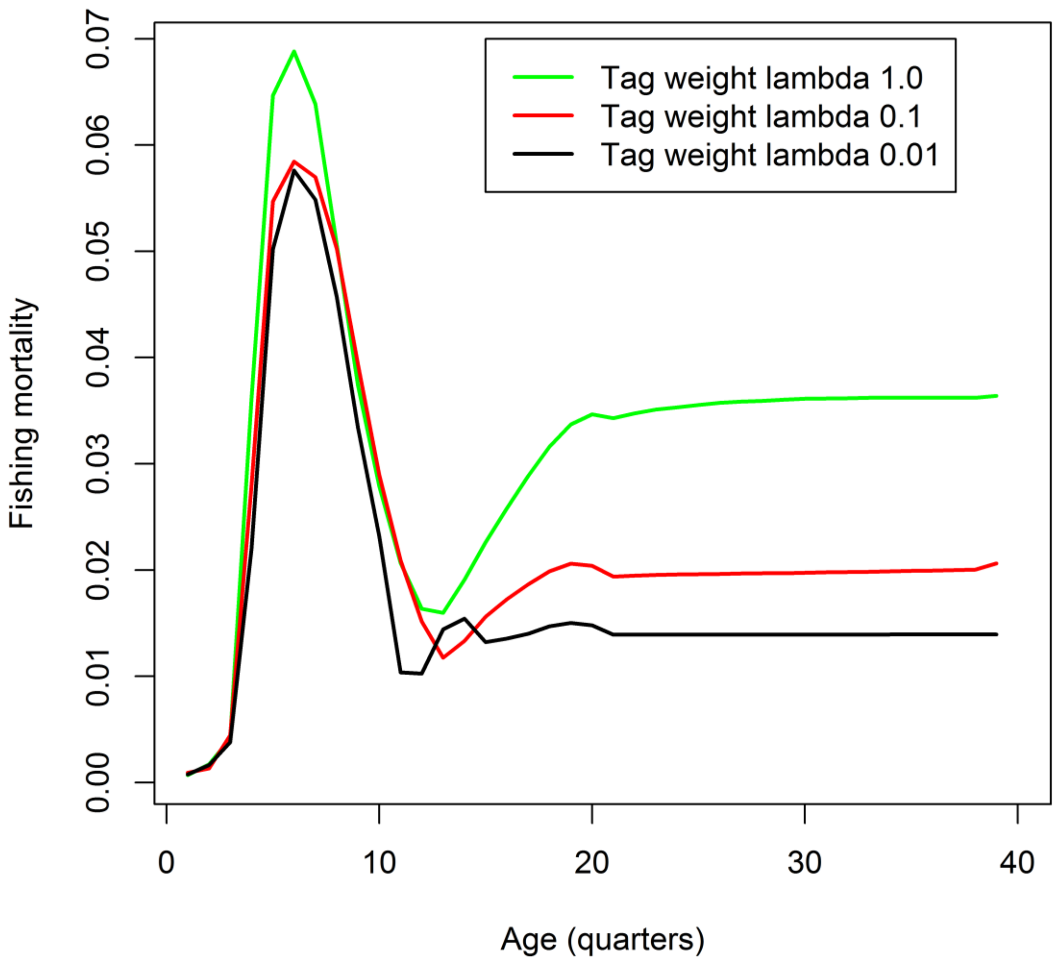

Figure 49. Average fishing mortality by age class from the reference period (2013 and 2014) for the three different levels of weighting of the tagging data.

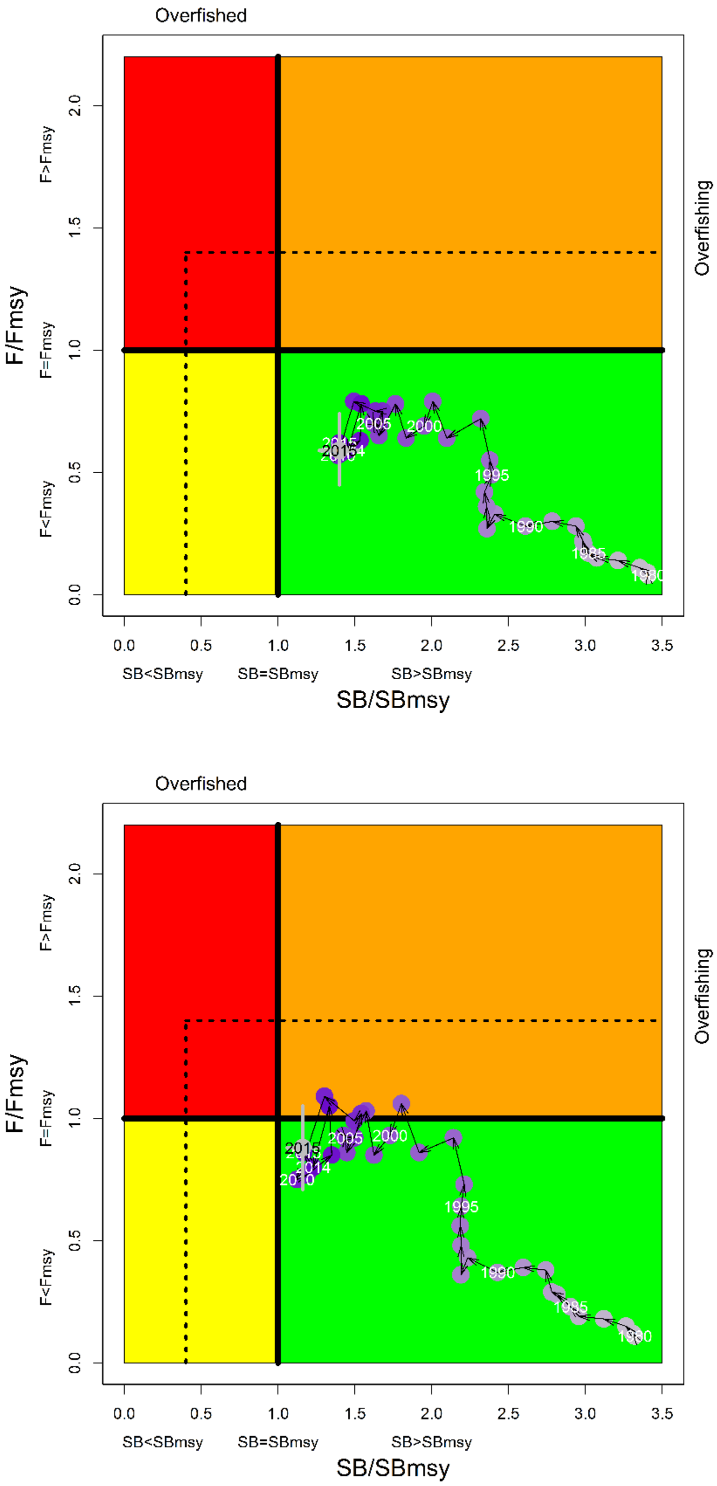

Figure 50. **Kobe plots for the *reference* model option (top) and the model option with *__higher__* tag weighting (lambda 1.0) (bottom). The grey lines represent 95% confidence intervals of the stock status metrics for the terminal year (2015).**

