SCIENTIFIC COMMITTEE
ELEVENTH REGULAR SESSION
Pohnpei, Federated States of Micronesia
5–13 August 2015

Stock assessment for south Pacific albacore tuna

WCPFC-SC11-2015/SA-WP-06
Rev 1 (4 August 2015)

S J Harley[1], N Davies[2], L Tremblay-Boyer[1], John Hampton[1], and S McKechnie[1]

[1]Oceanic Fisheries Programme, Secretariat of the Pacific Community
[2]Te Takina Ltd

Contents

Revision 1: 4 August 2015

The previous version incorrectly reported that two values of the tag reporting rate were included in the structural uncertainty grid and that the grid comprised 72 model runs.

This was for an earlier version of the assessment before we completed the correction to the tag releases as was the practice for the 2014 tropical tuna assessments. We have removed these incorrect details from Table 4 and Section 5.2.

Executive Summary

This paper describes the 2015 stock assessment of south Pacific albacore tuna (*Thunnus alalunga*) – the first assessment since 2012 (Hoyle et al., 2012). There have been many developments since the last assessment in terms of both the fishery and the integrated stock assessment model known as MULTIFAN-CL which is used to assess this stock. The current stock assessment includes much new data and new features reflecting recommendations from previous south Pacific albacore tuna assessments as well as relevant recommendations from the review of the 2011 bigeye tuna assessment (Davies et al., 2015).

This assessment is supported by the analysis of operational longline data to construct both the CPUE time series (Tremblay-Boyer et al., 2015b) and regional weights (Tremblay-Boyer et al., 2015a) and the analysis of longline size data (Scott and McKechnie, 2015). Finally the assessment includes results from a wide-scale study of the biological parameters of albacore (Williams et al., 2012; Farley et al., 2013b) – in particular results from the age and growth study aimed to address uncertainty around growth which has troubled previous assessments.

The main developments in the 2015 assessment are described in Table 1. The three most significant changes are: (1) the use of a spatially explicit model covering the southern region of the WCPFC Convention area; (2) the inclusion of direct age-length observations and tagging data from the 2009-10 releases; and (3) changing natural mortality from 0.4 to 0.3 per annum for consistency with albacore stock assessments conducted elsewhere.

The major structural changes (e.g.,the spatial and fishery structures) to the assessment mean that full consideration of the impacts of individual changes from the 2012 assessment is not possible. However, generally the results and main conclusions of the current assessment are similar to those from the 2012 assessment.

In addition to a single reference case model which we present here, we report the results of "one-off" sensitivity models to explore the impact of key data and model assumptions for the reference case model on the stock assessment results and conclusions. We also undertook a structural uncertainty analysis (grid) for consideration in developing management advice where all possible combinations of those areas of uncertainty from the one-off models were included.

The main conclusions of the current assessment are consistent with the previous assessment conducted in 2012. The main conclusions based on results from the reference case model and with consideration of results from performed sensitivity model runs, are as follows:

1. The new regional structure used for the 2015 assessment is better aligned with those of the assessments for bigeye and yellowfin tunas and provides an improved basis for further development of this assessment and providing advice to WCPFC;

2. There is some conflict between some of the data sources available for this assessment including

conflicts between the length-frequency data and the CPUE series and between the troll length frequency samples and the age-length data;

3. Current catch is either at or less than MSY;

4. Recent levels of spawning potential are most likely above the level which will support the MSY, and above $20\% SB_{F=0}$;

5. Recent levels of fishing mortality are lower than the level that will support the MSY;

6. Increasing fishing mortality to F_{MSY} levels would require a significant increase in effort, yield only very small (if any) increases in long-term catch, and would greatly reduce the vulnerable biomass available to the longline fleet;

7. Recent levels of spawning potential are lower than candidate bio-economic-related target reference points currently under consideration for south Pacific albacore tuna; and

8. Stock status conclusions were most sensitive to alternative assumptions regarding the weighting off different data sets and natural mortality, identifying these as important areas for continued research.

1 Introduction

This paper presents the 2015 stock assessment of south Pacific albacore tuna (*Thunnus alalunga*) covering the southern hemisphere component of the Western and Central Pacific Fisheries Commission (WCPFC) convention area and fisheries for the period 1960–2013. Since 1999, the stock has been assessed regularly and the most recent assessments are documented in Hoyle et al. (2008b); Hoyle and Davies (2009); Hoyle (2011), and Hoyle et al. (2012).

The overall objectives of the assessment are to estimate population parameters, such as time series of recruitment, biomass and fishing mortality, which indicate the stock status and fishing impacts. We summarize the stock status in terms of reference points adopted or under consideration by the WCPFC. The methodology used for the assessment is commonly known as MULTIFAN-CL[3] (Fournier et al., 1998; Hampton and Fournier, 2001; Kleiber et al., 2014). MULTIFAN-CL is a software program that implements a size-based, age- and spatially-structured population model. Model parameters are estimated by maximizing an objective function, consisting of both likelihood (data) and prior information components.

This assessment report should not be seen as a standalone document and should be read in conjunction with several supporting papers, specifically this assessment is supported by an analysis of operational longline data to construct both CPUE time series (Tremblay-Boyer et al., 2015b) and regional weights (Tremblay-Boyer et al., 2015a) and the analysis of longline size data (Scott and McKechnie, 2015). The assessment also includes results from a wide-scale study of the biological parameters of albacore (Williams et al., 2012; Farley et al., 2013b) – in particular results from the age and growth study aimed to address uncertainty around growth which has troubled previous assessments. This is the first MULTIFAN-CL stock assessment that includes conditional age-length observations with the aim of improving growth estimates.

2 Background

2.1 Biology and ecology

Albacore tuna comprise a discrete stock in the South Pacific (Murray, 1994). Mature albacore – above a minimum fork length (FL) of about 80 cm - spawn in tropical and sub-tropical waters between latitudes 10° S and 25° S during the austral summer (Ramon and Bailey, 1996). Juveniles are caught in surface fisheries in New Zealand's coastal waters, and in the vicinity of the sub-tropical convergence zone (STCZ, at about 40° S) in the central Pacific, about one year later at a size of 45–50 cm FL.

From this region, albacore appear to gradually disperse north, but may migrate seasonally between

[3] http://www.multifan-cl.org

tropical and sub-tropical waters. These seasonal migrations have been inferred from monthly trends in longline catch rates in sub-equatorial waters (Langley, 2004). Catch rates in sub-equatorial waters peak during December-January and May-July, indicating that albacore migrate south during early summer, and north during winter. This movement tends to correspond with the seasonal shift in the 23–28° C sea surface temperature isotherm location.

Daily otolith growth increments indicate that initial growth is rapid, with albacore reaching 45–50 cm (FL) in their first year (Leroy and Lehody, 2004; Williams et al., 2012). Subsequent growth is slower, at approximately 10 cm per year from ages 2–4, declining thereafter (Williams et al., 2012). Maximum recorded length is about 120 cm (FL). Recent analyses of length-at-age from otolith data have identified important patterns in south Pacific albacore growth (Williams et al., 2012; Farley et al., 2013b). Males grow to larger sizes than females, and their lengths-at-age start to diverge above about 85 cm, the length at sexual maturity (Figure 9). Lengths at age of both sexes also vary with longitude, with both growth rates and maximum sizes increasing toward the east and reaching a maximum at about 160° W.

The instantaneous natural mortality rate is believed to be between 0.2 and 0.5 per year, with significant numbers of fish reaching 10 years or more. Currently, the longest period at liberty for a recaptured tagged albacore in the South Pacific is 11 years, but in the North Pacific (the same species, but viewed a separate biological stock) there has been one recapture of 15 years (ISC Albacore Working Group, 2011).

2.2 Fisheries

Distant-water longline fleets of Japan, Korea, Chinese Taipei, and China, and the domestic longline fleets of a number of Pacific Island countries, catch adult albacore over a large proportion of their geographic range (Figure 5). The Chinese Taipei fleet in particular have targeted albacore consistently since the 1960s. Since the mid-1990's, the longline catch has increased considerably with the development (or expansion) of small-scale longline fisheries targeting albacore in several Pacific Island countries, notably American Samoa, Cook Islands, Fiji, French Polynesia, New Caledonia, Samoa, Tonga, and Vanuatu. The last few years have seen a further increase in longline catch. A troll fishery for juvenile albacore has operated in New Zealand's coastal waters since the 1960s and in the central Pacific (in the region of the STCZ) since the mid-1980s. Driftnet vessels from Japan and Chinese Taipei targeted albacore in the central Tasman Sea and in the central Pacific near the STCZ during the 1980s and early 1990s (Figure 5). Surface fisheries are highly seasonal, occurring mainly from December-April. Longline fisheries operate throughout the year, although there is a strong seasonal trend in the catch distribution, with the fishery operating in southern latitudes (south of 35° S) during late summer and autumn, moving northwards during winter.

The South Pacific albacore fishery was slow to develop, with catch fluctuating at low levels from the 1960s through to the late 1990s. Post-2000 catch increased to over 60,000 mt, and subsequently to over 80,000 mt (Figure 3). The longline fishery harvested most of the catch, about 25,000–30,000 mt per year on average, prior to about 1998. The increase in longline catch to approximately 70,000 mt in 2005 was due to the development of small-scale longline fisheries in Pacific Island Countries and Territories, and a recent increase in the numbers of fish caught is also apparent in the Chinese and Chinese Taipei longline fisheries. Catch from the troll fishery are relatively small, generally less than 10,000 mt per year. The driftnet catch reached 22,000 mt in 1989, but has since declined to zero following a United Nations moratorium on industrial-scale drift-netting.

3 Data compilation

Data used in this South Pacific albacore assessment consist of fishery-specific catch, effort and length-frequency data, tag release-recapture data, and conditional age-length observations. Details of these data and their stratification are described below.

3.1 Spatial stratification

Hoyle et al. (2012) provides a detail history of the various spatial structures considered in south Pacific albacore tuna assessments so here we will just focus on the approach used in the current assessment and how this differs to that of the 2012 assessment (Hoyle et al., 2012).

The geographic area encompassed in the 2012 assessment was the Pacific Ocean south of the equator, from 140° E to 70° W. The model had a single region for the purpose of mixing and availability, but had regionally defined fisheries based on a latitudinal split at 25° S and longitudinal splits at 180° and 110° W.

In considering the spatial structure for the current assessment we attempted to strike a balance between the biology and ecology of south Pacific albacore tuna and the management needs of the WCPFC, including development of reference points, harvest control rules, and bio-economic modelling. The eight region spatial structure for the assessment is provided in Figure 1: it maintains the latitudinal split at 25° S, but includes a further latitudinal split at 10° S to separate off the predominately tropical longline fishery; has longitudinal splits at 170° E and 150° W for consistency with the bigeye tuna and yellowfin tuna assessments and 110° W; and has two regions covering the overlap area between the Inter-American Tropical Tuna Commission and WCPFC convention areas – again split at 25° S. Given the unique shape of the overlap area just south of the equator, it was considered unnecessary to have a very small separate area north of 10° S.

One implication of the exclusion of the eastern Pacific Ocean is that all reference points will be relevant for the southern WCPFC convention area. A further consideration was the paucity of operational longline data for the far east region due to one CMM not making available their CPUE data for the stock assessment. Good spatial coverage of CPUE data is necessary for estimating the "regional weights" that are used in MULTIFAN-CL assessments. These weights reflect the relative abundance of albacore tuna in each model region (see Section 3.4.1).

3.2 Temporal stratification

The time period covered by this assessment is the first quarter of 1960 to final quarter of 2013. Within this period, data were compiled into quarters (Jan-Mar, Apr-Jun, Jul-Sep, and Oct-Dec). This differs from the annual time-step assumed for the previous assessment.

As was done for the tropical tuna assessments in 2014 we excluded data for the most recent year as provisional estimates for the catch and effort data for longline fisheries (available in time for the assessments) have generally been subject to significant revision either during or shortly after the completion of the assessment.

3.3 Definition of fisheries

MULTIFAN-CL requires all catch and effort to be allocated to "fisheries". Ideally, the fisheries are defined to have selectivity and catchability characteristics that do not vary greatly over time. For most pelagic fisheries assessments, fisheries can be defined according to gear type, fishing method and region.

The 2012 south Pacific albacore tuna assessment considered several flag-related longline fisheries, but for simplification in the current assessment, as we have also implemented spatial structure not present in the 2012 assessment, we restricted ourselves to a single fishery for each gear type in each region (where appropriate, e.g., surface fisheries were only included for the southern regions). So the current assessment included eight longline fisheries, and three each of driftnet and troll fisheries.

Details of the flags, and their respective catches within each longline fishery are provided in detail in Tremblay-Boyer et al. (2015a). The geographic distribution of the recent catch is presented in Figure 5.

3.4 Catch and effort data

Catch and effort data were compiled according to the fisheries defined in Table 3. All catches were expressed in numbers of fish, with the exception of the driftnet fishery, where catches were expressed in weight (metric tonnes). For longline fisheries, effort was standardized as described below, while for troll and driftnet fisheries, the number of vessel days of fishing activity was used.

Overall annual catches by gear type, and then further broken down by region are provided in Figures 3 and 4. We can see that catch has been increasing in recent years as mentioned in Section 2.2 and that most of the recent increase has occurred in regions 2 and 5, the area almost exclusively comprised of EEZs between 10-25° S and west of 150° W. We can see that the troll catches are exclusively in the southern regions, primarily region 6.

3.4.1 Longline effort and CPUE

Full details of the analyses undertaken to calculate the standardized CPUE indices and regional weights used in this assessment are provided in Tremblay-Boyer et al. (2015b,a) and only briefly summarized here.

Available operational longline catch and effort data[4] was separated into the different regions and then clustering methods were applied based on catches of albacore tuna, bigeye tuna, yellowfin tuna, and swordfish to identify different targeting behaviour. On a region by region basis we examined the the results of the clustering and selected catch and effort records from particular clusters (predominantly albacore tuna dominant clusters) to include the subsequent standardization.

The final indices used in the assessment were based on negative binomial regression models incorporating vessel effects and cluster as a factor (where data for multiple clusters were included). Temporal CV's were estimated by using the canonical method of Francis (1999) which estimates CV's for all time periods in the standardization model, including the reference period. These CV's underestimate the variation represented in the standardized effort component of MULTIFAN-CL's likelihood and so they were rescaled for each region separately so that they had a mean of 0.2 over the time periods 1998–2012 (the period for which CPUE was available for all regions).

The analysis to derive the regional weights was also based upon the analysis of operational longline catch and effort data. The aim of the analysis was to standardize the data in such a way that spatial differences in CPUE across the entire model domain reflects differences in relative abundance. The main differences to the analysis used to generate the CPUE indices was restriction to albacore tuna targeting clusters, inclusion of information on thermocline depth across the model domain, and using various methods to interpolate spatial regions where data was either uncertain or missing.

The final regional weights were based on a spatial surface generated from a time-aggregated data set of all available operational data where interpolation of missing squares was not used. We also used similar approach, but restricted to operational data since 2008 (the most recent increase in catches), but this analysis had much poorer spatial coverage and was therefore considered less reliable, but included in the structural sensitivity analysis described in Section 5.2.

3.5 Size data

Available length-frequency data for each of the defined fisheries were compiled into 100, 1-cm size classes (30–129 cm). Data were collected from a number of sources, and can be summarized as

[4] the data set was similar to that used for the Pacific-wide analysis of McKechnie et al. (2015) except Japan did not authorize the inclusion of any data held only by them (i.e., not otherwise held by SPC) in the analysis

follows.

3.5.1 Longline

Albacore catch size composition data have been routinely collected from the fishery since the early 1960's. These data are characterized by inconsistent temporal and spatial resolution, may be subject to very small sample sizes and consequently exhibit high variability in some periods. In addition the spatial and temporal distribution of length sampling has not always reflected that of the fishery. As a consequence, some samples may have very different size compositions to the majority of the catch and are representative only of periods and locations where very small catches have been made. Further details of the breakdown by flag and time of length-frequency samples for different fisheries is provided in Tremblay-Boyer et al. (2015a) and Scott and McKechnie (2015).

In previous assessments, a data re-weighting approach has been applied to improve the consistency of the size frequency data and to ensure that it is as representative as possible of catches across the full spatial and temporal extent of the fishery. Following the recommendations of the external review of the bigeye assessment (Ianelli et al., 2012), a revised re-weighting method (McKechnie, 2014) was developed and applied for the bigeye and yellowfin assessments conducted in 2014. The revised method re-weights the size composition data according to the proportion of temporally smoothed catch taken within each 10° x 20° spatial cell within a region in a given year-quarter. In addition, a minimum weighting threshold was imposed on the lowest allowable weighting to reduce the influence of size data from cells with very little catch. Re-scaled length frequency data based on an 11 year-quarter moving average for catch scaling and a minimum weighting threshold of 0.1 were used for the assessment (Scott and McKechnie, 2015).

3.5.2 Troll and other surface fisheries

New Zealand domestic troll data (fisheries in regions 3 and 6) were collected from port sampling programmes conducted by the Ministry of Fisheries and, more recently, the New Zealand National Institute of Water and Atmospheric Research (NIWA).

Length-frequency data from troll fishing operations in the sub-tropical convergence zone (STCZ) (fisheries 6 and 8) were collected and compiled through the Albacore Research Tagging Project (1991-1992) and by port sampling programmes in Levuka, Fiji; Pago Pago, American Samoa; and Papeete, French Polynesia; and, during the 1990–1991 and 1991–1992 seasons, by scientific observers.

Driftnet data were provided by the NRIFSF for Japanese driftnet vessels. Data from Japanese vessels were also collected by observers and by port sampling in Noumea, New Caledonia. It is assumed that these data are representative of all driftnet activity.

3.6 Tagging data

Limited tagging data were available for incorporation into the assessment. Data consisted of tag releases and returns from a South Pacific Albacore Research Group tagging programme in the mid-1980s and OFP albacore tagging programmes conducted during the austral summers of 1990–1992 and 2009–2010 (Figure 12; from Hoyle et al., 2012). Albacore were captured primarily by trolling and tagged using standard tuna tagging equipment and techniques by trained scientists and scientific observers. During the 1980s and 1990s, the majority of tag releases were made by scientific observers on-board New Zealand and US troll vessels fishing in New Zealand waters and in the central South Pacific sub-tropical convergence zone region. The more recent tagging was conducted in New Zealand waters.

For the albacore assessment, tag releases were stratified by release region (all albacore releases occurred in regions 3, 6 and 8), time period of release (quarter) and the same size classes used to stratify length-frequency data. Releases were classified into 32 tag release groups (region, year and quarter). Release numbers were modified to account for returns that could not be classified to recapture fisheries and/or time periods, and for tagging-related effects on the survival of tagged fish (see Section 4.3). Following adjustment, the total effective release numbers were 4,280. Returns from each size class of each tag release group (140 tag returns in total) were then classified by recapture fishery and recapture time period (quarter).