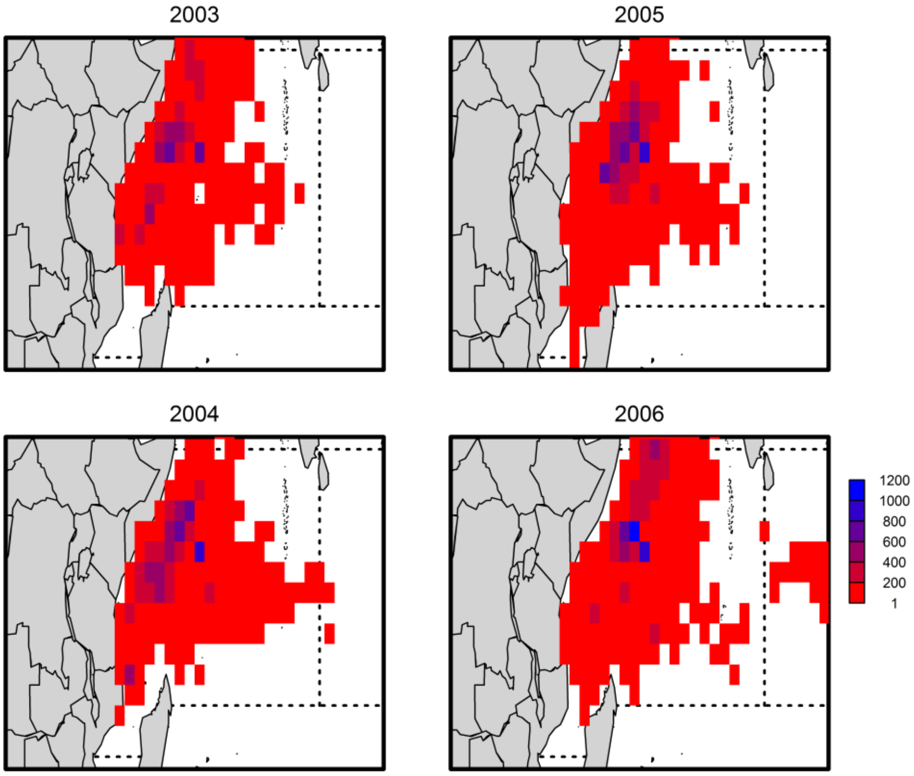

Figure 51. Bigeye tuna catch (mt) from the FAD purse seine fishery by 2 degree of latitude/longitude from 2003–2006.

APPENDIX 1. ANALYSIS OF LENGTH COMPOSITION DATA FROM THE LONGLINE FISHERIES.

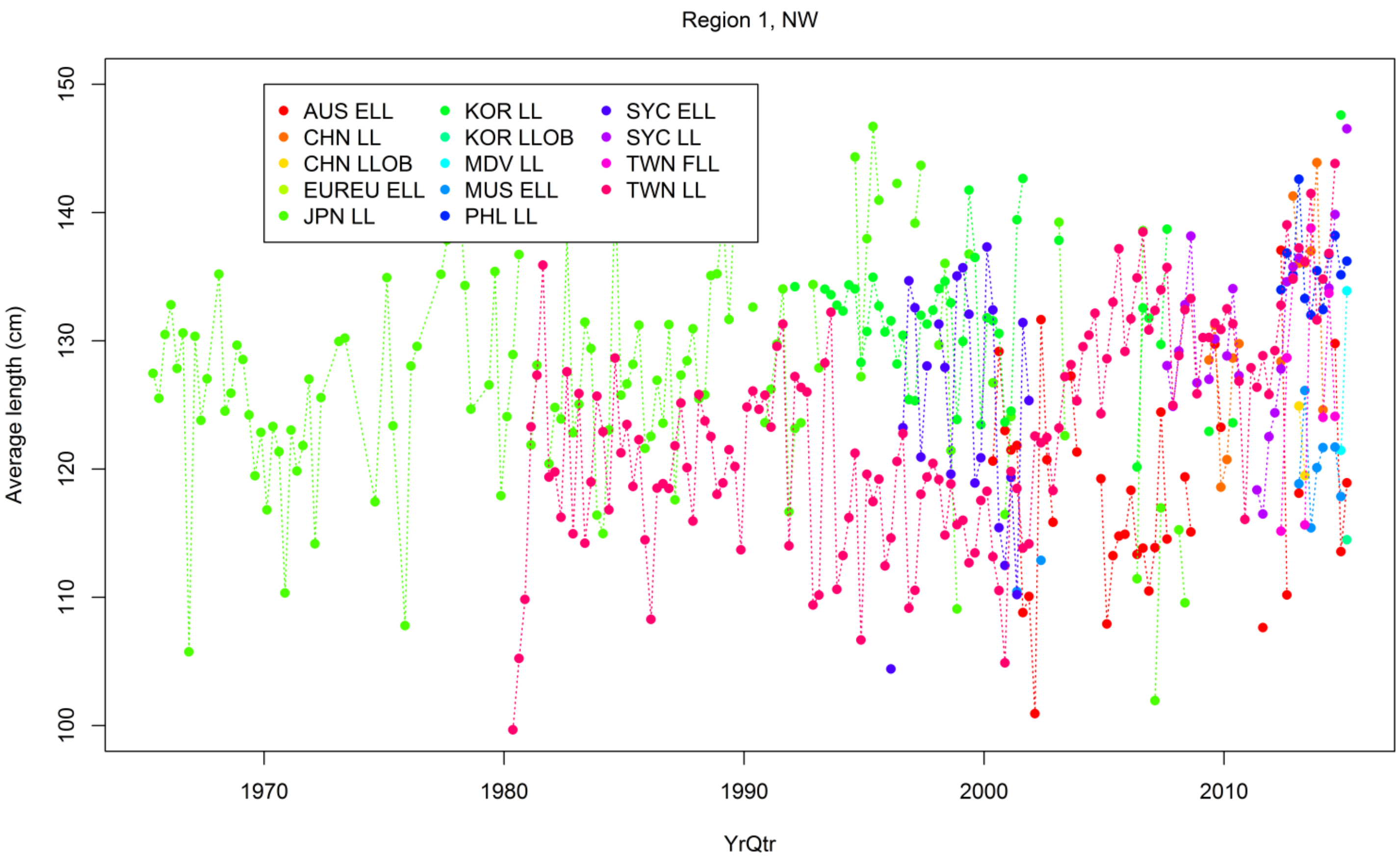

Figure A1: Average length of bigeye tuna (FL, cm) of fish sampled from the longline fishery in the western equatorial region, by fleet nationality and year/quarter.