Tag releases principally comprised juvenile fish (aged 1–4 years); few fish larger than 80 cm (FL) were tagged (Figure 13; from Hoyle et al., 2012). The length composition of fish from tag recoveries was comparable to the length at release, albeit slightly larger, allowing for growth during the period at liberty. Many of the tag recoveries were from longline fisheries in the southern regions (3, 6, 8) (Figure 13; from Hoyle et al., 2012). Relatively few tags were returned from the troll fisheries. Most tag recoveries occurred during the five years following release although there were several in excess of nine years after release.

3.7 Conditional length at age data

Observations from otolith readings were available from an ageing study of south Pacific albacore by Farley et al. (2013a) which comprised $n = 1969$ ages-at-length (males, females and unknown, combined), with age expressed as the decimal year. These observations were stratified according to the fishery definitions, spatial and temporal structures used in the assessment model, based upon the individual sample details (method, flag, date and latitude/longitude). Samples were collected over the years 2009 and 2010, but were aggregated into the single year of 2010, as no inter-annual variation in growth was to be considered in the model (constant growth is assumed in MULTIFAN-CL). This produced 21 fishery-quarter samples (6 longline fisheries, and 2 troll fisheries), from which only those containing more than 50 age-length observations were retained, resulting in 14 samples.

The two samples from troll fisheries were subsequently excluded from input to the model because of inconsistencies detected with the modal structure of the troll length composition data (this is discussed further in the Discussion section). The ages expressed in decimal years were translated into quarters to be consistent with the temporal structure assumed in the model. The 12 samples made up a total of $n = 1580$ age-length observations input to the model.

4 Model description

The model can be considered to consist of several components, (i) the dynamics of the fish population; (ii) the dynamics of the fisheries; (iii) the dynamics of tagged fish; (iv) observation models for the data; (v) parameter estimation procedure; and (vi) stock assessment interpretations. Detailed technical descriptions of components (i) – (iv) in respect of the MULTIFAN-CL modelling software are given in Hampton and Fournier (2001) and Kleiber et al. (2014), and are not repeated here.

4.1 Population dynamics

The model partitions the population into 8 spatial regions (see Section 3.1) and 48 quarterly age-classes. The last age-class comprises a "plus group" in which mortality and other characteristics are assumed to be constant. The population is "monitored" in the model at quarterly time steps, extending through a time window of 1960–2013. The main population dynamics processes are as follows:

4.1.1 Recruitment

Recruitment in terms of the MULTIFAN-CL model is the appearance of age-class 1 fish in the population. In previous South Pacific albacore assessments that incorporated spatial structure, albacore recruitment was assumed to occur in the southern-most regions, where juvenile albacore are first caught in troll fisheries. However, recent spatial modelling using the SEAPODYM approach (Lehodey et al., 2012) indicated that a wider spatial distribution of young albacore was likely. We therefore changed the previous approach and allowed the proportion of albacore recruitment occurring in all model regions to be estimated using the SEAPODYM results as starting values.

In previous assessments, it was assumed that recruitment is an annual event that occurs in the summer months. For this assessment, we adopted a quarterly temporal structure for the model dynamics thus providing for quarterly estimates of recruitment. This was done mainly to better align the albacore assessment model with those for the tropical tunas, to better facilitate integrated management analyses across the species.

Annual recruitment was assumed to be related to spawning potential according to the Beverton-Holt stock-recruitment relationship (SRR). A weak penalty was applied to deviation of the annual recruitment from the SRR so that it would have a negligible effect on recruitment and other model estimates. Note that in previous MULTIFAN-CL assessments, SRR penalties were computed for each quarterly recruitment estimate (in quarterly recruitment models). The move to computing penalties based on annual recruitment was recommended by the 2011 Bigeye Tuna Peer Review (Ianelli et al., 2012) and this has been implemented for the first time in this assessment.

Typically, fisheries data are very uninformative about SRR parameters and it is generally necessary to constrain the parameterisation in order to have stable model behaviour. As in other recent tuna assessments, the "steepness" coefficient (h) of the SRR was fixed at a moderate value of 0.8, with h defined as the ratio of the equilibrium recruitment produced by 20% of the equilibrium unexploited spawning potential to that produced by the equilibrium unexploited spawning potential (Mace and Doonan, 1988). In other words, the prior belief is that when the equilibrium spawning potential is reduced to 20% of its unexploited level, equilibrium recruitment would be reduced to 80% of its unexploited level. Steepness values of 0.65 and 0.95 were considered as sensitivity analyses, and as part of the structural uncertainty grid.

4.1.2 Initial population

The population age structure in the initial time period in each region was assumed to be in equilibrium and determined as a function of the average total mortality during the first 20 quarters. This assumption avoids having to treat the initial age structure, which is generally poorly determined, as independent parameters in the model. The initial age structure was applied to the initial recruitment estimates to obtain the initial populations in each region.

We noticed in this assessment, that any seasonality in recruitment is not considered in the initial conditions (see Figure 29 upper left panel), but we do not believe that this has any impact on the model results as the population age-structure quickly departs from equilibrium conditions.

4.1.3 Growth

The standard assumptions made concerning age and growth are (i) the lengths-at-age are normally distributed for each age-class; (ii) the mean lengths-at-age follow a von Bertalanffy growth curve; (iii) the standard deviations of length for each age-class are a log-linear function of the mean lengths-at-age; and (iv) the probability distributions of weights-at-age are a deterministic function of the lengths-at-age and a specified weight-length relationship. These processes are assumed to be regionally invariant.

As noted above, the population is partitioned into quarterly age-classes with an aggregate class for the maximum age (plus-group). The aggregate age class makes possible the accumulation of old and large fish, which is likely in the early years of the fishery when exploitation rates were very low.

Recent tropical tuna assessments have allowed the mean lengths of some younger first eight age-classes to be independent parameters, with the remaining mean lengths following a von Bertalanffy growth curve. These deviations attract a small penalty to avoid over-fitting the size data. We examined this in the current assessment and found that the improvement in fit to the size data did not warrant the additional model parameters.

4.1.4 Movement

Movement was assumed to occur instantaneously at the beginning of each quarter via movement coefficients that connect regions sharing a common boundary. Note that fish can move between non-contiguous regions in a single time step due to the "implicit transition" computational algorithm employed (see Hampton and Fournier (2001) and Kleiber et al. (2014) for details). Movement is parameterised as the proportion of fish in a given region that move to the adjacent region. Across each inter-regional boundary in the model, movement is possible in both directions for the four quarters, each with their own movement coefficients. Thus the number of movement parameters is 2×no.regions×4quarters. The seasonal pattern of movement persists from year to year with no allowance for longer-term variation in movement. Usually there are limited data available to estimate age-specific movement and the movement coefficients are normally invariant with respect to age.

A prior of 0.1 is assumed for all movement coefficients, inferring a relatively high mixing rate between regions. A small penalty is applied to deviations from the prior. Evaluation of age-specific movement − both linear and nonlinear were considered during model development, but while improvements in model fit were obtained, they we not significant based on AIC and were not included in the final model runs.

4.1.5 Natural mortality

The previous assessment for south Pacific albacore tuna assumed a constant value for mean natural mortality (M) of 0.4 per annum, but for the current assessment M was fixed at 0.3 (Figure 8) per annum (or 0.3/4 per quarter) to be consistent consistent with other stock assessments for albacore tuna, e.g., ICCAT (Dr L Kell, *pers. comm.*) and the North Pacific assessment (ISC Albacore Working Group, 2011).

The higher natural mortality likely to occur for young fish is not included in the model. Previous analyses applying higher natural mortality for young fish have shown little effect on management parameters.

We consider M to be a key source of uncertainty and therefore examined values of 0.25 and 0.4 for sensitivity.

4.1.6 Maturity

The maturity-at-age schedule assumed in the model (Figure 9) was taken directly from the previous assessment (Hoyle et al., 2012). The calculation was based on analyses of biological data, since published, (Farley et al., 2013b,a, 2014). The ogive was based upon the relative reproductive potential of females rather than the relative biomass of both sexes above the age of female maturity. This corresponded closely with the ogive of the weighted average of the proportion of females mature-at-age provided by Farley et al. (2014). The weighted estimates of the length and age at 50% maturity were 87 cm FL and 4.5 years, respectively. Whereas the ogive used for the 2012 assessment expressed age in respect of "years", this was translated into ages in "quarters" for consistency with the quarterly temporal structure of the model used in this assessment. Ogive values for the intermediate quarters were obtained by interpolating between the annual values using the `na.spline` function of the R package `zoo()`.

4.2 Fishery dynamics

4.2.1 Selectivity

Selectivity is fishery-specific and assumed to be time-invariant and length-based to the extent that ages with similar lengths must have similar selectivities at age. The selectivities at age were estimated using a cubic spline parameterisation. Each selectivity function was parameterised with four nodes, allowing some flexibility in the functional form while minimizing the number of parameters required to be estimated. The estimated selectivities at age have a range of 0–1. All selectivities were constrained such that the selectivity of the last two age classes was equivalent. Selectivities coefficients for the longline fisheries were computed over age-class ranges of 5–48, for troll fisheries over ranges of 3-46 and for driftnet fisheries over ranges of 5–46. Coefficients outside these ranges were set to zero.

In the previous South Pacific albacore assessment (Hoyle et al., 2012), the albacore population was modelled in a single spatial region, requiring that seasonal selectivity for the longline fisheries be modelled to account for the movement-driven seasonal availability of different-sized albacore to the different fisheries. In the current assessment, spatial structure and movement is explicit, so this enabled non-seasonal selectivity for longline fisheries to be employed. Explicit spatial structure also allowed simplifying assumptions to be made regarding the form of the selectivity curve for the longline fisheries in each region. For each fishery, we assumed that the oldest albacore were fully recruited. To encourage this behaviour, selectivity coefficients were penalized to be non-decreasing for successively older age-classes. We initially tested the assumption that longline fisheries in all regions shared common selectivity coefficients. However, it was found that this resulted in severe lack of fit to the size data for several fisheries and therefore independent selectivity coefficients were allowed for each fishery.

4.2.2 Catchability

Catchability was assumed to be constant over time for all longline fisheries because the effort for these fisheries was derived from the standardized CPUE indices described in Tremblay-Boyer et al. (2015b).

As in the previous assessment the catchability for all other fisheries was allowed to vary over time (akin to a random walk) using a structural time-series approach. Random walk steps were taken annually, and deviations were constrained by a prior distribution of mean zero and a variance equivalent to a coefficient of variation (CV) of 0.7 on a log scale.

4.2.3 Effort deviates

Effort deviations are constrained by prior distributions having a mean of zero and a specified variance, and are used to model the random variation in effort (i.e. fishing mortality relation). Time varying variances were applied to the effort deviations. For fisheries with standardized CPUE, temporal penalties were adjusted to be proportional to the temporal CV's estimated in the CPUE standardization (Tremblay-Boyer et al., 2015b). We assumed that the CPUE observations for the period 1998–2012 (the period when standardized CPUE was available for all longline fisheries) had an average CV 0f 0.2. As can be seen from Figure 6, prior to this period most of the CPUE values are much less certain.

4.3 Dynamics of tagged fish

4.3.1 Tag reporting

Tag-reporting rates are estimated with relatively uninformative Bayesian priors, because little independent information is available. Reporting rates were allowed to vary across fisheries, but were assumed to be common across tag release groups. As noted in Section 4.3, we adjusted tag release numbers to account for tag returns that could not be included in the data because of insufficient recapture information. This is an important source of "non-reporting" particularly for models where the tag return data need to be stratified by spatial regions. This adjustment was made to preserve the observed rate of tag recapture by release group in the original data (Berger et al., 2014). The adjustment was made by release group, but overall 140 of 260 (54%) tag returns in total could be allocated to recapture fisheries in the model.

We also further adjusted the tag release numbers downwards by a factor of 0.5 to account for likely mortality of albacore due to the stress of capture, handling and tagging. Initial tagging mortality (being a type 1 tag loss) operates in a similar fashion to non-reporting and was therefore dealt with by adjusting the releases. The choice of 0.5 as the adjustment factor was somewhat arbitrary, but was influenced by recent work on tagger effects on tagged fish survival and associated correction factors in what are thought to be more robust tropical tunas (Berger et al., 2014). In that study it was found that the median correction factors for tagger effects were 0.68–0.76 for the tropical tunas. Given that albacore are believed to be more sensitive to capture and handling, it was felt that a stronger correction was appropriate in this case. We also tested a factor of 0.7 and found that this resulted in relatively minor impacts on model results.

4.3.2 Tag mixing

The population dynamics of the fully recruited tagged and un-tagged populations are governed by the same model structures and parameters. The populations differ in respect of the recruitment process, which for the tagged population is the release of tagged fish, i.e. an individual tag and release event is the recruitment for that tagged population. Implicitly, we assume that the probability of recapturing a given tagged fish is the same as the probability of catching any given untagged fish in the same region and time period. For this assumption to be valid either the distribution of fishing effort must be random with respect to tagged and untagged fish and/or the tagged fish must be randomly mixed with the untagged fish. The former condition is unlikely to be met because fishing effort is almost never randomly distributed in space. The second condition is also unlikely to be met soon after release because of insufficient time for mixing to take place. Depending on the disposition of fishing effort in relation to tag release sites, the probability of capture of tagged fish soon after release may be different to that for the untagged fish. It is therefore desirable to designate one or more time periods after release as "pre-mixed" and compute fishing mortality for the tagged fish based on the actual recaptures, corrected for tag reporting (see below), rather than use fishing mortalities based on the general population parameters. This in effect de-sensitises the likelihood function to tag recaptures in the pre-mixed periods while correctly discounting the tagged population for the recaptures that occurred. We assumed that albacore mix fairly slowly with the untagged population at the region level and that this mixing process is complete by the end of the fourth quarter after release.

4.4 Likelihood components

There are four data components that contribute to the log-likelihood function — the total catch data, the length-frequency data, the age-length data, and the tagging data. The observed total catch data are assumed to be unbiased and relatively precise, with the SD of residuals on the log scale being 0.007.

The probability distributions for the length-frequency proportions are assumed to be approximated by robust normal distributions, with the variance determined by the effective sample size and the observed length-frequency proportion.

The size frequency data are assigned an effective sample size lower than the number of fish sampled. Reduction of the effective sample size recognizes that (i) length-frequency samples are not truly random (because of clumping in the population with respect to size) and would have higher variance as a result; and (ii) the model does not include all possible process error, resulting in further under-estimation of variances. The relative weighting of the longline size frequency is comparable to the approach used in the 2014 tropical tuna assessments.

Ageing data from biological sampling has been formally included in this assessment to assist in estimating growth parameters because these provide direct observations of the distribution of fish ages within length classes. Typically, these observations, c_{ljm} are collected from a particular fishing method or fishery m using a sampling design stratified in respect of length, and which assumes the observations at age j are random within each length class l.

Using the normal distribution, the model growth function predicts mean lengths μ_j and standard deviations σ_j for each age class j. Then if length is l let

$$q_{lj} = \frac{e}{\sigma_j}^{\frac{-(\mu_j - l)^2}{2\sigma_j^2}}.$$

The predicted catch age composition, p_{jm}, (that takes account of the selectivity pattern) for fishery m is used to derive the predicted distribution of age-at-length for that fishery given the growth estimates of length-at-age:

$$\rho_{ljm} = \frac{p_{jm} q_{lj}}{\sum_k p_{km} q_{lk}}.$$

The observed age composition within each length interval is assumed to be multinomially distributed, and therefore the negative log-likelihood for length interval l is

$$-\sum_j \vartheta_m c_{ljm} \log(\rho_{ljm})$$

where ϑ_m is the effective sample size for fishery m, and the total for fishery m is summed among all length intervals. The total likelihood for an age-length is the sum over the length intervals in the sample.

A log-likelihood component for the tag data was computed using a negative binomial distribution. The negative binomial is preferred over the more commonly used Poisson distribution because tagging data often exhibit more variability than can be attributed by the Poisson. We have employed a parameterisation of the variance parameters such that as they approach infinity, the negative binomial approaches the Poisson. Therefore, if the tag return data show high variability (for example, due to contagion or non-independence of tags), then the negative binomial is able to recognize this. This should then provide a more realistic weighting of the tag return data in the overall log-likelihood and allow the variability to impact the confidence intervals of estimated parameters. However, early attempts at estimating fishery-specific variance parameters from the data yielded values at either bound, suggesting insufficient information was available. A fixed value at the mid-point of the variance range was therefore assumed for all fisheries. Stock assessment results were relatively insensitive to the choice of the variance level.

4.5 Parameter estimation and uncertainty

The parameters of the model were estimated by maximizing the log-likelihoods of the data plus the log of the probability density functions of the priors and smoothing penalties specified in the model. The maximization to a point of model convergence was performed by an efficient optimization using exact derivatives with respect to the model parameters (auto-differentiation, Fournier et al., 2012). Estimation was conducted in a series of phases, the first of which used arbitrary starting values for most parameters. A bash shell script, `doitall`,implements the phased procedure for fitting the model. Some parameters were assigned specified starting values consistent with available biological information. The values of these parameters are provided in the `alb.ini` file[5].

In this assessment only one approach, a structural uncertainty analysis, was used to describe the uncertainty in key model outputs. We did not undertake the Hessian calculation for the current assessment – it is extremely time consuming and only including in two figures to illustrate uncertainty in spawning potential and recruitment time series.

For the structural uncertainty analysis, a crosswise grid of model runs was undertaken which incorporated many of the options of uncertainty explored by the key model runs and one-off sensitivity analyses. This procedure attempts to describe the main sources of structural and data uncertainty in the assessment.

For highly complex population models fitted to large amounts of often conflicting data, it is common for there to be difficulties in estimating absolute abundance (Lee et al., 2014). Therefore, a profile likelihood analysis was done of the marginal posterior likelihood in respect of the total population scaling parameter. Reasonable contrast in the profile was taken as indicating sufficient information existed in the data for estimating absolute abundance, and also offered confirmation of the global minimum obtained by the maximum likelihood estimate.

[5]These files, along with the other input data and results files, will be posted to `http://www.spc.int/oceanfish/en/ofpsection/sam/sam` at the conclusion of SC11 once a base case model for provision of management advice has been decided.

Due to the low number of observations for recent cohorts, recruitment estimates in the terminal model time periods may be poorly estimated. This was investigated using retrospective analysis where data from the terminal time periods (the last three years) were successively removed and the model fitted to each case (see Annex). The terminal recruitments and biomass estimates were compared among the retrospective models for their robustness to the loss of data.

4.6 Stock assessment interpretation methods

Several ancillary analyses using the converged model were conducted in order to interpret the results for stock assessment purposes. The methods involved are summarized below and the details can be found in Kleiber et al. (2014).