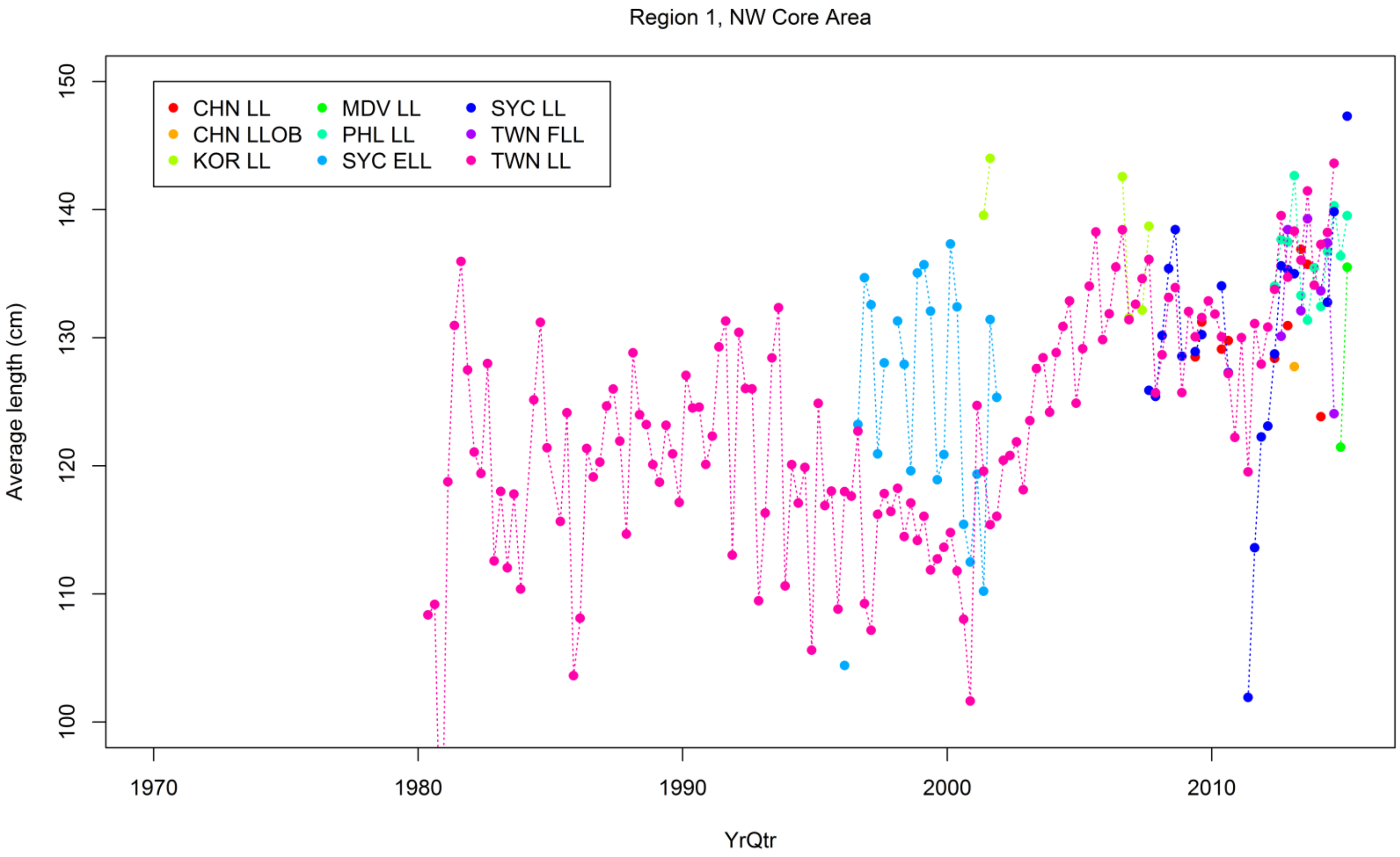

Figure A2: Average length of bigeye tuna (FL, cm) of fish sampled from the longline fishery in the core fishery area of the western equatorial region, by fleet nationality and year/quarter.

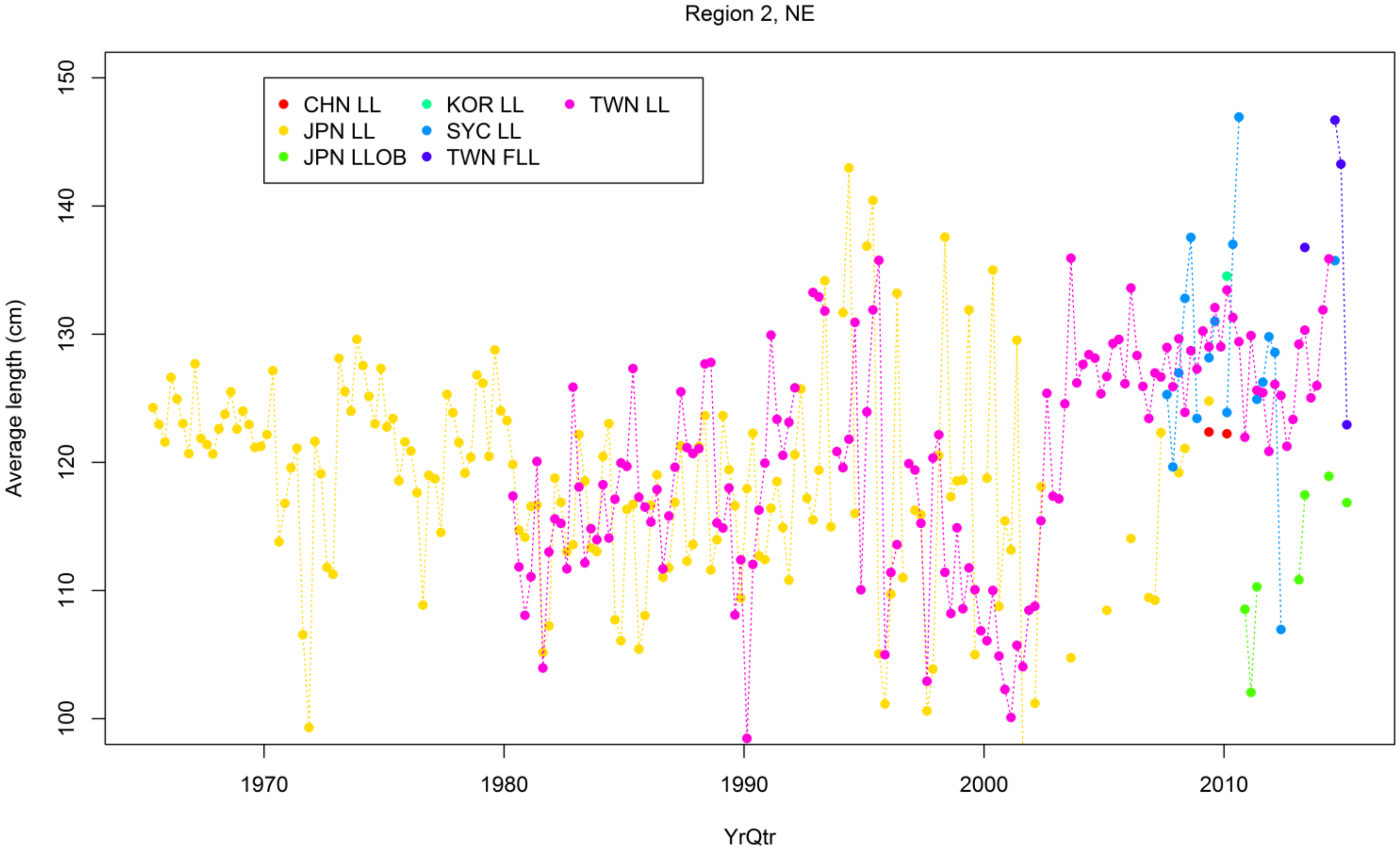

Figure A3: Average length of bigeye tuna (FL, cm) of fish sampled from the longline fishery in the eastern equatorial region, by fleet nationality and year/quarter.