4.6.1 Reference points

The unfished spawning biomass ($SB_{F=0}$) in each time period was calculated given the estimated recruitments and the Beverton-Holt spawner-recruit relationship. This offers a basis for comparing the exploited population relative to the population subject to natural mortality only. WCPFC adopted $20\%SB_{F=0}$ as a limit reference point for the albacore stock. For this assessment $SB_{F=0}$ is calculated as the average over the period 2003–2012. The other key reference point, $F_{current}/F_{MSY}$, is described in Section 4.6.3.

4.6.2 Fishery impact

Many assessments estimate the ratio of recent to initial biomass as an index of fishery depletion. The problem with this approach is that recruitment may vary considerably throughout the time series, and if either the initial or recent spawning biomass estimates (or both) are "non-representative" because of recruitment variability or uncertainty, then the ratio may not measure fishery depletion, but simply reflect recruitment variability.

We approach this problem by computing the spawning biomass time series (at the region level) using the estimated model parameters, but assuming that fishing mortality was zero. Because both the real spawning biomass SB_t and the unexploited spawning biomass $SB_{0,t}$ incorporate recruitment variability, their ratio at each time step of the analysis $SB_t/SB_{0,t}$ can be interpreted as an index of fishery depletion. The computation of unexploited biomass includes an adjustment in recruitment to acknowledge the possibility of reduction of recruitment in exploited populations through stock-recruitment effects. This analysis was conducted in respect of groups of fisheries so as to describe the relative fishing impacts of each group on the population.

We note that this approach is similar to that undertaken for the estimation of the limit reference point (LRP), $20\%SB_{F=0}$, except that for the LRP the level of $SB_{F=0}$ is the average over a particular time window rather than the value predicted for a given year.

4.6.3 Yield analysis

The yield analysis consists of computing equilibrium catch (or yield) and biomass, conditional on a specified basal level of age-specific fishing mortality (F_a) for the entire model domain, a series of fishing mortality multipliers, *fmult*, the natural mortality-at-age (M_a), the mean weight-at-age (w_a) and the SRR parameters. All of these parameters, apart from *fmult*, which is arbitrarily specified over a range of 0-50 in increments of 0.1, are available from the parameter estimates of the model. The maximum yield with respect to fmult can easily be determined and is equivalent to the MSY.

As in previous tropical tuna assessments the SRR was not estimated over all the estimated recruitments – due to the poorer quality of data – and more uncertain recruitment estimates in the 1960s – these were excluded from the SRR fitting procedure.

For the standard yield analysis, the F_a are determined as the average over some recent period of time. In this assessment, we use the average over the period 2009–2012. We do not include 2013 in the average as fishing mortality tends to have high uncertainty for the terminal data year of the analysis and the catch and effort data for this terminal year are usually incomplete.

The MSY-based reference points were also computed using the average annual F_a from each year included in the model (1960–2013). This enabled temporal trends in the reference points to be assessed and a consideration of the differences in MSY levels under historical patterns of age-specific exploitation.

5 Model runs

5.1 Developments from the last assessment

Many changes have been implemented to the 2015 reference case model when compared to the 2012 reference case model. These changes came about through implementation of recommendations from the previous assessment (Hoyle et al., 2012), the independent bigeye tuna review (Ianelli et al., 2012), and efforts to make the south Pacific albacore tuna assessment more structurally similar to to those for bigeye tuna and yellowfin tuna to assist in evaluation of management options, including bio-economic modelling.

The major model changes are outlined in Table 1 and include changes to the spatial coverage and inclusion of sub-regional stratification in additions to considerable simplification to the fisheries structure. Subsequently it was not considered possible to provide ‘step by step’ changes, and the two forms of retrospective analyses described in the Annex provide the best basis for examining how the assessment has changed in response to modelling improvements and new data.

5.2 Sensitivity analyses and structural uncertainty

Several hundred runs were undertaken in conducting the 2015 south Pacific albacore tuna assessment, but in terms of presenting information on the bounds of plausible model sensitivity we have focused on a small of uncertainty axes which are described in further detail below. These axes were used for both 'one-off' changes from the reference case model and for the structural sensitivity analyses (after Hoyle et al. (2008a)) where all-possible combinations of the assumptions tested in the one-off sensitivity runs were considered. This resulted in a grid of 36 models (Table 4).

Natural mortality

The previous assessment assumed a value of 0.4 for the reference case and 0.3 and 0.5 for sensitivity. Considering the values used in other albacore tuna assessments we have used 0.3 for the reference case and 0.25 and 0.4 for sensitivity. The north Pacific albacore tuna assessment (ISC Albacore Working Group, 2011) examined values of 0.25 and 0.35 around their best estimate of 0.3, but we wanted to include 0.4 as it had been the reference case in the previous assessment.

Size data relative weighting

In integrated stock assessment models such as this, the choice of weight for the size data likelihood component (SZ_{dw}) is somewhat arbitrary. It is therefore standard procedure to test the assumption used for the reference case in a sensitivity analysis. The relative influence of the length composition data for all fisheries was reduced (i.e., a lower SZ_{dw}) by assigning an effective sample size of 0.02 (0.05 in the reference case) times the individual samples, with a maximum sample size of 20 (50 in the reference case). This explores the relative influence of size composition data upon the model estimates and illustrates data conflicts.

Regional weights

Regional weights are important for determining the average distribution of vulnerable biomass across regions. For the reference case model was had used all available operational CPUE data to construct the regional relative abundance surface. For sensitivity we considered regional weights based on data only for the period since 2008. This period was chosen as this represents the period where catches increased, however, catch data for this period is less supported by operational catch and effort data so it is considered less plausible.

Steepness

As was the case in the tropical tuna assessments we assumed a value of 0.8 for the reference case, but examined values of 0.65 ($h_{0.65}$) and 0.95 ($h_{0.95}$) in the grid.

6 Results

6.1 Model diagnostics (reference case)

A brief review of the fit of the model to four data sources follows: the standardized CPUE for the longline fisheries, size composition data, age-length data, and tagging data. The penalty for fitting the catch data is sufficiently high that the fit is essentially perfect so is not discussed further.

Longline CPUE

The fit to the standardized indices (Figure 6) is provided in Figures 10 and 11. While all fits look satisfactory based on Figure 6, some patterns in the effort deviates are evident in Figure 11. In regions three and four the model underestimates the overall CPUE decline, while in region five the model predicts lower CPUE than observed in recent years.

Size composition data

Two diagnostics are presented to illustrate the fit of the model to the observed size composition data: Figure 12 shows the aggregated (across all observations for a fishery) observed and predicted length frequencies for each fishery, and Figure 13 shows the predicted and observed median lengths and weight over time.

Given the flexible functional selectivity forms and lack of grouped selectivity curves across fisheries, the general fits presented in Figure 12 appear satisfactory. One concern, which is discussed further later in the report, was the inability of the model to adequately predict up the length modes in the troll fisheries.

Hoyle et al. (2012) used time-block and seasonal selectivity curves in the 2012 assessment whereas we have taken a more simplistic approach, and this, or possibly the approaches used to pool length frequency data within fisheries, may have contributed to some of the lack of fit evident in Figure 13. There are an apparent increases, albeit slight, in the sizes of fish through time taken in longline fisheries in regions 4, 5, and 7. Data for longline fisheries in regions 1, 3, 6, and 8 are more variable so no obvious lack of fit is apparent.

Tagging data

Fits to the tagging data compiled in various ways are shown in Figures 14, 15, and 16. Tagging was carried out only in regions 3, 6 and 8, therefore only these rows of Figure 14 contain observed and predicted data. Given the small number of tag recaptures in the model overall, most of the numbers in Figure 14 are very small, usually representing just one or two recaptures and so is not a particularly useful diagnostic.

Figures 15 and 16 show the observed and predicted recoveries in a spatially aggregated form, by calendar time period and by time at liberty, respectively. Generally speaking, these figures indicate that the model is capturing reasonably well the time series behaviour of tag returns.

Conditional age-length data

Model predictions of mean length-at-age from the estimated growth function are generally consistent with the observations from conditional age-length data for the ages best represented, i.e., 15 to 25 quarters (Figure 19). Model predictions are somewhat higher than that observed for age-classes less than 15 quarters, and lower for age-classes greater than 30 quarters. This suggests the estimated growth is faster at young ages, and slower at old ages, than is implied from the observed conditional age-length data. However, the estimated variance of the model growth function is relatively broad (Figures 18 and 19), which facilitates a reasonable fit to the observed range of length-at-age in these age-classes; e.g., to 57 cm in age-class 10 quarters, and to 115 cm in age-class 35 quarters. The apparent lack of fit of the predicted mean length-at-age over these age-classes most likely reflects the effects of other data types included in the overall model fit that impact upon the growth estimates, in particular size composition data.

6.2 Model parameter estimates (reference case)

Tag Reporting Rates

The estimates of tag reporting rates (Figure 17) are virtually unconstrained, and the estimates reflect to a large extent the relative numbers of tags returned by the various fisheries. Reporting rates are estimated to be higher for L-ALL-2, L-ALL-5, L-ALL-7 (which have recorded the highest catches) and T-ALL-3, from which several of the more recent tag releases were reported. Only one estimate, for L-ALL-7, is at the upper bound of 0.9.

Growth

Model predictions of mean length-at-age from the estimated growth function exhibit rapid growth at young ages up to 12 quarters, followed by an inflection around the age (17 - 18 quarters) that corresponds to the length-at-50% maturity (around 90 cm, Farley et al., 2013b), with extremely slow growth thereafter to a maximum average length of around 100 cm, (Figure 18). Individual growth variability is large, with fish within older age classes having a possible length range of around 40 cm. This is slightly larger than that predicted in previous assessments (Hoyle et al., 2012), and the growth rate of younger fish is more rapid. However, the average maximum length is comparable, and most probably reflects the influence that size composition and conditional age-length data included in the fit contain few observations larger than around 115 cm (Figures 12 and 19).

Selectivity

Age-specific selectivity coefficients are shown in Figure 20. For longline, there is a clear pattern of younger albacore being selected in the southern regions (3, 6 and 8). This suggests that the model is unable to explain the availability of albacore of different sizes in the various regions through fish movement alone. For the longline fisheries in regions 2 and 5, where the highest catches are taken, the population is not fully selected until around quarterly age-class 30, compared to age-classes 10–15 for the southern regions. For the troll and driftnet fisheries, selectivity peaks around age-classes 7–8, and is effectively zero for age-classes older than about 15 quarters.

Catchability

Time-series changes in catchability were only estimated for the surface fisheries and of these only the troll fisheries have a reasonably long period of activity in the fishery (Figure 21). The strong increasing catchability trends for the troll fisheries in regions 3 and 6 and the declining trend in region 8 may represent true changes in these fisheries, or reflect conflicts between these unstandardized data and the standardized CPUE time series used for the longline fisheries.

Movement

A graphical representation of the quarterly movement coefficients is shown in Figure 22, which displays the proportions of fish in each region moving to every other region per quarterly time period. The main features indicated are relatively small amounts of movement among the western regions (1, 2 and 3) and these regions appear to be to a large extent isolated from the regions to the east. There is more substantial exchange between regions 4, 5 and 7. Not surprisingly, most of the higher movements occur between adjacent regions; however some amount of movement is possible between non-adjacent regions as well.

The outcome of the estimated movement patters is shown in Figure 23, which portrays the origin of the equilibrium albacore biomass in each region. This plot reinforces the lack of mixing estimated to occur within the western regions (1, 2 and 3). Regions 4, 5 and 7 appear to be comprised of albacore originating mainly in region 5, while regions 6 and 8 are comprised almost entirely of fish originating in regions 3 and 6.

It should be stressed that the information available to parameterise movement in this model is sparse, and it is likely that similarities and differences in relative longline CPUE trends and regional scaling to a large extent influence the movement pattern that is estimated.

6.3 Stock assessment results

Symbols used in the following discussion are defined in Table 2 and the key results are provided in Table 5.

6.3.1 Recruitment

The estimated distribution of recruitment across regions should be interpreted with caution as MULTIFAN-CL can use a combination of movement and regional recruitment to distribute the population in a way that optimizes the objective function. The reference case recruitment estimates (aggregated by year for ease of display) for each region and the entire assessment domain are shown in Figures 24 and 25. As separate y-axes are used in Figure 24, the second figure is important for context. The overall pattern can be described as having high and variable recruitment in the first five year of the model, followed by lower, but variable, recruitment since. The low recruitment at the end of the time series should be interpreted with caution due to the lack of data supporting its estimation and the setting of the final four deviates to the mean.

The model estimates that recruitment is almost entirely sourced from five regions: 1, 2, 3, 5 and 6.

6.3.2 Biomasss

Trends in biomass are represented using the estimated spawning potential, although some key total biomass reference points are included in the results tables.

The reference case spawning potential estimates (aggregated by year for ease of display) for each region and the entire assessment domain are shown in Figures 26 and 27. As separate y-axes are used in Figure 24, the second figure is important for context. The overall pattern can be described as starting from from a high level in 1960 and, aside from peaks in 1970 and around 2000, following a near constant decline over the time period. Aside from region 2, biomass declines overtime in all regions. The increase in region 2 is driven by the CPUE and catch for that region. The most important regions for biomass are 2 and 5, followed by 8, 7, and 6. The tropical regions (north of 10° S are estimated to have low albacore tuna biomass.

6.3.3 Fishing mortality

Average fishing mortality rates for juvenile age-classes has increased gradually throughout the time series with one large spike coinciding with the brief period of driftnet fishing (Figure 28). Average fishing mortality rates for adult age-classes increased gradually through to 2000 where it more than doubled at the time of rapidly increasing longline catch. It increased dramatically again with the most recent increase in longline fishing since 2010 throughout the time series while juvenile mortality increases strongly through to the late 1990s and has been relatively stable since.

Changes in fishing mortality-at-age and population age structure are shown for decadal time intervals in Figure 29. The irregular age-structure patterns reflect the seasonal occurrence of recruitment within a quarterly model. The relative impacts of the surface and longline fisheries is reflected in the small increase in F before age 10 quarters versus the high level of fishing mortality plateauing around age 25 quarters.

6.3.4 Fishery impact

We measure fishery impact at each time step as the ratio of the estimated spawning potential to that which would have occurred in the historical absence of fishing. This is a useful variable to monitor, as it can be computed both at the sub-regional level and overall. This information is plotted in two ways, first the fished and unfished spawning potential trajectories (Figure 30) and second as the depletion ratios themselves (Figures 31 and 37). The latter is relevant for the agreed limit reference point and discussed in more detail in Section 6.4.1.

It is also possible to ascribe the fishery impact to specific fishery components in order to see which types of fishing activity have the largest impact on the spawning potential (Figures 32 and 33). Unlike with bigeye and yellowfin tunas, there are fewer gears impacting on south Pacific albacore tuna so for the impact plot we examine the impacts of the longline fisheries in different latitude bands and for the surface fisheries combined.

The fishing impact / depletion levels were generally consistent across regions with between 40–60% fishery impact and the subtropical fisheries in regions 2, 5, and 7 were responsible for most of the fishery impact, followed by the longline fisheries in the south (regions 3, 6, and 8). While important in some regions and in the past, current impacts of the surface fisheries on the overall stock is low.

6.3.5 Yield analysis

The yield analyses conducted in this assessment incorporate the SRR (Figure 34) into the equilibrium biomass and yield computations. Importantly in the reference case model the steepness of the SRR was fixed at 0.8, so only the scaling parameter was estimated.

MSY was estimated at 76,800 mt and the $YF_{current}$ was 63,840 mt (Figure 35 and Table 5). Catches in 2013 were very slightly higher than MSY. The equilibrium unfished spawning potential was estimated at 396,500 mt and the spawning potential that would support the MSY was estimated to be 57,430 mt which is only 14% of SB_0. This very low value is a product of the relationship between the fishing-mortality-at-age and maturity-at-age profiles for south Pacific albacore tuna.

6.4 Stock status

6.4.1 Majuro plot and the Limit Reference Point (LRP)

$SB_{F=0}$ calculated for the period 2003–12 is the basis for the limit reference point and this is a spawning potential of 408,361 mt which is only 3% higher than SB_0 (Table 5). The limit reference point is 20%$SB_{F=0}$ and this is a spawning potential of 81,672 mt.Latest (2013) spawning potential is estimated to be 40% of $SB_{F=0}$ (Figures 36 and 37).

Fishing mortality has generally been increasing through time, and for the reference case $F_{current}/F_{MSY}$ (2009–12 average) is 0.39. This indicates that a 2.5 times increase in fishing mortality is necessary to produce the MSY (Table 5); this increase in effort would increase equilibrium catch by 20%, but likely reduce catch rates by almost 65% (comparing $SB_{latest}/SB_{F=0}$ to SB_{MSY} as a proxy of longline vulnerable biomass).

6.4.2 Against potential Target Reference Points (TRPs)

There are currently no agreed biomass-related target reference points for any species, but the WCPFC has requested investigation of spawning potential in the range of 40-60%$SB_{F=0}$ for skipjack for potential biomass-related target reference points and has examined economic-based target reference points for the south Pacific albacore tuna stock. Based on bio-economic modelling described in Pilling et al. (2015) the range of $SB_{F=0}$ that would support break-even or 10% profits is 0.65–0.80$SB_{F=0}$. This region has been shaded green on the Majuro plot (Figures 36). As reported above, current (2009–12 average) and latest (2013) spawning biomass are estimated to be 41% and 40% respectively of $SB_{F=0}$ and therefore are lower than these potential TRPs.

We note that it will be important to update the analyses of Pilling et al. (2015) with the new assessment to ensure that the benchmarks are comparable and incorporate any new economic information.

6.5 Sensitivity of the reference case

In this section we examine the sensitivity of the reference case model to the various data and modelling assumptions.

6.5.1 One-off changes from the structural uncertainty analysis

Comparisons of the recruitment, spawning potential, and depletion trajectories for the reference case and one-change sensitivity runs from the structural uncertainty analysis are provided in Figures 38–40, the key reference points are compared in Table 5 and the likelihood components in Table 7. Majuro plots for the one-off models are provided in Figure 41.

Natural mortality (M)

The model was very sensitive to the values of M considered. Low M gave very low MSY (62,440 mt) and closer examination of this model indicated that current catches (higher than the estimated MSY) were only being supported by above average recent recruitment, (aka bigeye tuna). Under this scenario $SB_{latest}/SB_{F=0}$ was down to 0.31. The model also had the highest $F_{current}/F_{MSY}$ at 0.59. The higher value of M gave the highest MSY of the one-off sensitivity analyses (112,400 mt), the highest level of $SB_{latest}/SB_{F=0}$ at 0.55, and the lowest $F_{current}/F_{MSY}$ at 0.20.

The lower value of M gave a much worse fit to the data – especially the CPUE, tagging, and age-length data, conversely the higher M run gave a better fit to these data sets.