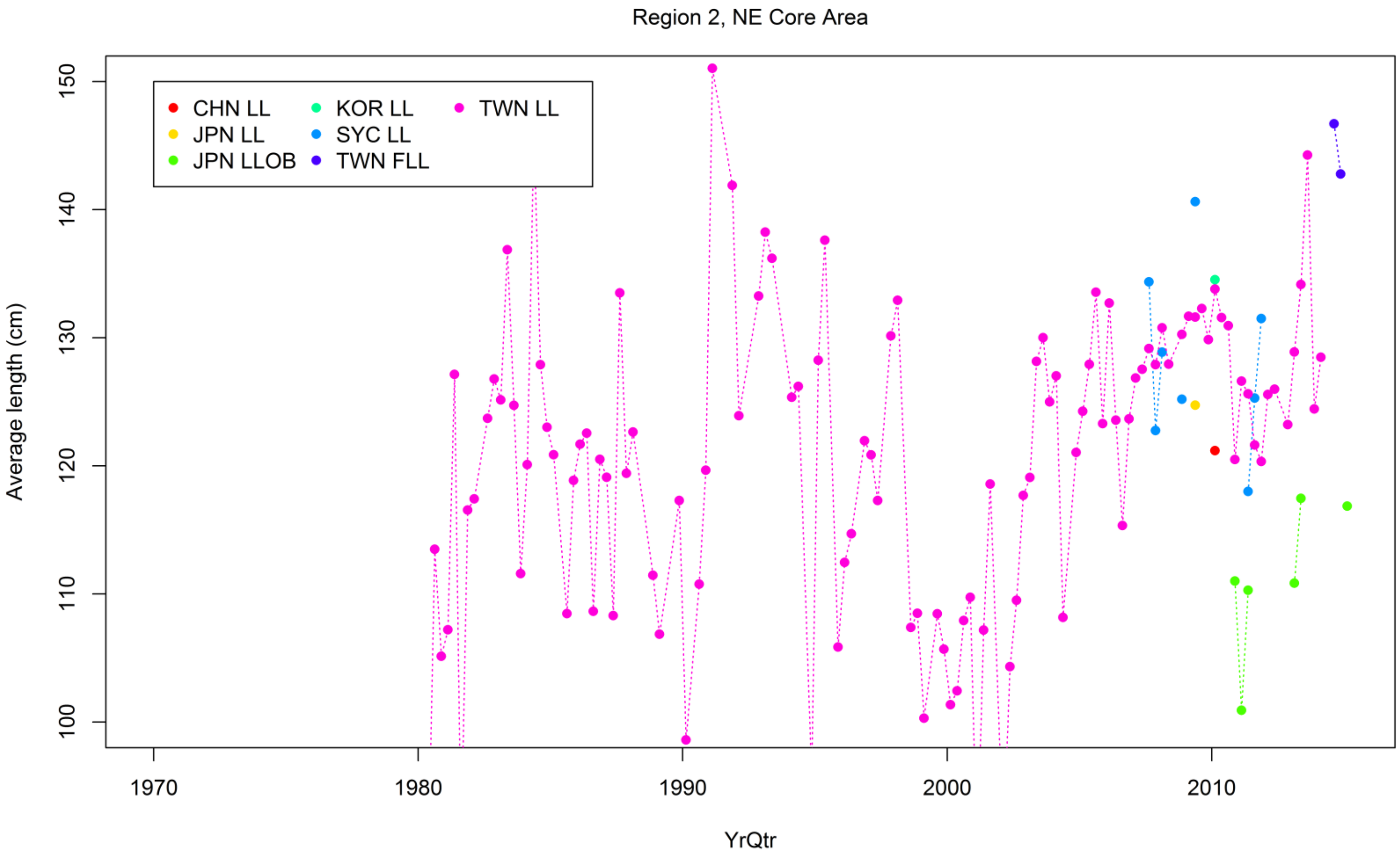

Figure A4: Average length of bigeye tuna (FL, cm) of fish sampled from the longline fishery in the core fishery area of the eastern equatorial region, by fleet nationality and year/quarter.

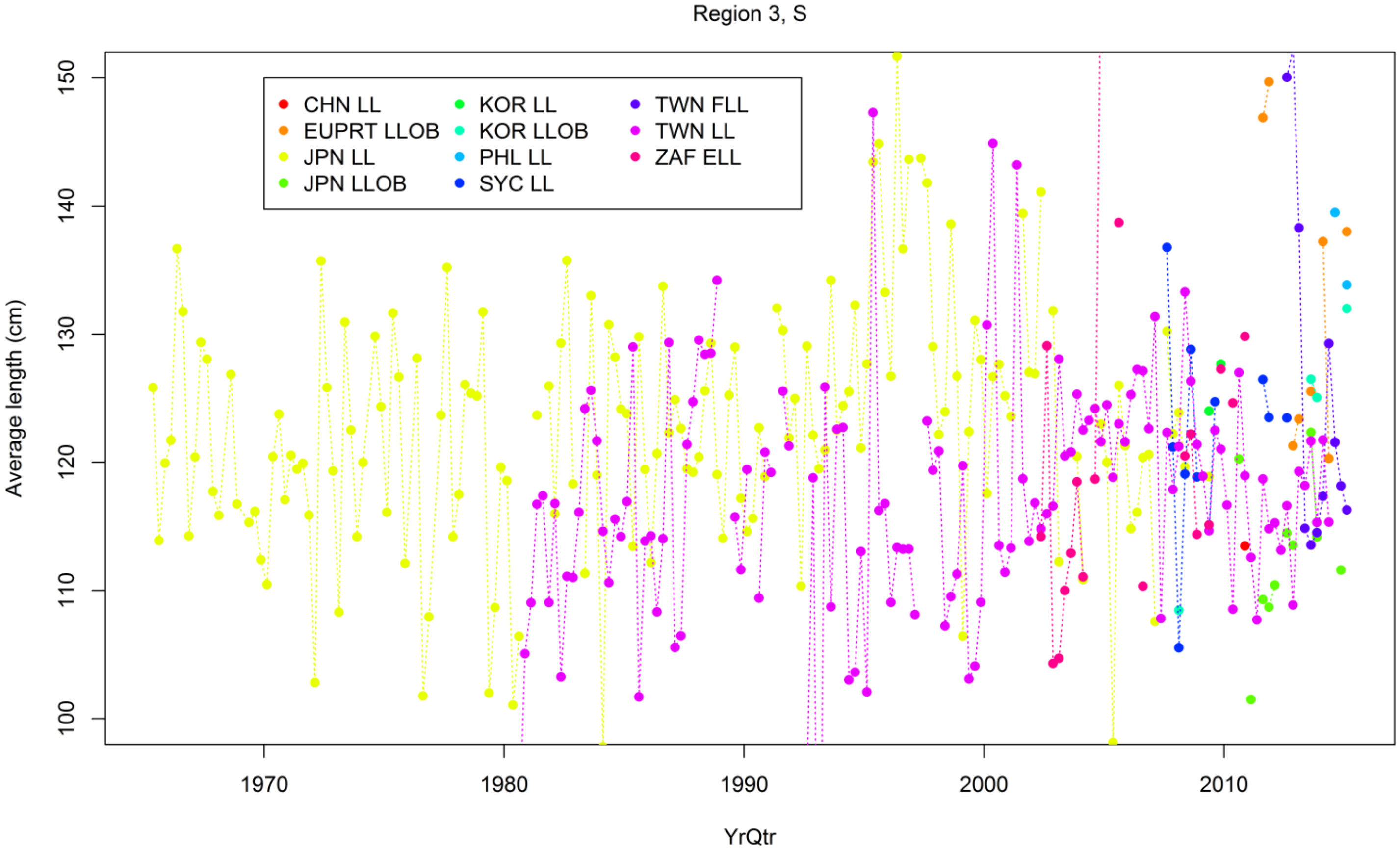

Figure A5: Average length of bigeye tuna (FL, cm) of fish sampled from the longline fishery in the southern region, by fleet nationality and year/quarter.

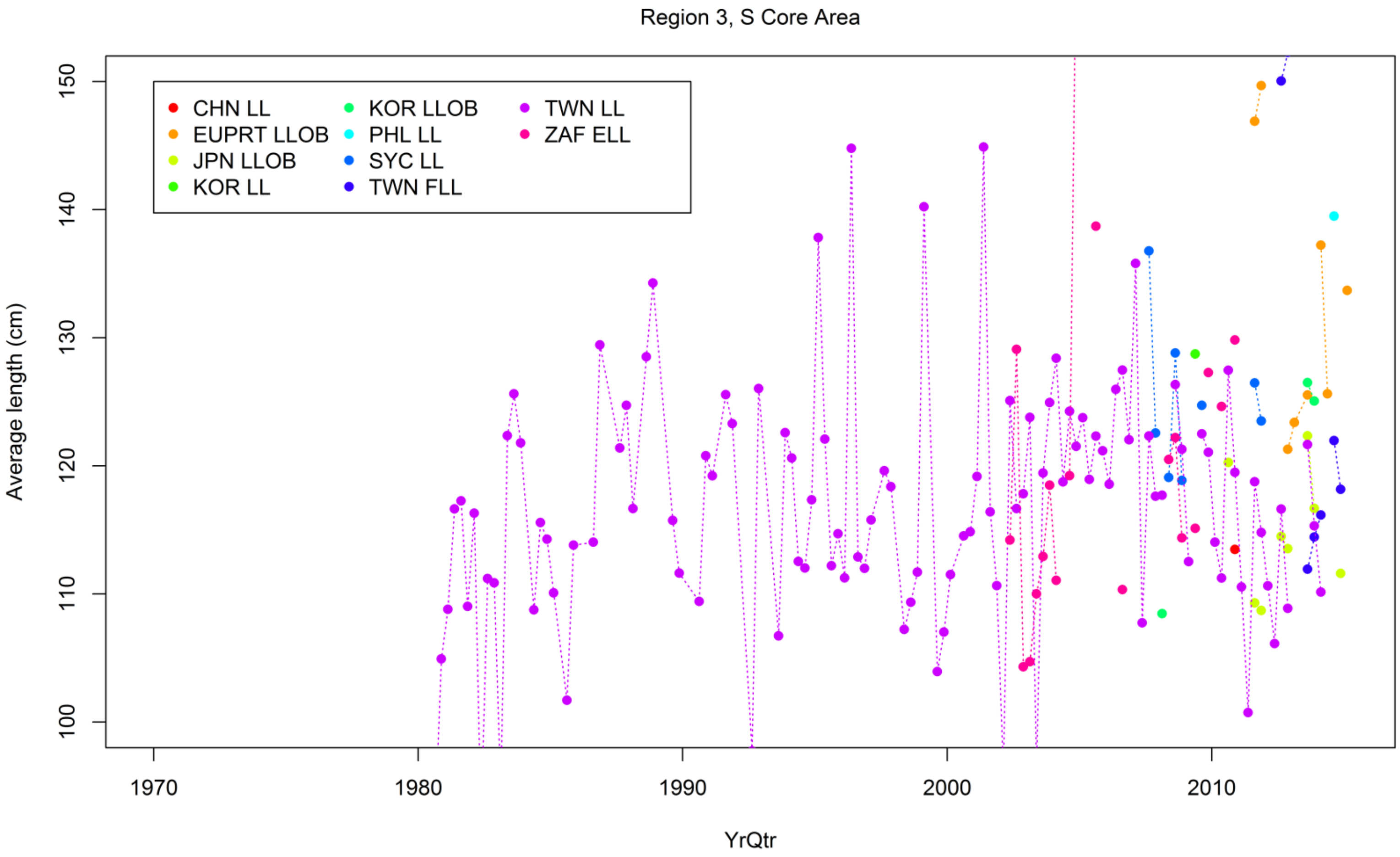

Figure A6: Average length of bigeye tuna (FL, cm) of fish sampled from the longline fishery in the core area of the southern region, by fleet nationality and year/quarter.

APPENDIX 2. SUMMARY RESULTS OF PRELIMINARY ASSESSMENT MODELLING.

Table A1. Summary of model options investigated during the preliminary phase of model development.

Model option	Specification	Rationale
Reference	Comparable to base model configuration described in the main section of the report, although some subtle differences: - Tag OD parameters fixed at 2.0 (not estimated). - No temporal split in PSLS selectivity. - Initial tag grouping age assignment.	
CPUE_downweightCV30	Increase CPUE CV from 0.2 to 0.3	Examine relative influence of the CPUE indices.
CPUE_LL1q	Estimate q deviates for 2008-2012 for LL1 CPUE indices.	To address conflict between LL1 and LL2 CPUE indices during that period (and tag data too). This effectively further down-weights the influence of the LL1 CPUE indices.
CPUE_old	CPUE indices from previous assessment. Limited to period from 1979.	Examine influence of the change in CPUE indices from Japanese series to combined CPUE indices from operational data.
LF_recentTWLL	Include recent (post 1997) TW LL LF data in the LL1-3 LF data sets. LL1-3 ESS retained at 1.0.	Examine influence of recent (post 1997) TW LL LF data.
LF_upweightLL	Increase max ESS for LL1-3 from 1 to 10. Does not include recent (post 1997) TW LL LF data.	Examine relative influence of the LL1-3 LF data.
region_recdevs	Estimate regional recruitment deviates for R1 for 2005-2012 enabling diverging trends in recent recruitment from R1 and R2.	Investigate whether divergent patterns in CPUE between R1 and R2 can be accomodated by variation in recruitment between the two regions.
region1	Single region model (including tag data set), all regional CPUE indices are retained in model.	Investigate influence of spatial configuration.
region2	Two region model combining regions 1 & 3 as a single region and retaining region 2. All CPUE indices retained in model.	Investigate influence of spatial configuration.
region2_cpue	As above but excluding CPUE indices from R3 (southern).	Investigate influence of spatial configuration.
Tag_downweight10	Reduce tag data releases and recoveries by n/10	Down weighting of the tag data set to investigate influence of these data.