Weight to the size data (SZ_dw)

Reducing the weight to the length-frequency data provided better fits to the age-length, CPUE, and tagging data. For CPUE, it provided a better fit to fisheries 3, 4, 6, and 7, and worse fits to fisheries 1 and 8. This model gave the highest absolute biomass and an MSY of 91,120 mt and more optimistic level of $SB_{latest}/SB_{F=0}$ of 0.47.

Regional weights

Using regional weights derived from the most recent few years lead to higher absolute levels of biomass and MSY (91,400 mt) and more optimistic level of $SB_{latest}/SB_{F=0}$ of 0.48. This model had a slightly worse overall fit, fitting better to the CPUE and tagging data, but worse to the length-frequency data.

Steepness (h)

Steepness, has similar, but less extreme affects on $F_{current}/F_{MSY}$ and $SB_{latest}/SB_{F=0}$ than the natural mortality sensitivity analyses, but the MSY value itself was relatively insensitive. By design the assumed value of steepness has little effect on model fitting and other model quantities.

6.5.2 Structural uncertainty analysis

Comparisons of the impacts of different axes of the structural uncertainty analysis are shown in two ways, first through a series of Majuro plots which show $F_{current}/F_{MSY}$ and SB_{latest}/SB_{MSY} with colour coding for each option within the axes (Figure 43), and second through a series of box and whisker plots (Figures 44–46).

The general patterns for each option within the five axes are the same as described in Section 6.5 so we do not repeat them again here. The positive (or negative) impacts of the different options were found to be somewhat additive, e.g., model runs with more options that individually produce generally "better" outcomes gave even "better" outcomes when combined.

Considering the multiplier on current effort required to achieve , *fmult*, the model with the lowest value (1.36) included low M and low steepness and reference case regional weights and length data weighting; conversely the model with the highest value (11.14) included high M and steepness, recent regional weights and length data down weighted. These same models gave $SB_{latest}/SB_{F=0}$ of 0.28 and 0.67 respectively.

6.6 Overall stock status conclusions

Based on the results from the reference case model, and the sensitivity analyses, including the structural uncertainty grid, we make the following conclusions regarding stock status:

- Current catch is either at or less than MSY;
- Recent levels of spawning potential are most likely above the level which will support the MSY and $20\% SB_{F=0}$;
- Recent levels of fishing mortality are lower than the level that will support the MSY;
- Increasing fishing mortality to F_{MSY} levels would require a significant increase in effort, yield only very small (if any) increases in long-term catch, and would greatly reduce the vulnerable biomass available to the longline fleet; and
- Recent levels of spawning potential are lower than candidate bio-economic-related target reference points currently under consideration for south Pacific albacore tuna.

7 Discussion and conclusions

In this section we will discuss the major changes from the 2012 assessment of Hoyle et al. (2012) in terms of the modelling approaches and conclusions (Section 7.1), focus on some of the specific areas of uncertainty which we believe are relevant to the interpretation of the assessment (Sections 7.2), and finally outline what we believe are the key research priorities to support further improved assessments for south Pacific albacore tuna in the future (Section 7.3).

7.1 Changes to the previous assessment

The gap between the 2015 and 2012 assessments provided considerable opportunity for changes in the fishery and stock assessment approaches – both within MULTIFAN-CL and generally. This assessment has befitted from the implementation of many of the recommendations from the Independent Review of the 2011 bigeye assessment (Ianelli et al., 2012; Davies et al., 2015).

The major changes to the current assessment are outlined in Table 1 and included having explicit spatial structure in the model and exclusion of waters outside the WCPFC Convention area, new data sets (age-length and recent albacore tuna tagging data), and some changes to assumed biological parameters.

The major changes to the spatial structure mean that it is not possible to provide 'step by step' analysis of the impacts of the various developments, but it is safe to conclude that the estimated MSY will be lower due to the exclusion of the EPO regions. At the same time this also means that the MSY estimate is more directly applicable to decisions of WCPFC.

The more complex structure – both spatially and the inclusion of a quarterly time step – added to the computational demands of this assessment and therefore the decision was made to simplify the fisheries structures used. Another implication of moving to a spatially-structured model was the need to consider the distribution of recruitment, regional weights, and movement. Considerable time was spent, without success, attempting to replicate the current hypotheses regarding recruitment and movement distributions. This was similar to the outcome of the SEAPODYM modelling (Lehodey et al., 2012), and eventually we decided to use the recruitment distribution estimated by SEAPODYM as the starting point for the model runs.

One change for which we were able to quantify the effect, was the change to a lower estimate of natural mortality. The one-off sensitivity analyses showed that this reduced the MSY from 112,400 mt to 76,800 mt. Almost all of the data sets included in the assessment 'preferred' a higher value for M, but we did not have time to carefully consider whether this improved fit was for aspects of these data that we were confident in or not. The key reason for making the change was to make our assumptions consistent with those used in other albacore tuna assessments (e.g.,the North Pacific and the Atlantic Oceans). We do not believe that there is any specific reason for M to be higher in the south Pacific, but this does not mean that the value of 0.3 or 0.4 is correct. There is a proposal for a joint Tuna Regional Fisheries Management Organization (tRFMO) collaboration on Management Strategy Evaluation (MSE) with a focus on albacore tuna (Dr Laurie Kell, *pers. comm.*). This forum will provide the opportunity for consideration of natural mortality and other important model parameters.

Overall the changes in the current assessment have not unexpectedly changed our view of the status of this resource. From the retrospective analyses (provided in the Annex), our results are generally similar to those from recent assessments, with the more pessimistic stock status well explained by the increase in catches since the previous assessment.

7.2 Sources of uncertainty

In this section we comment on some of the difficulties encountered in the assessment or issues that arose in the modelling which led to potential uncertainty. This includes discussion of some of the factors that were included in the uncertainty framework used in the assessment, i.e., sensitivity analyses and the structural uncertainty analysis (grid).

Longline catch and effort data are critical to MULTIFAN-CL assessments for deriving both CPUE indices and the regional weights. One potential uncertainty in these model inputs was the decision by Japan not to allow for the use of operational data provided for the Pacific-wide bigeye tuna analysis[6]. Japan was the primary fleet fishing in the south Pacific Ocean in the early part of the period covered by the south Pacific albacore tuna assessment. We are not able to determine the impact of the exclusion of these data on the stock assessment results and conclusions, but hope that this matter can be addressed prior to the next south Pacific albacore tuna assessment.

The assessment results were moderately sensitive to the assumed regional weights used to scale the CPUE in each region. The reference case assumption based on all available data is clearly the one best supported by data (i.e., there were far more spatial gaps in coverage for the analysis using a later time period), but the sensitivity of the results to the alternative series highlights regional weights as an important source of uncertainty. Once all collected data are available for analysis, we suggest revisiting the analysis of regional weights – potentially including other covariates (both relating to fishing practices and oceanographic conditions).

The assessment results were very sensitive to the weighting applied to the length frequency data which is a different outcome compared to the bigeye tuna assessment of Harley et al. (2014) where results were generally insensitive to the weighting. It is clear that there is conflict between signals from the size data and those from the other data sets, and this should be further examined in the next assessment. This could include consideration of relative data weighting (Francis, 2011) or likelihood functions assumed (Francis, 2014; Davies et al., 2015), as well as structuring some of the longline fisheries, e.g., separating out some fleets and/or considering time-blocks in selectivity.

As was done for the tropical tuna assessments in 2014 we excluded data for the most recent year as provisional estimates for the catch and effort data for longline fisheries (available in time for

[6] All other parties to the Pacific-wide bigeye tuna Memorandum of Understanding (MOU) were able to extend the agreement to cover the south Pacific albacore tuna assessment, but Japan indicated that it was not able to agree. Japan indicated that we must follow proper procedures (email from Dr Keisuke Satoh dated 22 May 2015), but did not respond to questions of clarification on what these procedures were.

the assessments) have generally been subject to significant revision either during or shortly after the completion of the assessment. This does make the assessment less 'current' than it would be if accurate data for 2014 were available. Nominal CPUE data for one long-term fleet indicates that 2014 was another very poor year for CPUE (Brouwer et al., 2015) so this will likely be estimated as a further decline in the longline vulnerable biomass in the next south Pacific albacore tuna assessment.

Uncertainty in growth was a dominating feature of the 2012 assessment (and ones before it) and had been the motivating factor for the various biological studies undertaken in recent years, e.g., Farley et al., 2013b; Williams et al., 2012. While the age-length data did reduce the uncertainty in the lengths of older fish, in the development of the assessment we found strong differences in growth of younger fish, specifically the interpretation of the modes within the troll length frequency data – our understanding that they were annual – and the conclusion from the age-length data that they might be six monthly, but definitely not annual. In essence, the modes in the troll length frequency data, if they are annual, would suggest that albacore of 50-70 cm taken in this fishery are growing at a rate of approximately 10 cm per year (Figure 12). By contrast, the age at length data suggest the growth rate to be approximately 20 cm per year (Figure 19). We subsequently excluded the age-length records for the troll fishery after discussions with those involved in the ageing work who indicated that the partial ages for the youngest fish may require further examination. However, this did not remove the inconsistency and further investigation of this issue in required.

7.3 Recommendations for further work

As discussed in the sections above, there are areas of uncertainty in the current assessment, and many of these can be addressed by further work. This section outlines some recommendations, some directed at those undertaking future assessments, and some at the SC and WCPFC itself.

WCPFC-specific recommendations

- WCPFC must consider establishing a process (either compulsory or voluntary) to ensure that all available operational longline logsheet data are available to support the best possible regional stock assessments, including the testing and development of methods for CPUE standardization and related analyses (e.g., regional weights).

Biological studies

- Conflict between the growth rates implied by direct ages from Farley et al. (2013a) and the troll length frequency data led to the exclusion of age-length records from small fish in the current assessment. We recommend re-examination of these otoliths from young fish to help resolve the issue and allow better modelling of growth rates for small albacore tuna;

- Collaborate with albacore tuna assessment scientists in other RFMOs to get a better understanding for the basis of assumed values for M considered in the assessments, and the available data sets that might better inform it;
- Consider routine collection of hard parts from the longline fishery to allow the construction of catch-at-age data sets which should improve the ability of assessments to track trends in fishing mortality. Given the nature of the growth curve, and selectivity of the longline fisheries, current length-frequency data are likely to provide far less information that age-frequency data; and
- Consideration of biological markers/indicators that might provide better information of rates of east-west movement of albacore tuna will be important for investigating the issue of spatial variation in growth described in Williams et al. (2012). Current tagging data are unlikely to be sufficient.

MULTIFAN-CL modelling and the stock assessment

- Examine the potential for orthogonal recruitment structure (i.e., estimating a constant seasonal plus a year effect) to provide a more parsimonious approach for modelling the strong seasonal recruitment signal present in the south Pacific albacore tuna assessment. Implementation of this approach should consider the initial conditions, recruitment deviates not estimated at the end of the time series, and recruitment in projections;
- Continue the analysis of operational data – expanding the data set to include all fleets if possible – to further examine clustering approaches and geospatial models for both CPUE standardization and estimation of regional weights. This should also include examination of potential bias in CPUE indices caused by vessels changing their fishing patterns based on expected CPUE and economic returns (after Brouwer et al., 2015);
- Further examination of the weighting of the different data sources (after Francis, 2011), including the new age-length data. This should include consideration of the best weights and/or likelihood functions to use for size frequency data (e.g., Francis, 2014).;
- Further examine different scenarios of recruitment distributions and movement to assess the potential for MULTIFAN-CL to approximate the current thinking on population sub-structure and migration behaviour – both in terms of available data and model structures;
- Consideration of finer scale modelling of the longline fisheries including time-block splits and separating longline fleets within a region where appropriate;

- Implement an approach that allows for better internal consistency between maturity-at-age and maturity-at-length when growth is estimated within the model;
- Examination of the pros and cons of moving to a sex-structured stock assessment model for south Pacific albacore tuna; and
- Update the bio-economic analysis of potential Target Reference Points (TRPs) with the current assessment and economic data.

7.4 Main assessment conclusions

The main conclusions of the current assessment are consistent with the previous assessment conducted in 2012. The main conclusions based on results from the reference case model and with consideration of results from performed sensitivity model runs, are as follows:

1. The new regional structure used for the 2015 assessment is better aligned with those of the assessments for bigeye and yellowfin tunas and provides an improved basis for further development of this assessment and providing advice to WCPFC;
2. There is some conflict between some of the data sources available for this assessment including conflicts between the length-frequency data and the CPUE series and between the troll length frequency samples and the age-length data;
3. Current catch is either at or less than MSY;
4. Recent levels of spawning potential are most likely above the level which will support the MSY, and above $20\% SB_{F=0}$;
5. Recent levels of fishing mortality are lower than the level that will support the MSY;
6. Increasing fishing mortality to F_{MSY} levels would require a significant increase in effort, yield only very small (if any) increases in long-term catch, and would greatly reduce the vulnerable biomass available to the longline fleet;
7. Recent levels of spawning potential are lower than candidate bio-economic-related target reference points currently under consideration for south Pacific albacore tuna; and
8. Stock status conclusions were most sensitive to alternative assumptions regarding the weighting off different data sets and natural mortality, identifying these as important areas for continued research.

Acknowledgements

We are grateful to all fisheries agencies in the region that have contributed data to the SPC and/or the WCPFC for this assessment. We thank Peter Williams and his team for assistance with collating the input data. Graham Pilling and Steve Brouwer for their review of a haggard earlier draft and Keith Bigelow for insights into the CPUE analysis and regional structure.

References

Berger, A. M., McKechnie, S., Abascal, F., Kumasi, B., Usu, T., and Nichol, S. J. (2014). Analysis of tagging data for the 2014 tropical tuna assessments: data quality rules, tagger effects, and reporting rates. WCPFC-SC10-2014/SA-IP-06, Majuro, Republic of the Marshall Islands, 6–14 August 2014.

Brouwer, S., Harley, S. J., and Pilling, G. (2015). The influence of catch rate on fishing effort – an investigation using south Pacific albacore as an example. WCPFC-SC11-2015/SA-IP-04, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Cadigan, N. G. and Farrell, P. J. (2005). Local influence diagnostics for the retrospective problem in sequential population analysis. *ICES Journal of Marine Science: Journal du Conseil*, 62(2):256–265.

Cadrin, S. and Vaughan, D. (1997). Retrospective analysis of virtual population estimates for Atlantic menhaden stock assessment. *Fishery Bulletin*, 95(3):256–265.

Davies, N., Fournier, D. A., Hampton, J., and Bouye, F. (2015). Recent developments and future plans for MULTIFAN-CL. WCPFC-SC11-2015/SA-IP-01, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Farley, J. H., Hoyle, S. D., Eveson, J. P., Williams, A. J., Davies, C. R., and Nicol, S. J. (2014). Maturity ogives for south pacific albacore tuna (*Thunnus alalunga*) that account for spatial and seasonal variation in the distributions of mature and immature fish. *PLoS ONE*, 9(1):e83017.

Farley, J. H., Williams, A. J., Clear, N. P., Davies, C. R., and Nicol, S. J. (2013a). Age estimation and validation for South Pacific albacore *Thunnus alalunga*. *Journal of Fish Biology*, 82(5):1523–1544.

Farley, J. H., Williams, A. J., Hoyle, S. D., Davies, C. R., and Nicol, S. J. (2013b). Reproductive dynamics and potential annual fecundity of South Pacific albacore tuna (*Thunnus alalunga*). *PLoS ONE*, 8(4):e60577. doi:10.1371/journal.pone.0060577.

Fournier, D., Hampton, J., and Sibert, J. (1998). MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences*, 55:2105–2116.

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielson, A., and Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, 27(2):233–249.

Francis, R. I. C. C. (1999). The impact of correlations in standardised CPUE indices. NZ Fisheries Assessment Research Document 99/42, National Institute of Water and Atmospheric Research. (Unpublished report held in NIWA library, Wellington.).

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. (2014). Replacing the multinomial in stock assessment models: A first step. *Fisheries Research*, 151:70–84.

Hampton, J. and Fournier, D. (2001). A spatially-disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Marine and Freshwater Research*, 52:937–963.

Harley, S. J., Davies, N., Hampton, J., and McKechnie, S. (2014). Stock assessment of bigeye tuna in the Western and Central Pacific Ocean. WCPFC-SC10-2014/SA-WP-01, Majuro, Republic of the Marshall Islands, 6–14 August 2014.

Hoyle, S., Hampton, J., and Davies, N. (2012). Stock assessment of albacore tuna in the South Pacific Ocean. WCPFC-SC8-2012/SA-WP-04, Busan, Republic of Korea, 7–15 August 2012.

Hoyle, S., Langley, A. D., and Hampton, J. (2008a). General structural sensitivity analysis for the bigeye tuna stock assessment. WCPFC-SC4-2008/SA-WP-03, Port Moresby, Papua New Guinea, 11–22 August 2008.

Hoyle, S. D. (2011). Stock assessment of albacore tuna in the south Pacific Ocean. WCPFC-SC7-2011/SA-WP-06, Pohnpei, Federated States of Micronesia, 9–17 August 2011.

Hoyle, S. D. and Davies, N. (2009). Stock assessment of albacore tuna in the south Pacific Ocean. WCPFC-SC5-2009/SA-WP-06, Port Vila, Vanuatu, 10–21 August 2009.

Hoyle, S. D., Langley, A. D., and Hampton, J. (2008b). Stock assessment of albacore tuna in the south Pacific Ocean. WCPFC-SC4-2008/SA-WP-08, Port Moresby, Papua New Guinea, 11–22 August 2008.

Ianelli, J., Maunder, M. N., and Punt, A. E. (2012). Independent review of the 2011 WCPO bigeye tuna assessment. WCPFC-SC8-2012/SA-WP-01, Busan, Republic of Korea, 7–15 August 2012.

ISC Albacore Working Group (2011). Stock Assessment of Albacore Tuna in the North Pacific Ocean in 2011. WCPFC-SC7-2011/SA-WP-10, Pohnpei, Federated States of Micronesia, 9–17 August 2011.

Kleiber, P., Hampton, J., Davies, N., Hoyle, S. D., and Fournier, D. (2014). *MULTIFAN-CL User's Guide*.

Langley, A. (2004). An examination of the influence of recent oceanographic conditions on the catch rate of albacore in the main domestic longline fisheries. Technical Report Working Paper SA-4., 17th Standing Committee on Tuna and Billfish. 9–18 August 2004. Majuro, Republic of Marshall Islands.

Lee, H. H., Piner, K. R., Methot, R. D., and Maunder, M. N. (2014). Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: An example using blue marlin in the Pacific Ocean. *Fisheries Research*, 158:138–146.

Lehodey, P., Senina, I., Titaud, O., Calmettes, B., Nicol, S., Hampton, J., Caillot, S., and Williams, P. (2012). Project 62: SEAPODYM applications in the WCPO. . WCPFC-SC9-2012/EB-WP-03, Busan, Republic of Korea, 7–15 August 2012.

Leroy, B. and Lehody, P. (2004). Note on the growth of the south Pacific albacore. INFO-BIO-2, 17th Standing Committee on Tuna and Billfish, Majuro, Republic of Marshall Islands, 9–18 August 2004.

Mace, P. M. and Doonan, I. J. (1988). A generalised bioeconomic simulation model for fish population dynamics. New Zealand Fisheries Assessment Research Document 88/4.

McKechnie, S. (2014). Analysis of longline size frequency data for bigeye and yellowfin tunas in the WCPO. WCPFC-SC10-2014/SA-IP-04, Majuro, Republic of the Marshall Islands, 6–14 August 2014.

McKechnie, S., Hampton, J., Abascal, F., Davies, N., and Harley, S. J. (2015). Sensitivity of WCPO stock assessment results to the inclusion of EPO dynamics within a Pacific-wide analysis. WCPFC-SC11-2015/SA-WP-04, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Murray, T. (1994). A review of the biology and fisheries for albacore, *Thunnus alalunga*, in the South Pacific Ocean. In Shomura, R. S., Majkowski, J., and Langi, S., editors, *Interactions of Pacific tuna fisheries*, FAO Fish. Tech. Pap, pages 188–206.

Pilling, G., Reid, C., Harley, S. J., and Hampton, J. (2015). Compatibility and consequences of alternative potential Target Reference Points for the south Pacific albacore stock. WCPFC-SC11-2015/MI-IP-07, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Ramon, D. and Bailey, K. (1996). Spawning seasonality of albacore, *Thunnus alalunga*, in the Sensitivityouth Pacific Ocean. *Fishery Bulletin*, 94:725–733.

Scott, R. D. and McKechnie, S. (2015). Analysis of longline length frequency compositions for South Pacific albacore tuna. WCPFC-SC11-2015/SA-IP-07, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Tremblay-Boyer, L., McKechnie, S., and Harley, S. J. (2015a). Spatial and fisheries structure and regional weights for the 2015 south Pacific albacore tuna (*Thunnus alalunga*) assessment. WCPFC-SC11-2015/SA-IP-07, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Tremblay-Boyer, L., McKechnie, S., and Harley, S. J. (2015b). Standardized CPUE for south Pacific albacore tuna (*Thunnus alalunga*) from operational longline data. WCPFC-SC11-2015/SA-IP-03, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Williams, A. J., Farley, J. H., Hoyle, S. D., Davies, C. R., and Nicol, S. J. (2012). Spatial and sex-specific variation in growth of albacore tuna (*Thunnus alalunga*) across the South Pacific Ocean. *PLoS ONE*, 7(6):e39318. doi:10.1371/journal.pone.0039318.

Williams, A. J., Nicol, S. J., and Leroy, B. (2010). South Pacific albacore tagging project: 2010 summary report. WCPFC-SC6-2010/GN-IP-06, Nuku'alofa, Tonga, 10–19 August 2010.

Table 1: Major changes from the 2012 reference case model

Component	2012 assessment	2015 assessment
Spatial extent	Entire Pacific Ocean south of the equator	The WCPFC Convention are south of the equator
Regional structure	One region with six regionally defined fisheries	Eight regions with their own fisheries.
Fishery structure	Some fleet specific fisheries	Gear-specific fisheries definitions
Time step	Annual	Quarterly
Growth information	Length-frequency data	Age-length observations and length frequency data
Tagging data	Releases and recaptures from late 1980s and 1990s programmes	As before, but with relrease and recapture data from the most recent tagging (Williams et al., 2010).
Natural mortality	0.4 per annum	0.3 per annum

Table 2: Description of symbols used in the yield and stock status analyses. For the purpose of this assessment, 'current' is the average over the period 2009–2012 and 'latest' is 2013.

Symbol	Description
C_{latest}	Catch in the latest year
$F_{current}$	Average fishing mortality-at-age for a recent period
F_{MSY}	Fishing mortality-at-age producing the maximum sustainable yield (MSY)
MSY	Equilibrium yield at F_{MSY}
$F_{current}/F_{MSY}$	Average fishing mortality-at-age for a recent period relative to F_{MSY}
SB_0	Equilibrium unexploited spawning potential
SB_{latest}	Spawning potential in the latest time period
$SB_{F=0}$	Average spawning potential predicted to occur in the absence of fishing for the period 2002–11
SB_{MSY}	Spawning potential that which will produce the maximum sustainable yield (MSY)
$SB_{latest}/SB_{F=0}$	Spawning potential in the latest time period relative to the average spawning potential predicted to occur in the absence of fishing for a period.
SB_{latest}/SB_{MSY}	Spawning potential in the latest time period relative to that which will produce the maximum sustainable yield (MSY)
$20\%SB_{F=0}$	WCPFC adopted limit reference point – spawning potential in the absence of fishing average over years $t-11$ to $t-1$.

Table 3: Summary of the groupings of fisheries within the assessment for selectivity curve, catchability (used for the implementation of regional weights), tag recaptures, and tag reporting rates.

	Region	Selectivity	Catchability	Tag.recaptures	Tag.reporting
L-All-1	1	1	1	1	1
L-All-2	2	2	1	2	2
L-All-3	3	3	1	3	3
L-All-4	4	4	1	4	4
L-All-5	5	5	1	5	5
L-All-6	6	6	1	6	6
L-All-7	7	7	1	7	7
L-All-8	8	8	1	8	8
T-All-3	3	9	2	9	9
T-All-6	6	10	3	10	10
T-All-8	8	11	4	11	11
D-All-3	3	12	5	12	12
D-All-6	6	13	6	13	13
D-All-8	8	14	7	14	14

Table 4: Description of the structural sensitivity grid used to characterise uncertainty in the assessment. The reference case option is denoted in **bold face**.

Axis	Levels	Options
Natural mortality	3	0.25, **0.3**, or 0.4 per year
Length data weighting	2	sample sizes divided by **20** or 50
Regional weights	2	data from **1975 onwards** or 2008
Steepness	3	0.65, **0.80**, or 0.95

Table 5: Key reference points for the reference case model and one-off sensitivity analyses

	Ref.Case	+2008 Reg.wei	Sz.dwnwt	Low M	High M	h 0.65	h 0.95
C_{latest}	$77,046$	$77,425$	$76,103$	$76,622$	$77,436$	$77,054$	$77,044$
$YF_{current}$	$63,840$	$66,280$	$66,680$	$58,680$	$69,280$	$69,600$	$60,640$
MSY	$76,800$	$91,400$	$91,120$	$62,440$	$112,400$	$79,000$	$77,640$
$fmult$	2.550	3.370	3.280	1.710	5.120	2.010	3.450
$F_{current}/F_{MSY}$	0.390	0.300	0.300	0.590	0.200	0.500	0.290
SB_0	$396,500$	$471,200$	$467,800$	$396,600$	$395,900$	$445,500$	$370,200$
$SB_{F=0}$	$408,361$	$464,971$	$465,028$	$426,025$	$397,801$	$436,769$	$393,139$
SB_{latest}/SB_0	0.410	0.470	0.470	0.330	0.550	0.370	0.450
$SB_{latest}/SB_{F=0}$	0.400	0.480	0.470	0.310	0.550	0.380	0.420

Table 6: Percentiles of key reference points from the grid

	5%	50%	95%
C_{latest}	$75,341$	$77,231$	$78,243$
$YF_{current}$	$58,500$	$67,900$	$76,190$
MSY	$65,950$	$91,660$	$149,900$
$fmult$	1.700	3.650	8.970
$F_{current}/F_{MSY}$	0.110	0.280	0.590
SB_0	$373,725$	$469,500$	$649,000$
$SB_{F=0}$	$396,636$	$464,999$	$608,470$
SB_{latest}/SB_0	0.330	0.480	0.630
$SB_{latest}/SB_{F=0}$	0.320	0.480	0.640

Table 7: Likelihood components for the reference case model and one-off sensitivity analyses

	Ref.Case	+2008 Reg.wei	Sz.dwnwt	Low M	High M	h 0.65	h 0.95
bhsteep	0.788	0.651	0.624	1.090	0.628	0.766	0.770
effdev	959.313	947.237	890.171	962.187	894.663	959.619	959.278
catdev	11.707	11.404	12.743	11.382	11.897	11.737	11.729
lencomp	-280, 938.500	-280, 900.500	-221, 876.800	-280, 915.200	-281, 054.900	-280, 938.500	-280, 937.900
wtcomp	0	0	0	0	0	0	0
tagdata	633.193	631.045	625.085	651.840	615.154	633.218	633.259
agelengthdata	3, 168.354	3, 170.084	3, 136.996	3, 175.970	3, 123.680	3, 168.482	3, 168.885
total_likelihood	276, 006.600	275, 984.200	217, 111.100	275, 949.000	276, 244.900	276, 006.300	276, 006.100

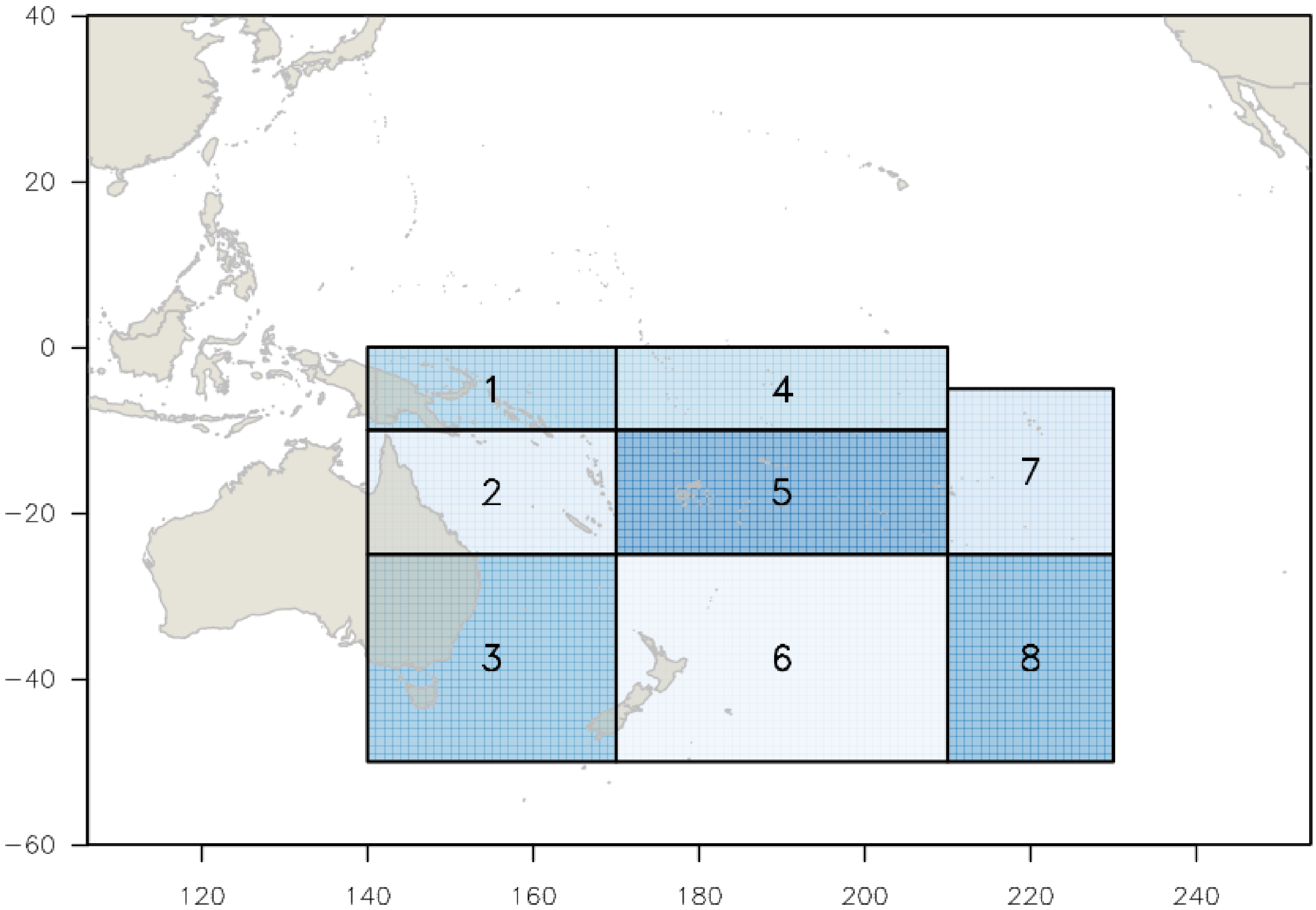


Figure 1: Regional structure of the reference case model.

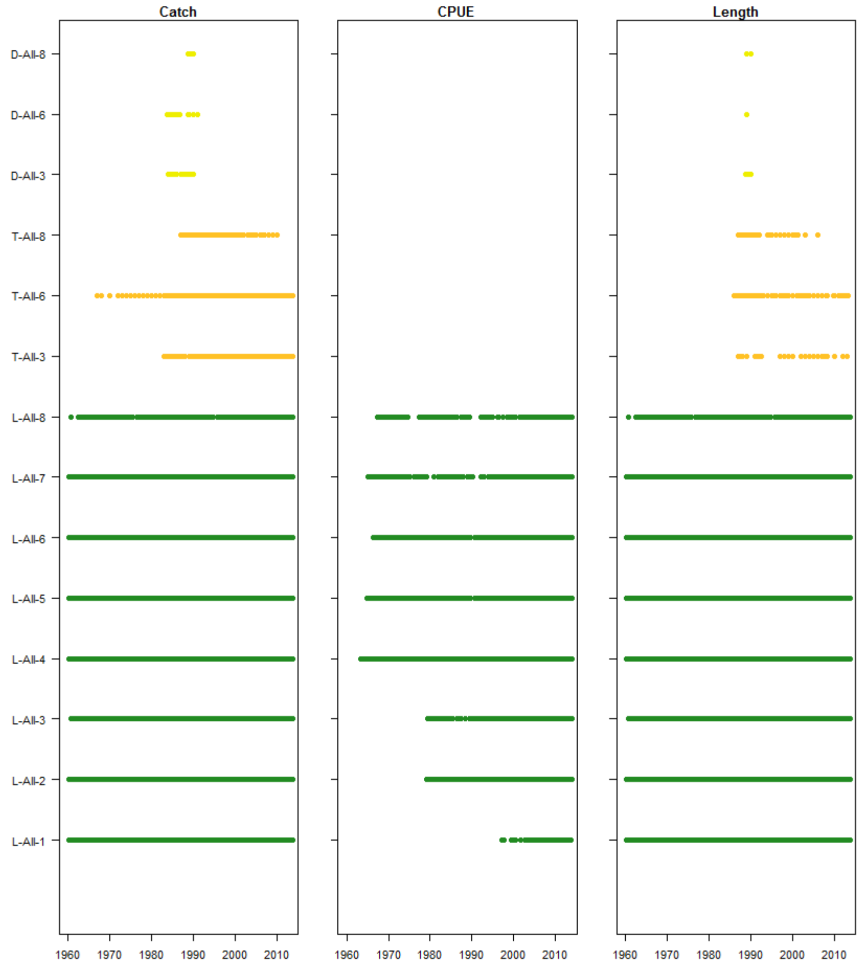


Figure 2: Presence of catch, standardised CPUE, and length frequency data by year and fishery for the reference case model. The different colours refer to longline (green), troll (orange) and other driftnet (yellow).

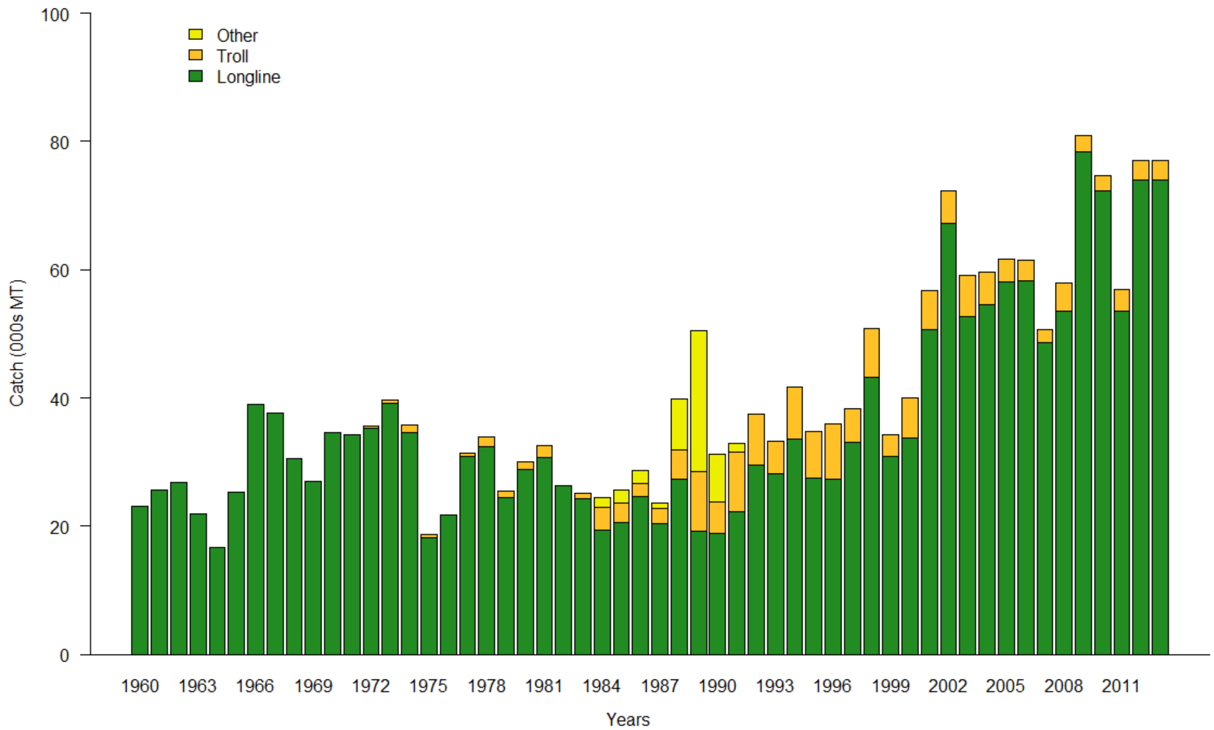


Figure 3: Total annual catch (1000s mt) by fishing gear from the reference case model.