Tag_excludeLike	Exclude tag components from total likelihood	Examine influence of tag data set.
Tag_MixQ2	Reduce the tag mixing period from 4 quarters to 2 quarters.	Examine influence of tag mixing period.
Tag_MixQ6	Increase the tag mixing period from 4 quarters to 6 quarters.	Examine influence of tag mixing period.
Tag_OD50	Change tag OD parameter from 2 to 50	Investigate behaviour of OD parameter.
Tag_ODestimated	Estimate the individual tag OD parameters for each tag group.	Investigate behaviour of OD parameter.
Tag_PSonly	Exclude tag recoveries from non PS fisheries an set RR for those fisheries = 0.	Examine influence of tag recoveries from fisheries for which there are no reliable estimates of tag reporting rate (especially LL).
Tag_ReviseAgeGrps	Minor changes in the procedure to assign of tags releases to age classes (tag grps).	Improvement in data processing.
Select_PSLS_recent	Split the PSLS selectivity by time blocks pre 2007 and 2007 onwards.	Improve fit to the PSLS data in the more recent period (smaller fish).
Start1952	Commence model in 1952 assuming unexploited conditions. Recruitment deviation period equivalent to base model.	Investigate sensitivity to initial conditions (1952 vs 1975).

Table A2. Details of objective function components for the range of preliminary model options.

Model option	Total	Component					
		Survey	Length_comp	Tag_comp	Tag_negbin	Recruitment	Parm_priors
Reference	3,767.48	-297.65	1,263.39	1,384.60	1,389.80	-14.84	41.77
CPUE_downweightCV30	3,752.88	-303.30	1,252.29	1,389.56	1,390.85	-17.87	40.78
CPUE_LL1q	3,746.27	-319.99	1,262.13	1,383.45	1,392.56	-16.09	41.88
CPUE_old	3,722.37	-353.84	1,264.91	1,396.48	1,393.55	-20.89	41.67
LF_recentTWLL	3,823.26	-298.29	1,299.72	1,397.29	1,400.03	-18.00	41.98
LF_upweightLL	4,136.40	-293.89	1,620.74	1,386.60	1,398.28	-17.41	41.60
region_recdevs	3,763.00	-297.27	1,263.03	1,382.55	1,386.96	-18.75	41.63
region1	3,827.52	-295.03	1,251.77	1,450.22	1,421.68	-13.35	11.79
region2	3,716.56	-296.76	1,254.76	1,364.44	1,389.58	-14.61	18.74
region2_cpue	3,738.04	-275.76	1,258.91	1,365.46	1,387.06	-16.54	18.55
Tag_downweight10	1,698.16	-297.89	1,213.81	189.58	578.61	-24.37	37.87
Tag_excludeLike	890.21	-295.46	1,177.19	0.00	0.00	-28.96	36.19
Tag_MixQ2	5,351.41	-251.96	1,295.60	2,294.20	1,955.21	18.76	39.24
Tag_MixQ6	3,092.25	-298.93	1,240.85	1,069.59	1,061.38	-22.63	41.45
Tag_OD50	4,434.66	-277.01	1,279.68	1,448.99	1,956.31	-16.33	42.46
Tag_ODestimated	3,875.08	-293.06	1,255.70	1,396.76	1,440.50	-18.22	92.91
Tag_PSonly	2,671.86	-300.47	1,255.96	485.82	1,208.38	-17.62	39.24
Tag_ReviseAgeGrps	4,050.07	-298.94	1,270.28	1,386.68	1,665.61	-15.99	41.93
Select_PSLS_recent	3,739.41	-282.01	1,233.77	1,386.88	1,370.55	-10.45	40.20
Start1952	3,767.28	-297.73	1,263.45	1,384.61	1,389.71	-14.96	41.79

Table A3. Estimates of management quantities for the preliminanry set of models.

Model	SB_0	SB_{MSY}	SB_{MSY}/SB_0	SB_{2015}	SB_{2015}/SB_0	SB_{2015}/SB_{MSY}	F_{2015}/F_{MSY}	MSY
Reference	946,538	263,458	0.278	308,185	0.33	1.17	0.74	104,050
CPUE_downweightCV30	1,055,770	284,471	0.269	368,111	0.35	1.29	0.64	113,970
CPUE_LL1q	953,590	264,958	0.278	307,972	0.32	1.16	0.74	105,359
CPUE_old	1,039,380	284,072	0.273	384,148	0.37	1.35	0.62	114,200
LF_recentTWLL	989,134	274,610	0.278	337,893	0.34	1.23	0.69	108,050
LF_upweightLL	993,755	275,158	0.277	330,776	0.33	1.20	0.68	109,716
region_recdevs	958,657	266,534	0.278	309,093	0.32	1.16	0.73	105,561
region1	893,773	254,392	0.285	270,887	0.30	1.06	0.79	101,668
region2	947,277	263,624	0.278	302,786	0.32	1.15	0.71	104,706
region2_cpue	920,873	254,332	0.276	261,819	0.28	1.03	0.82	99,573
Tag_downweight10	1,016,420	277,026	0.273	345,519	0.34	1.25	0.66	112,302
Tag_excludeLike	1,820,340	402,251	0.221	905,841	0.50	2.25	0.35	228,858
Tag_MixQ2	880,981	225,224	0.256	310,021	0.35	1.38	0.77	95,551
Tag_MixQ6	989,295	274,532	0.278	323,936	0.33	1.18	0.69	110,435
Tag_OD50	843,451	232,466	0.276	234,538	0.28	1.01	0.87	103,142
Tag_ODestimated	921,848	255,954	0.278	285,896	0.31	1.12	0.76	105,649
Tag_PSonly	1,002,270	279,802	0.279	352,677	0.35	1.26	0.66	109,312
Tag_ReviseAgeGrps	974,786	271,095	0.278	338,033	0.35	1.25	0.70	105,577
Select_PSLS_recent	930,063	254,337	0.273	307,152	0.33	1.21	0.80	99,533
Start1952	946,713	263,517	0.278	307,689	0.33	1.17	0.74	104,089

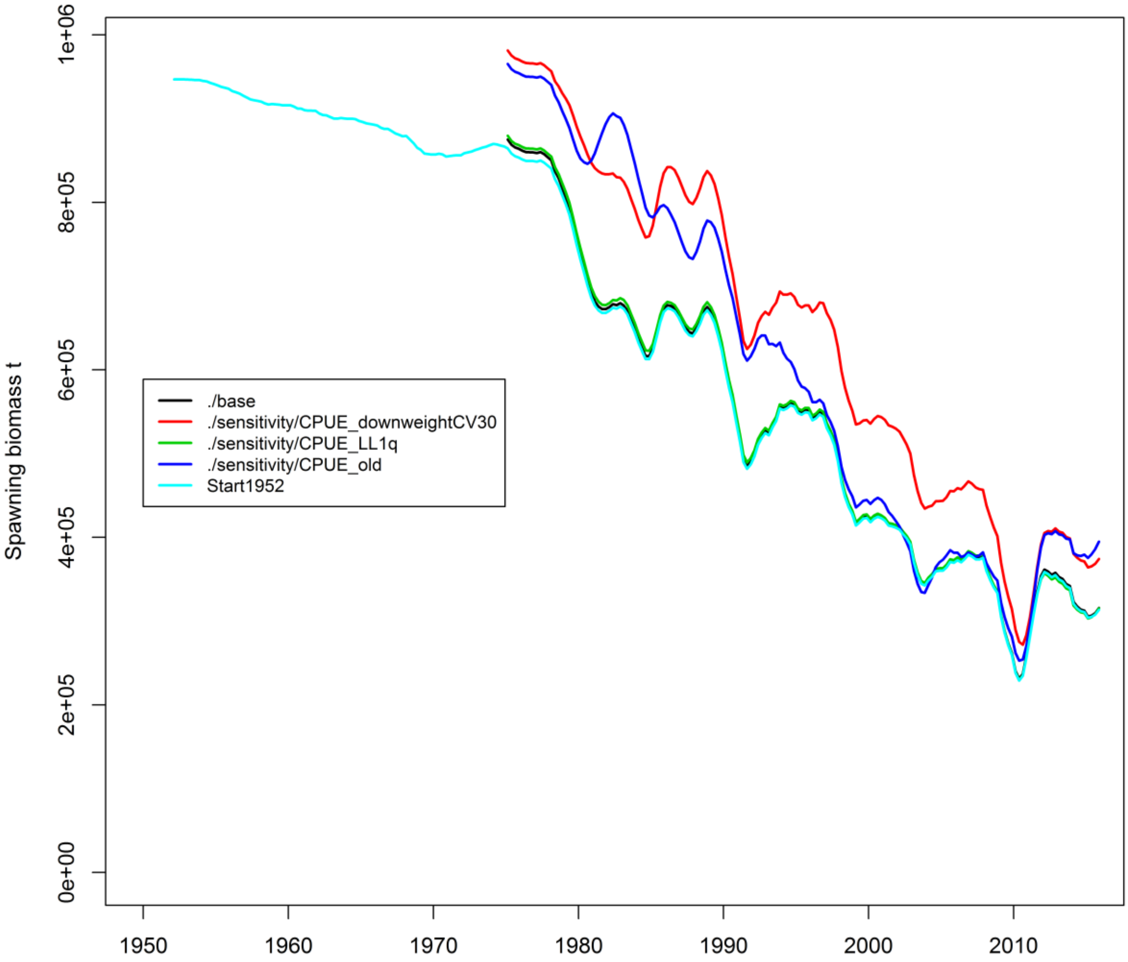

Figure A7. Spawning biomass trajectories for the range of preliminary model options investigating different longline CPUE scenarios. For comparison, the reference model (base) and the Start1952 model trajectories are also plotted.

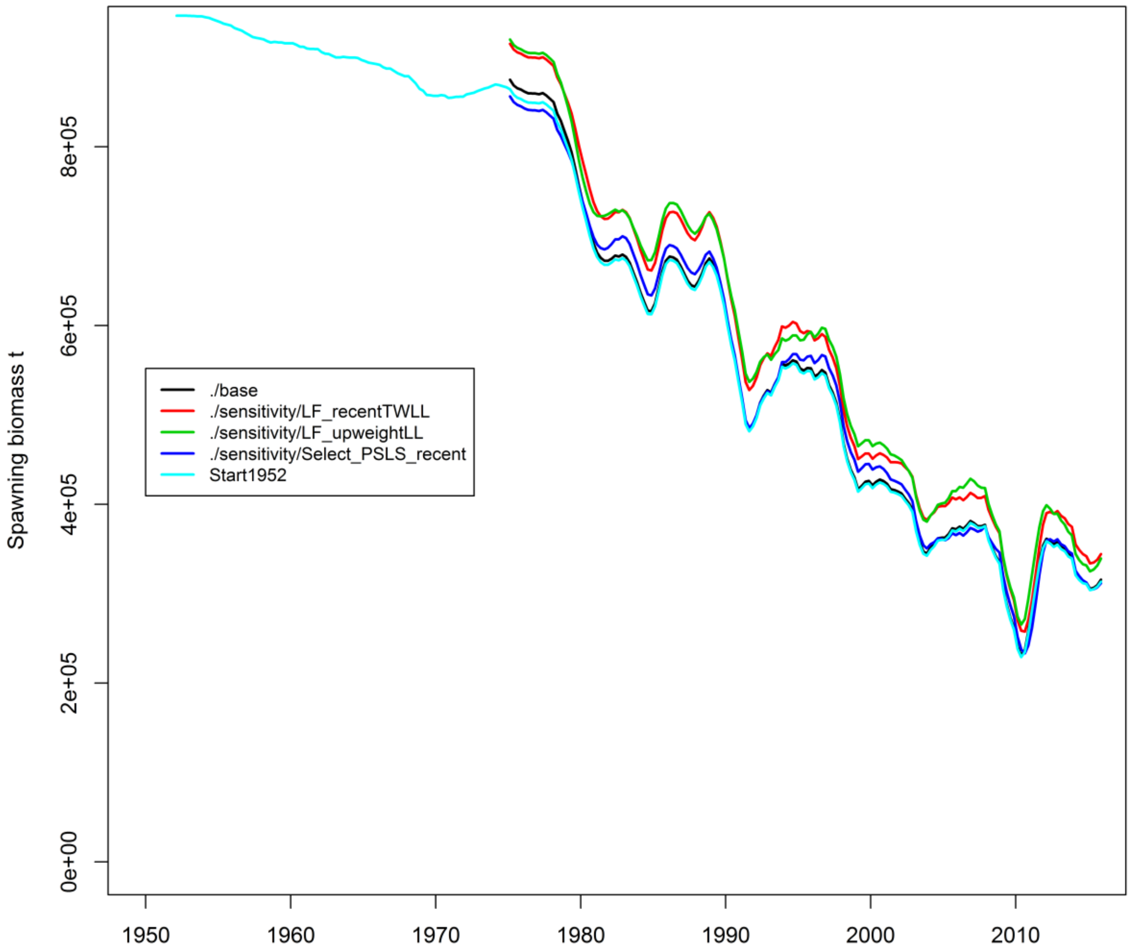

Figure A8. Spawning biomass trajectories for the range of preliminary model options investigating different scenarios for the treatment of the length frequency data sets. For comparison, the reference model (base) and the Start1952 model trajectories are also plotted.