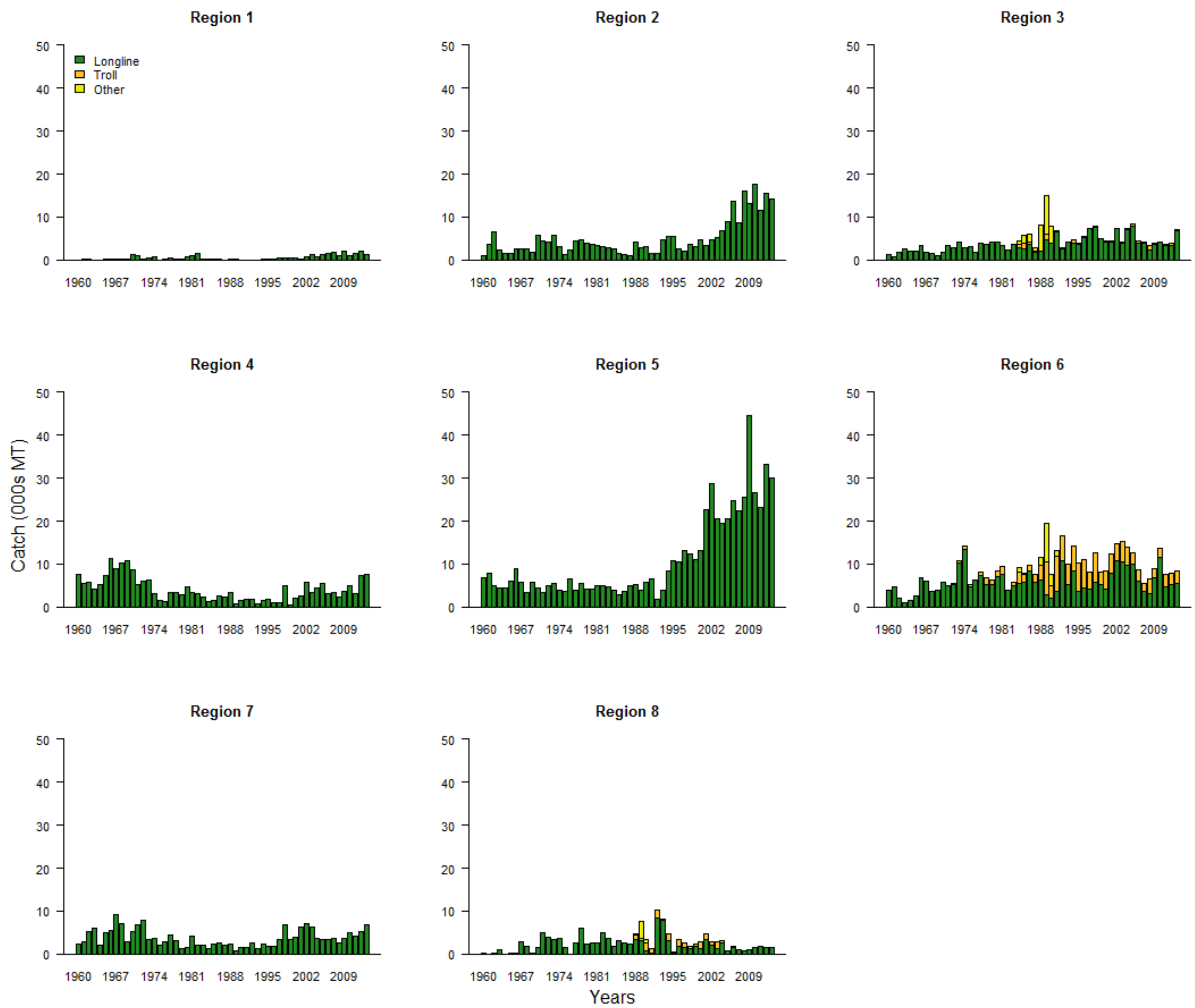


Figure 4: Total annual catch (1000s mt) by fishing method and assessment region from the reference case model.

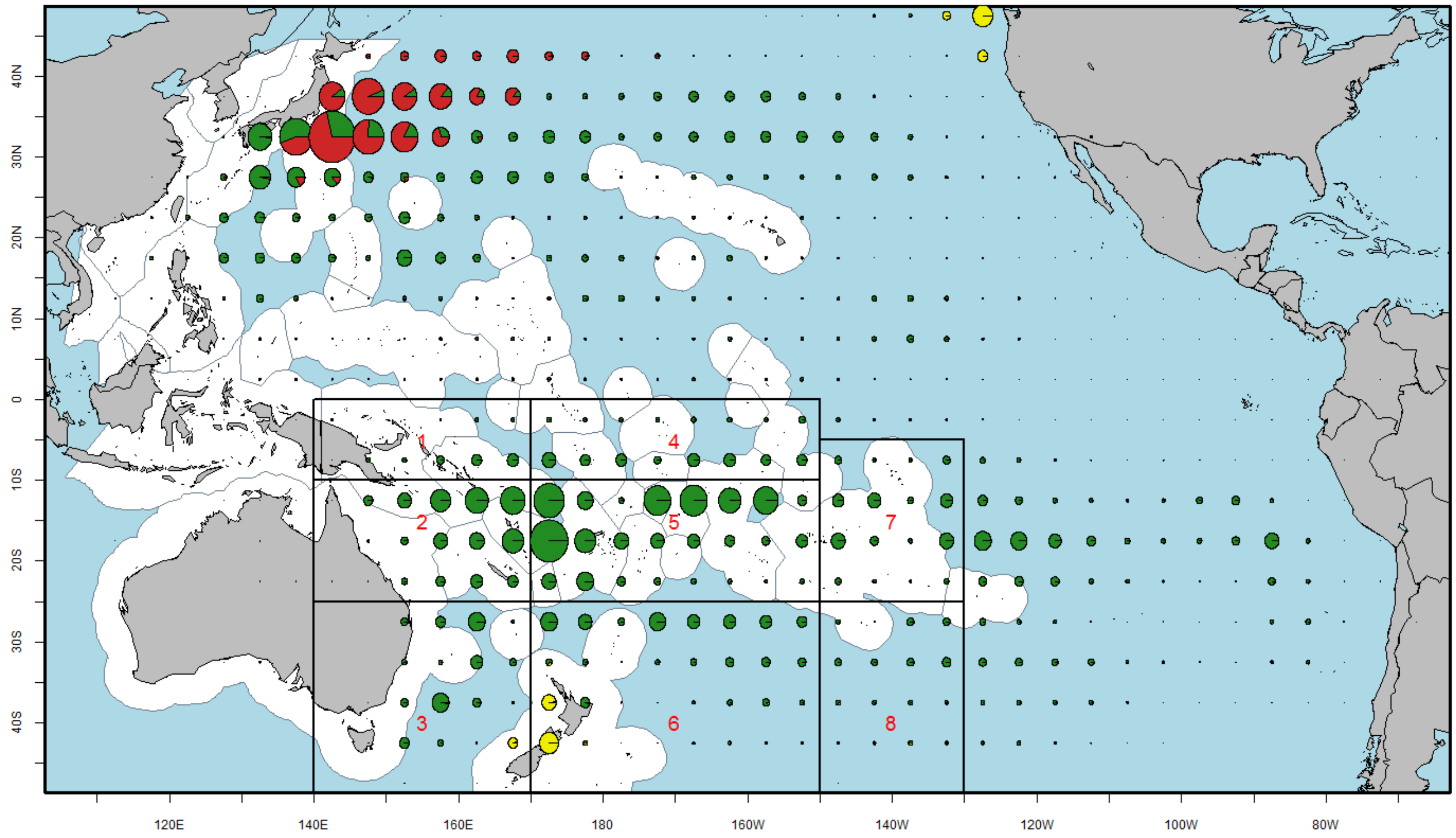


Figure 5: Catch distribution (2004-2013) by 5° square and fishing method: longline (green), pole-and-line (red), and other (yellow) for the entire Pacific Ocean. Overlayed are the regions for this assessment.

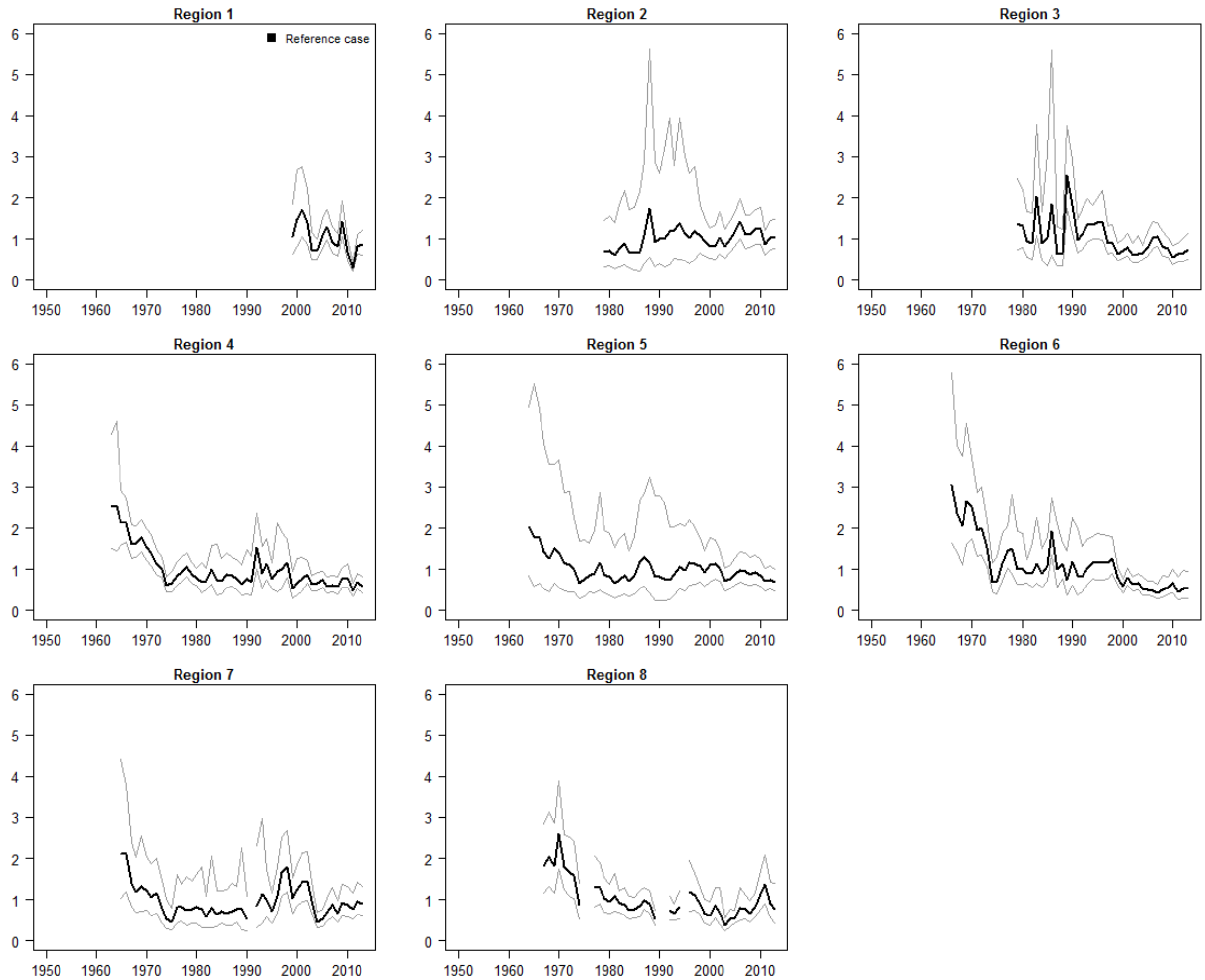


Figure 6: GLM standardised catch-per-unit-effort (CPUE) for the principal longline fisheries (L ALL 1–8) from the reference case model. Indices are scaled by the respective regional weights. See Tremblay-Boyer et al. (2015b) and Tremblay-Boyer et al. (2015a) for further details of the CPUE and region scalars respectively. The light grey lines prepresent the 95% confidence intervals assumed in the assessment.

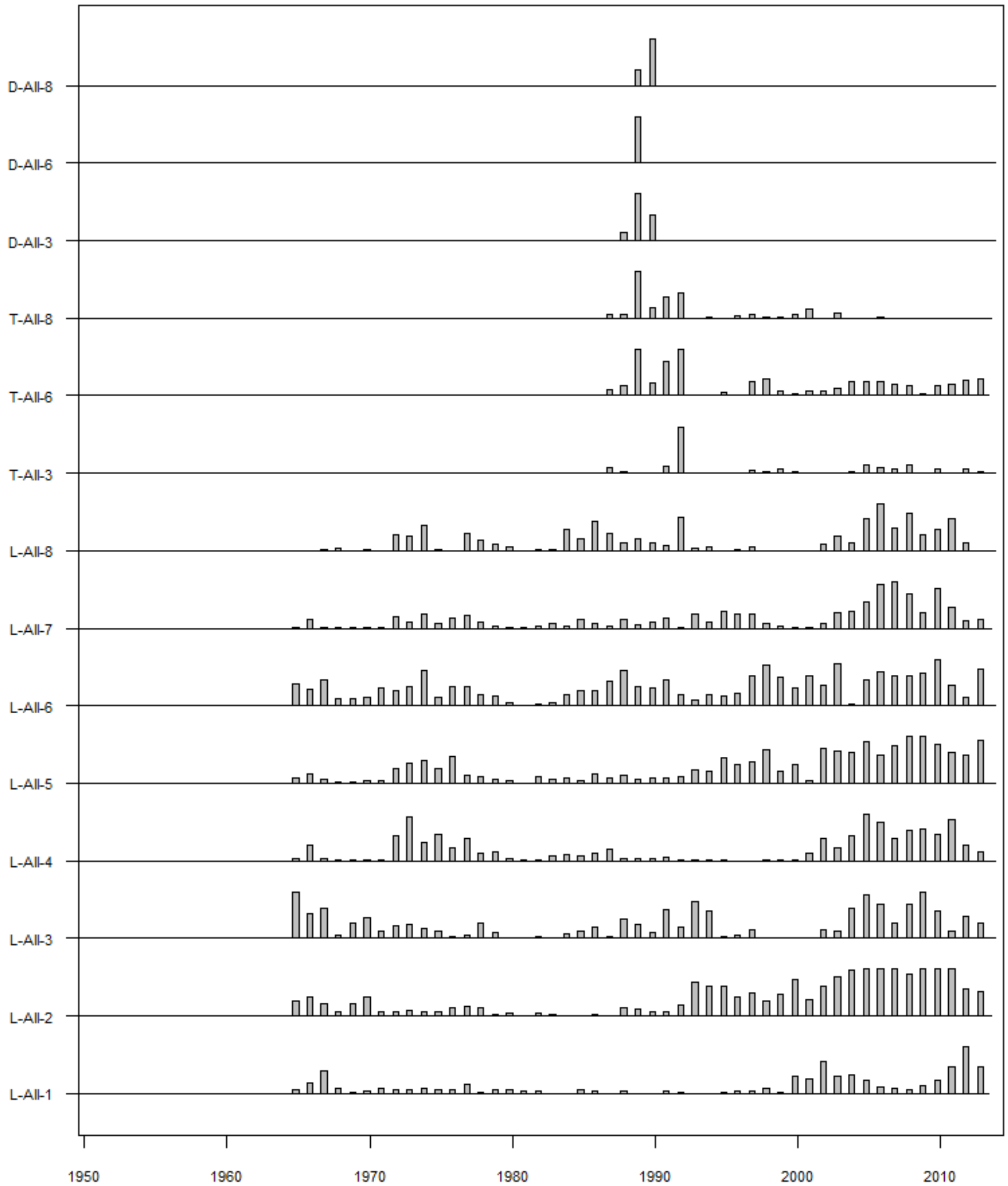


Figure 7: Number of length frequency samples from the reference case model. The maximum value is 22,475, but note that in the reference case model a maximum sample size of 1000 is allowed.

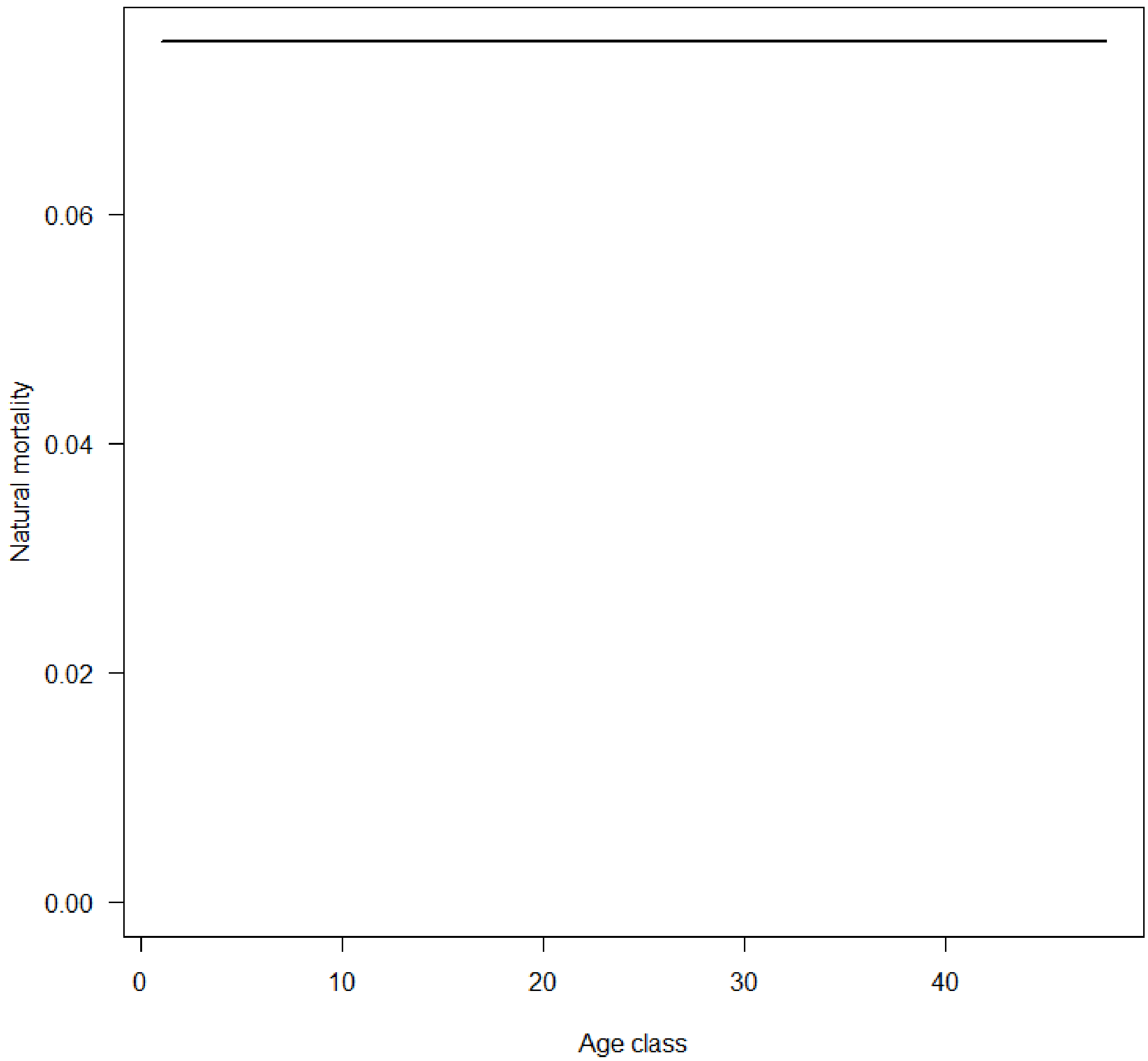


Figure 8: Quarterly natural mortality-at-age as assumed in the reference case and estimated in the one-change sensitivity

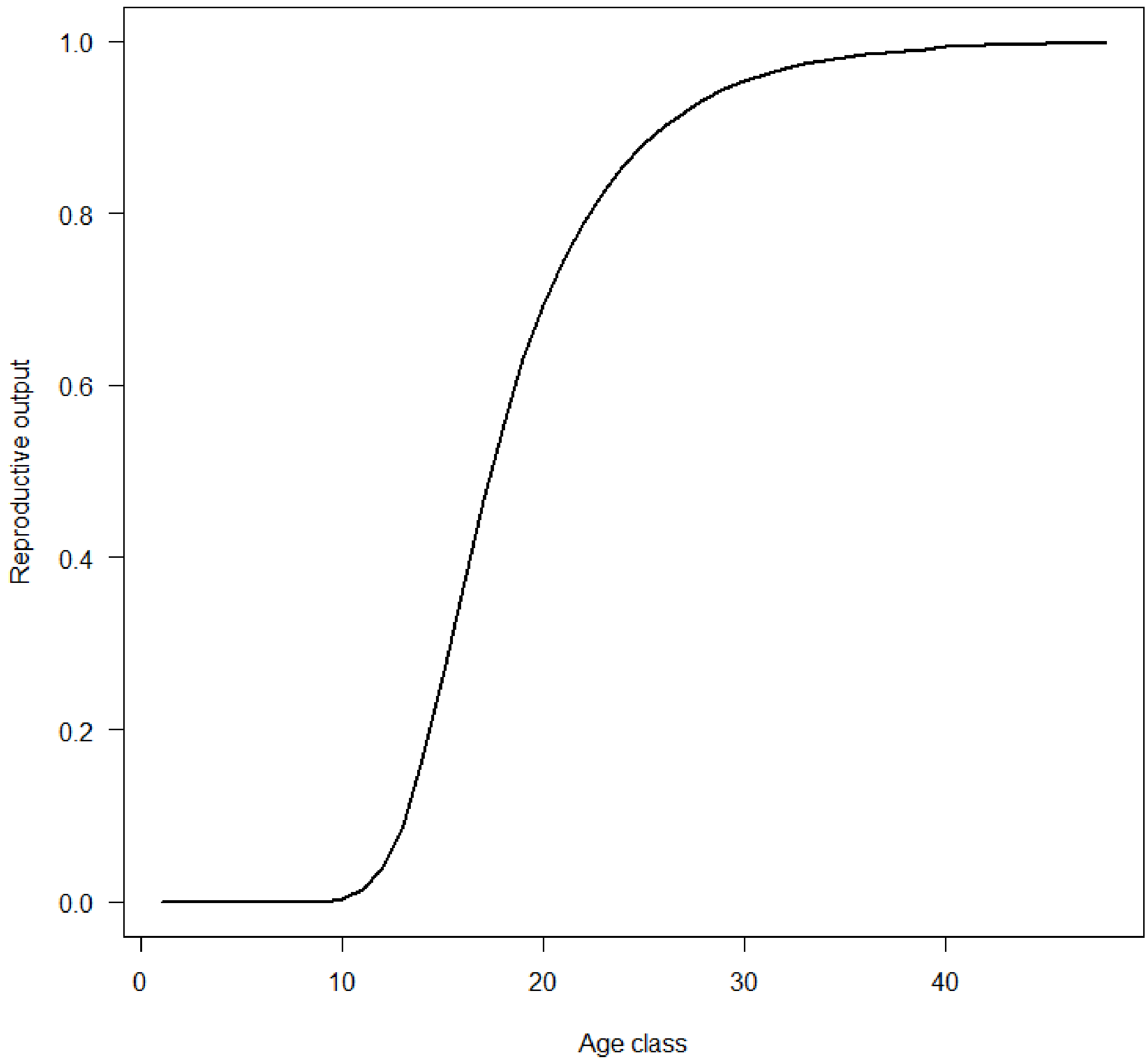


Figure 9: Maturity-at-age as assumed in the reference case model.

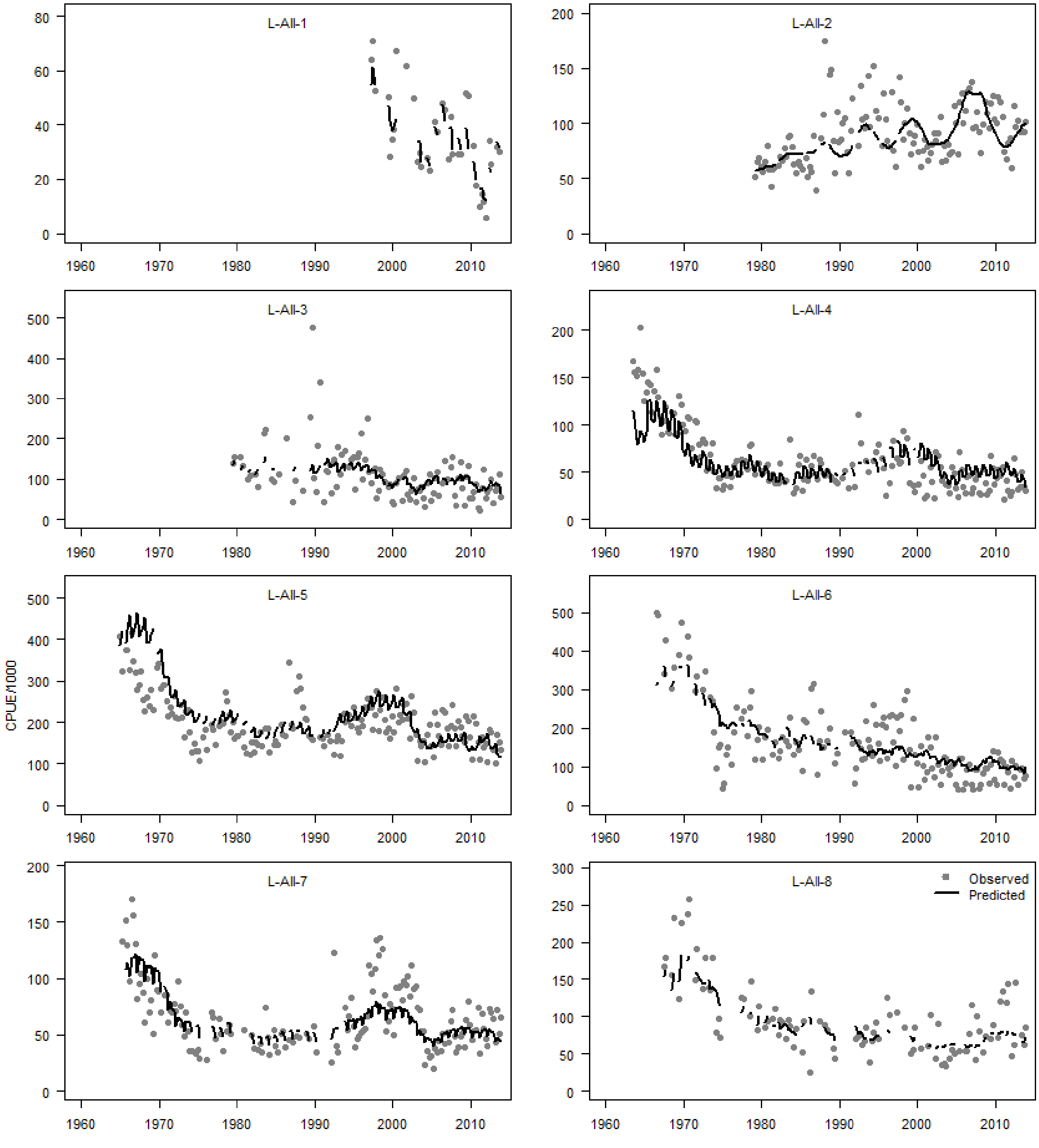


Figure 10: Observed (grey dots) and predicted (line) CPUE for the longline fisheries from the reference case.

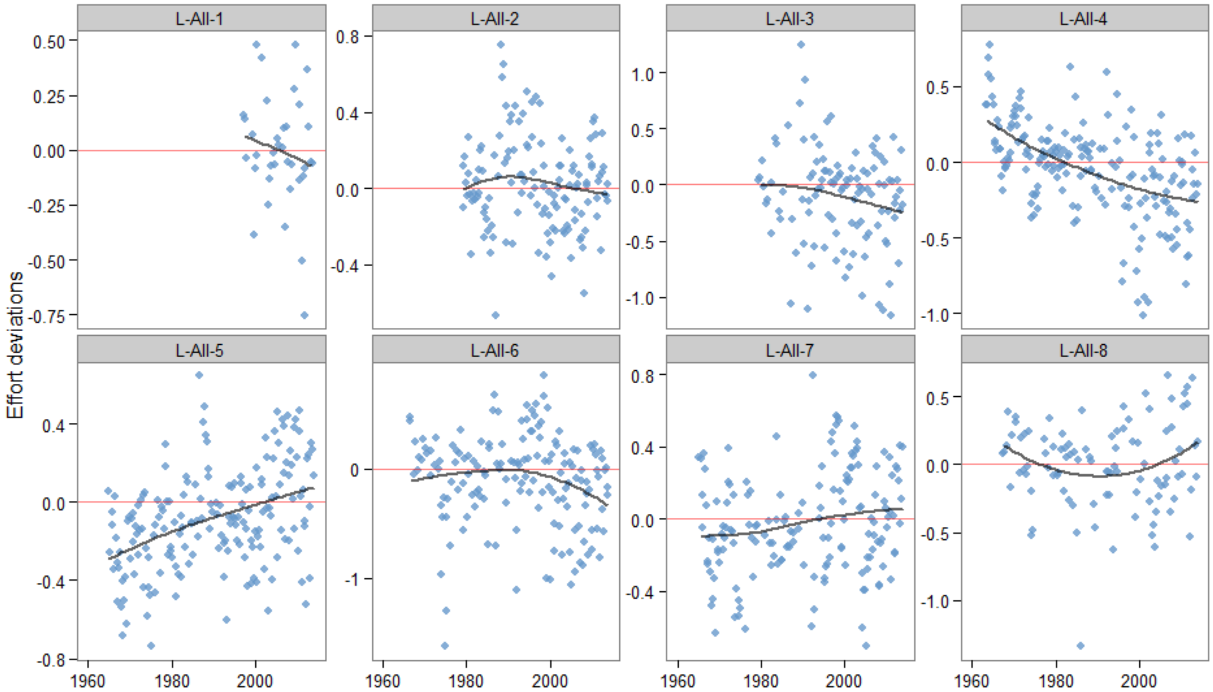


Figure 11: Effort deviations by time period for each LL-ALL fishery for the reference case. The dark line represents a lowess smoothed fit to the effort deviations. A small number of values lie outside the bounds of the plot.

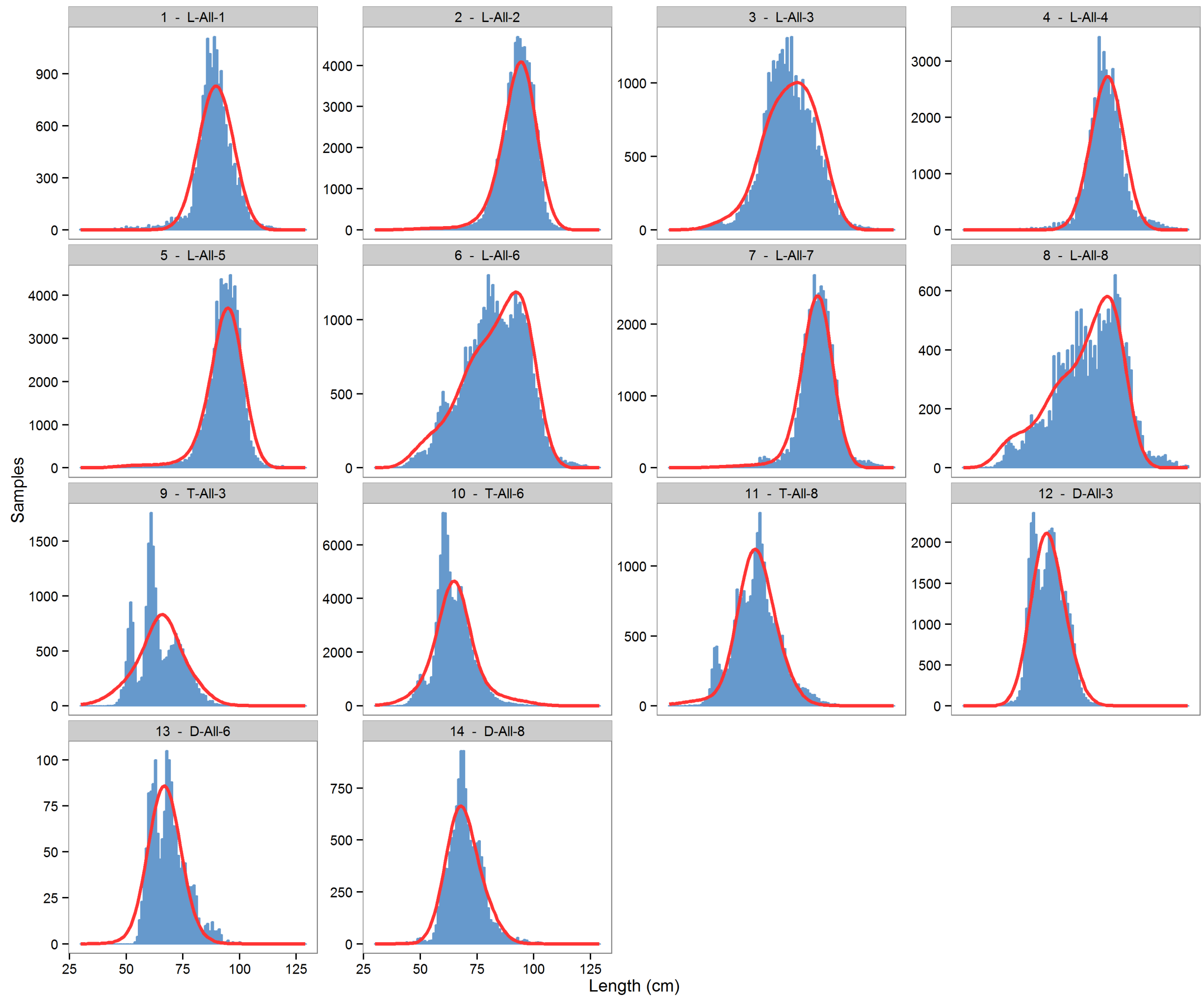


Figure 12: Composite (all time periods combined) observed (blue histograms) and predicted (red line) catch at length for all fisheries with samples for the reference case.

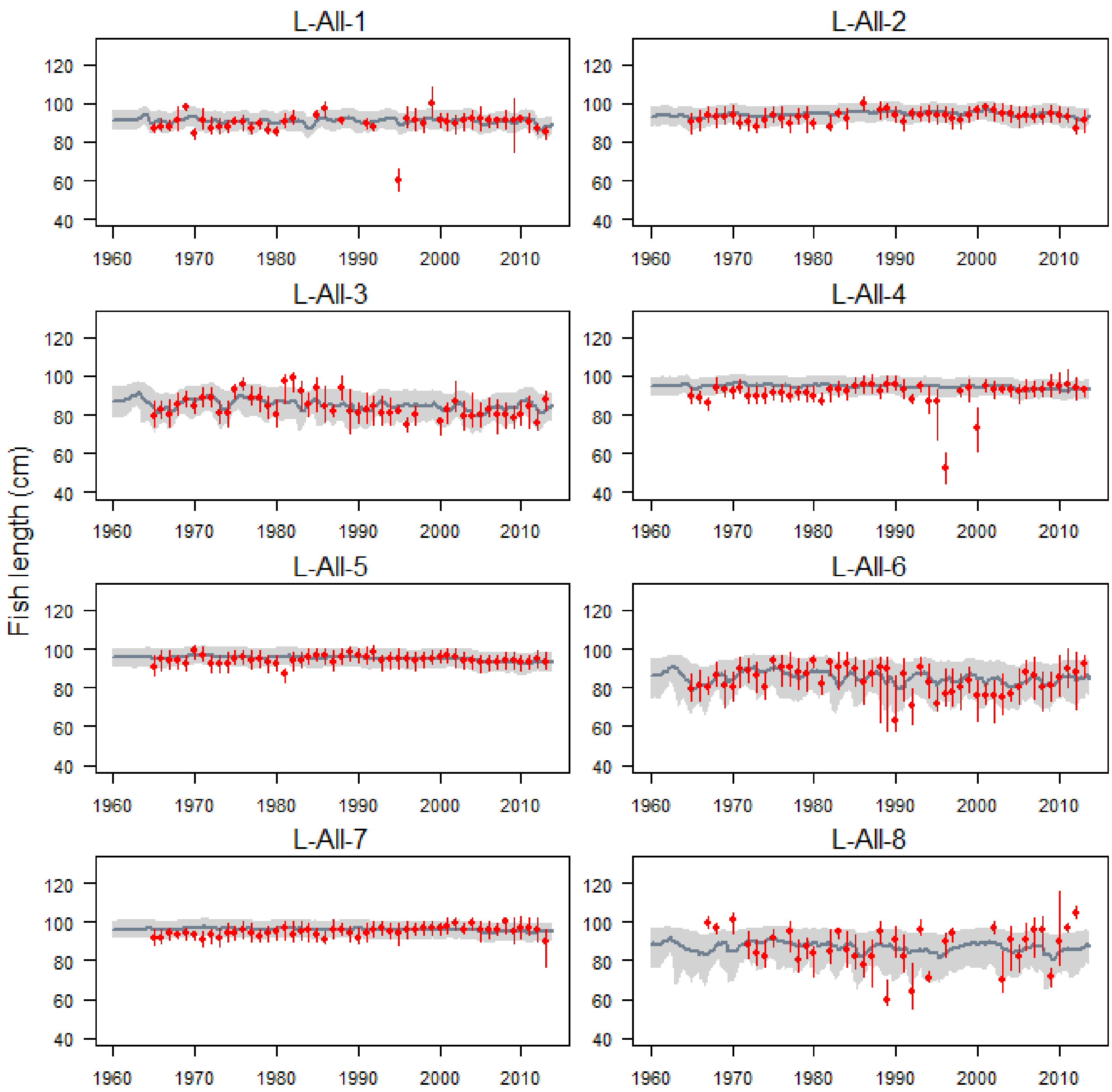


Figure 13: A comparison of the observed (red points) and predicted (grey line) median fish length (FL, cm) for all fisheries with samples for the reference case. The confidence intervals represent the values encompassed by the 25% and 75% quantiles. Sampling data are aggregated by year and only length samples with a minimum of 30 fish per year are plotted.