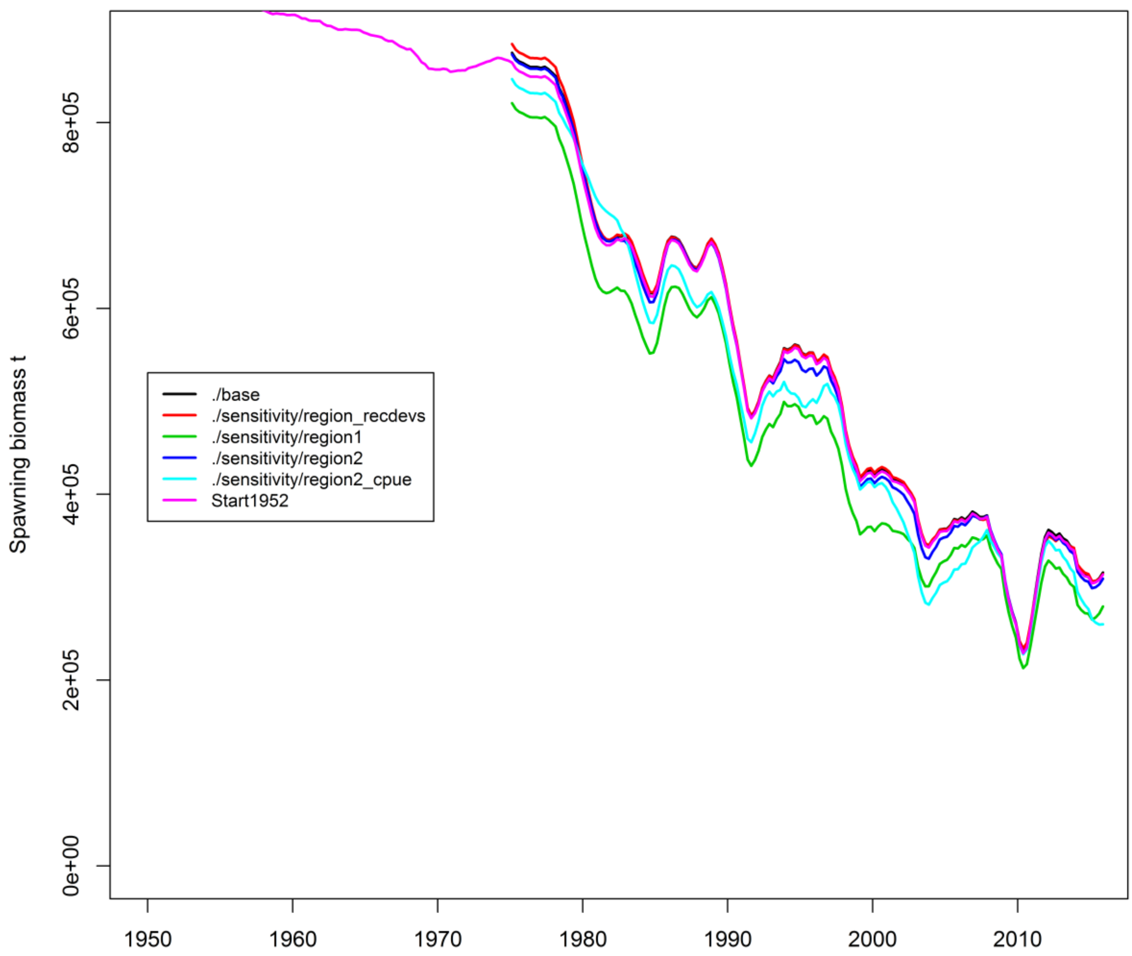

Figure A9. Spawning biomass trajectories for the range of preliminary model options investigating different regional configurations. For comparison, the reference model (base) and the Start1952 model trajectories are also plotted.

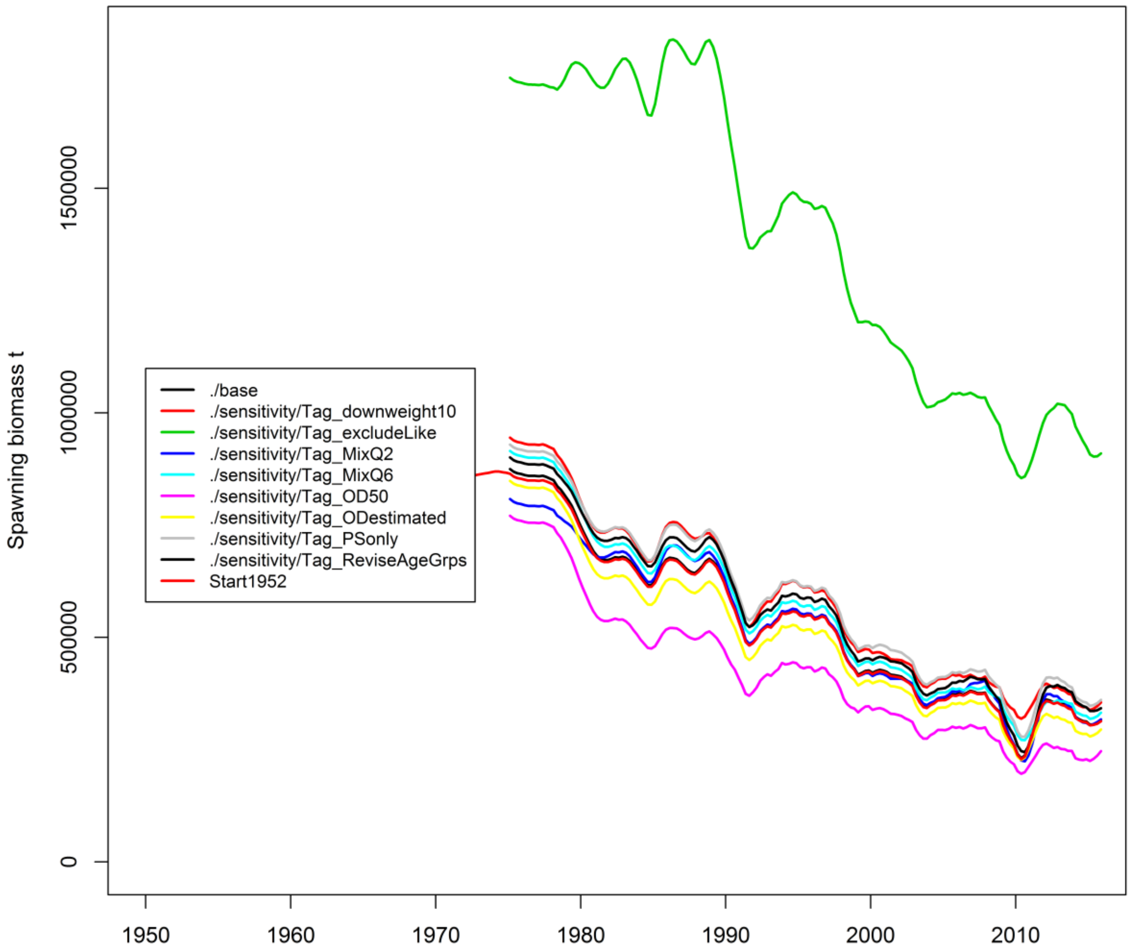

Figure A10. Spawning biomass trajectories for the range of preliminary model options investigating different treatment of the tag release/recovery data and assumptions. For comparison, the reference model (base) and the Start1952 model trajectories are also plotted.