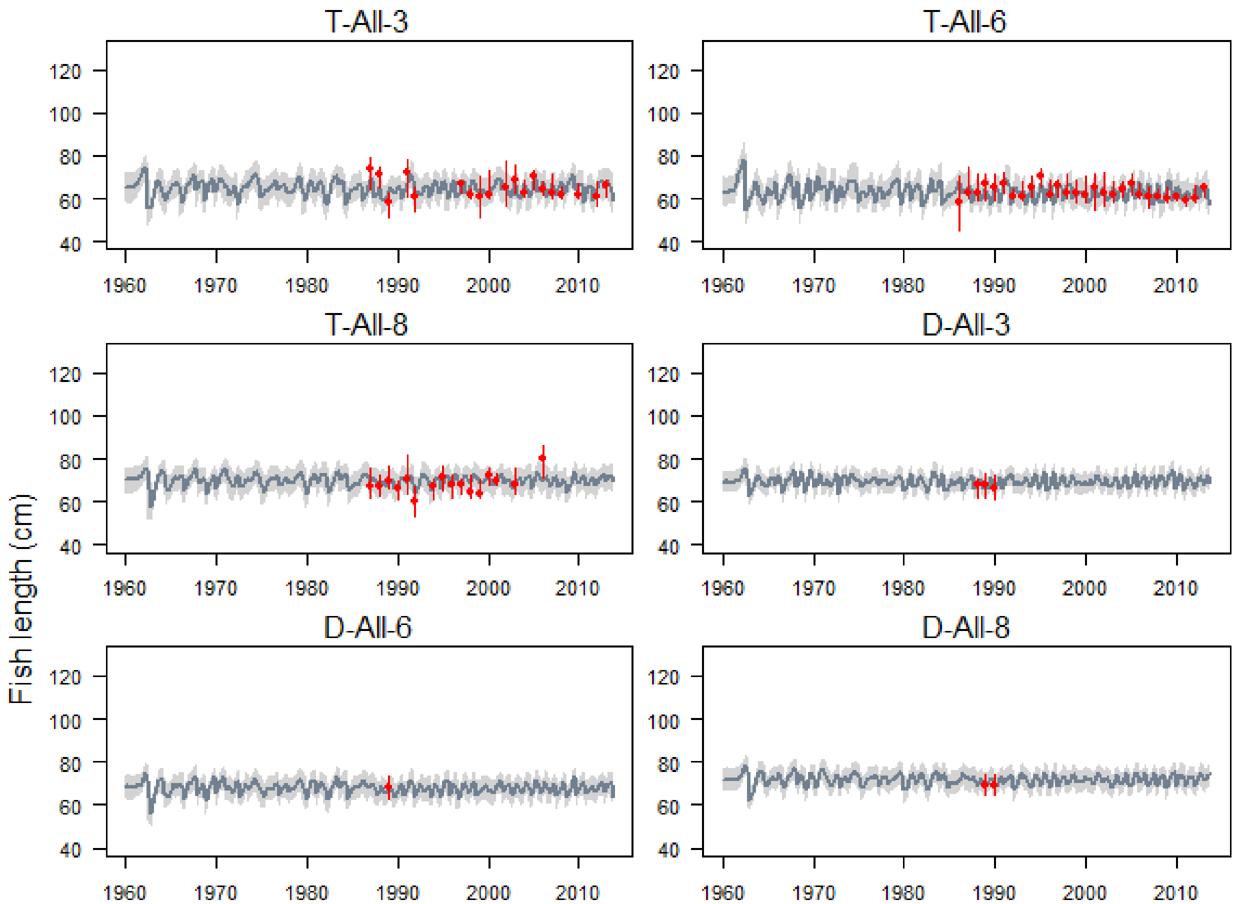


Figure 13: Continued

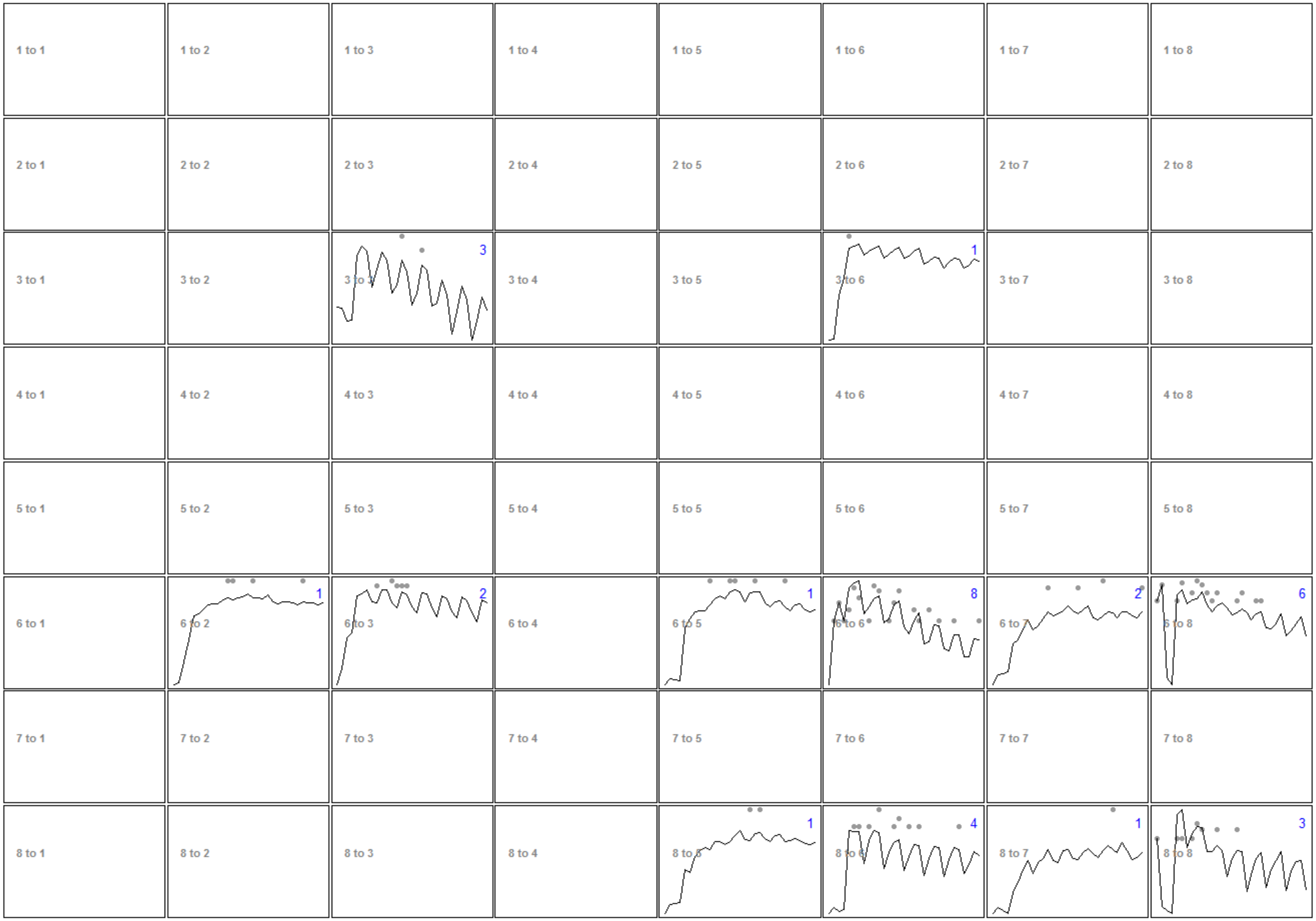


Figure 14: Observed and predicted tag returns (on the log-scale) by time-at-liberty showing the region of release (y-axis) and recapture (x-axis) for the reference case model. The y-axis is the number of quarters at liberty and the number in blue shows the maximum number of recaptures in a time period for that region-region combination. Both observed and predicted values exclude recaptures from the 4 quarter mixing period after release.

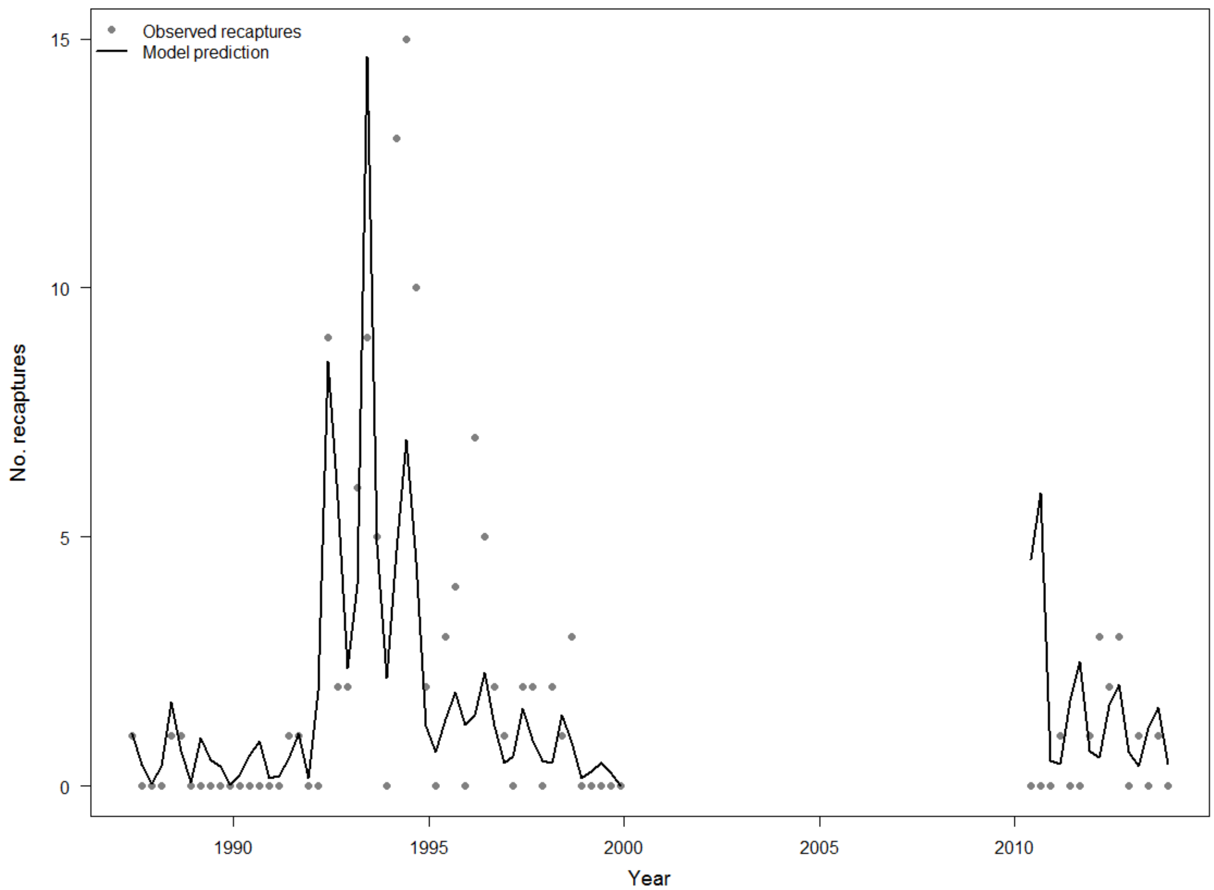


Figure 15: Observed and predicted tag returns (with tags recaptured during the four-quarter mixing period excluded) over time for the reference case model across all tag release events

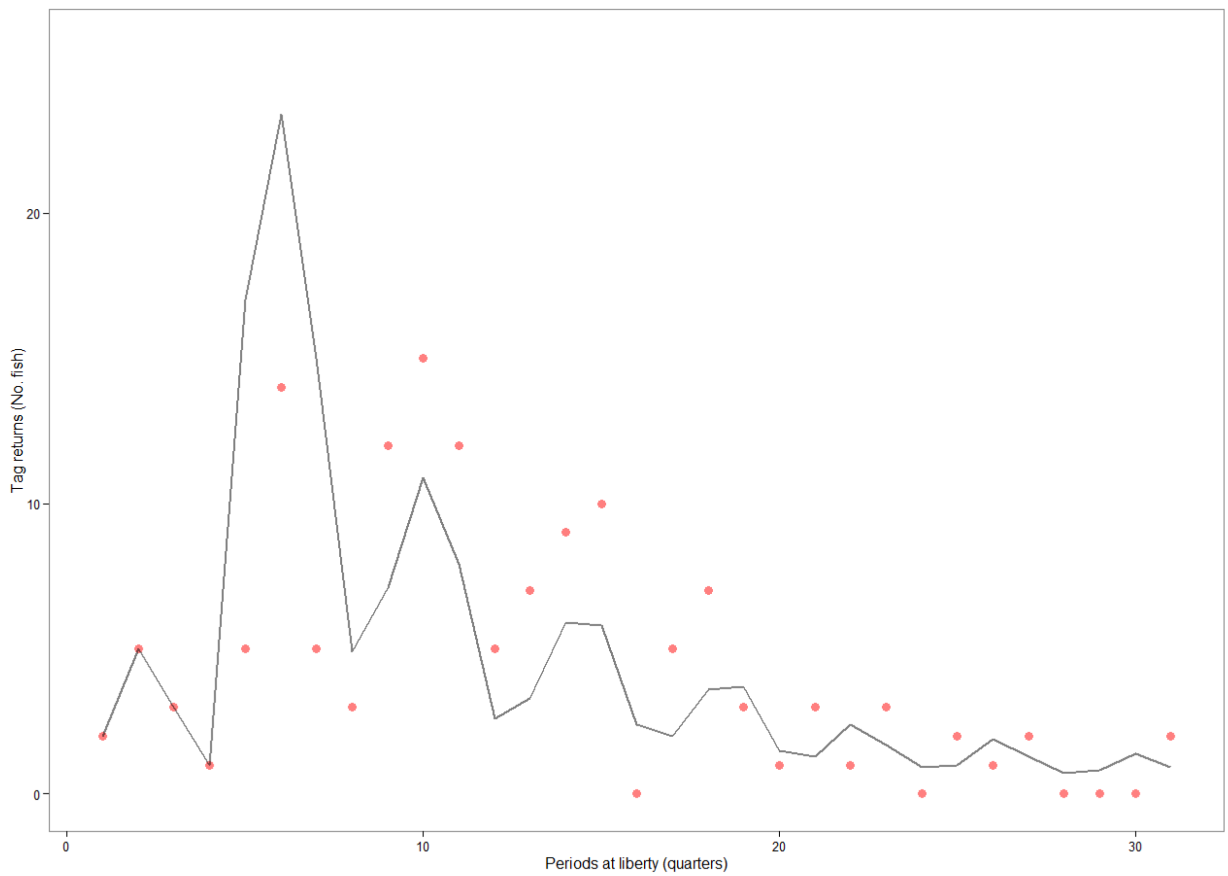


Figure 16: Observed and predicted tag attrition for the reference case across all tag release events. The exact correspondence between observed and predicted values for periods 1–4 is a direct result of the mixing assumption.

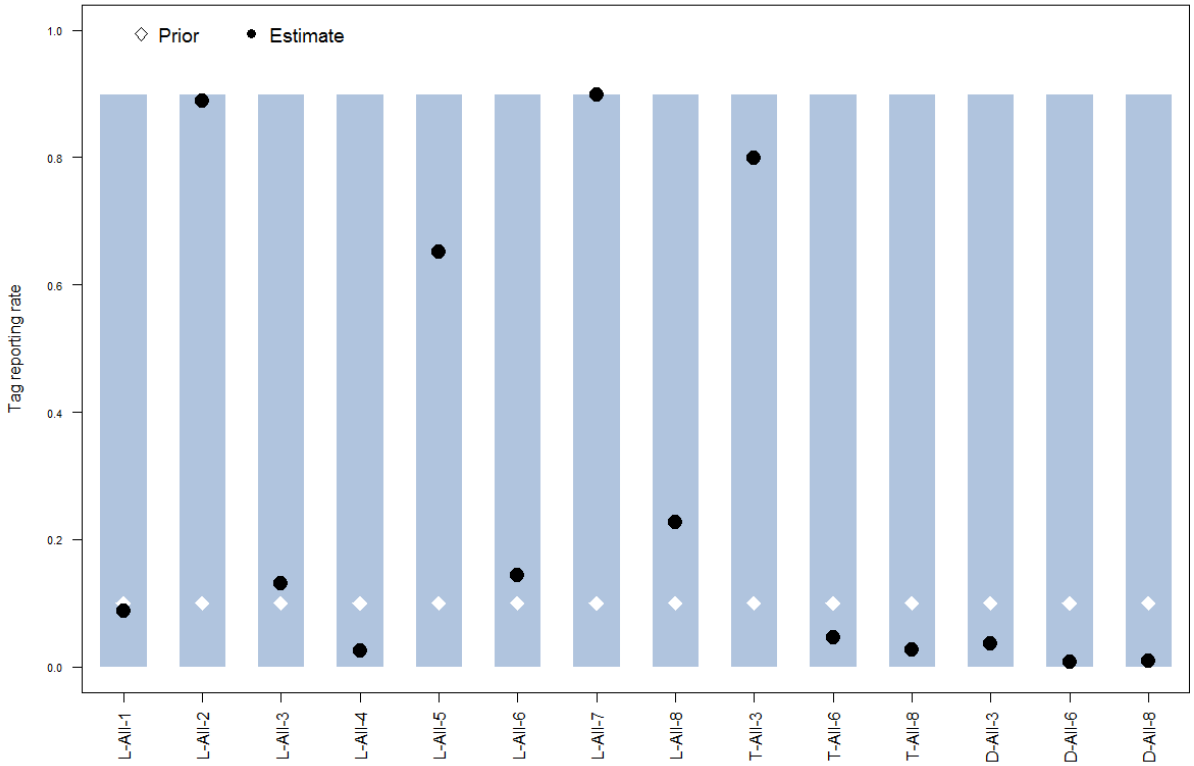


Figure 17: Estimated reporting rates for the reference case. Reporting rates can be estimated separately for each release program and recapture fishery group (histograms). The prior mean ±1.96 SD is also shown for each reporting rate group and the assumed upper bound.

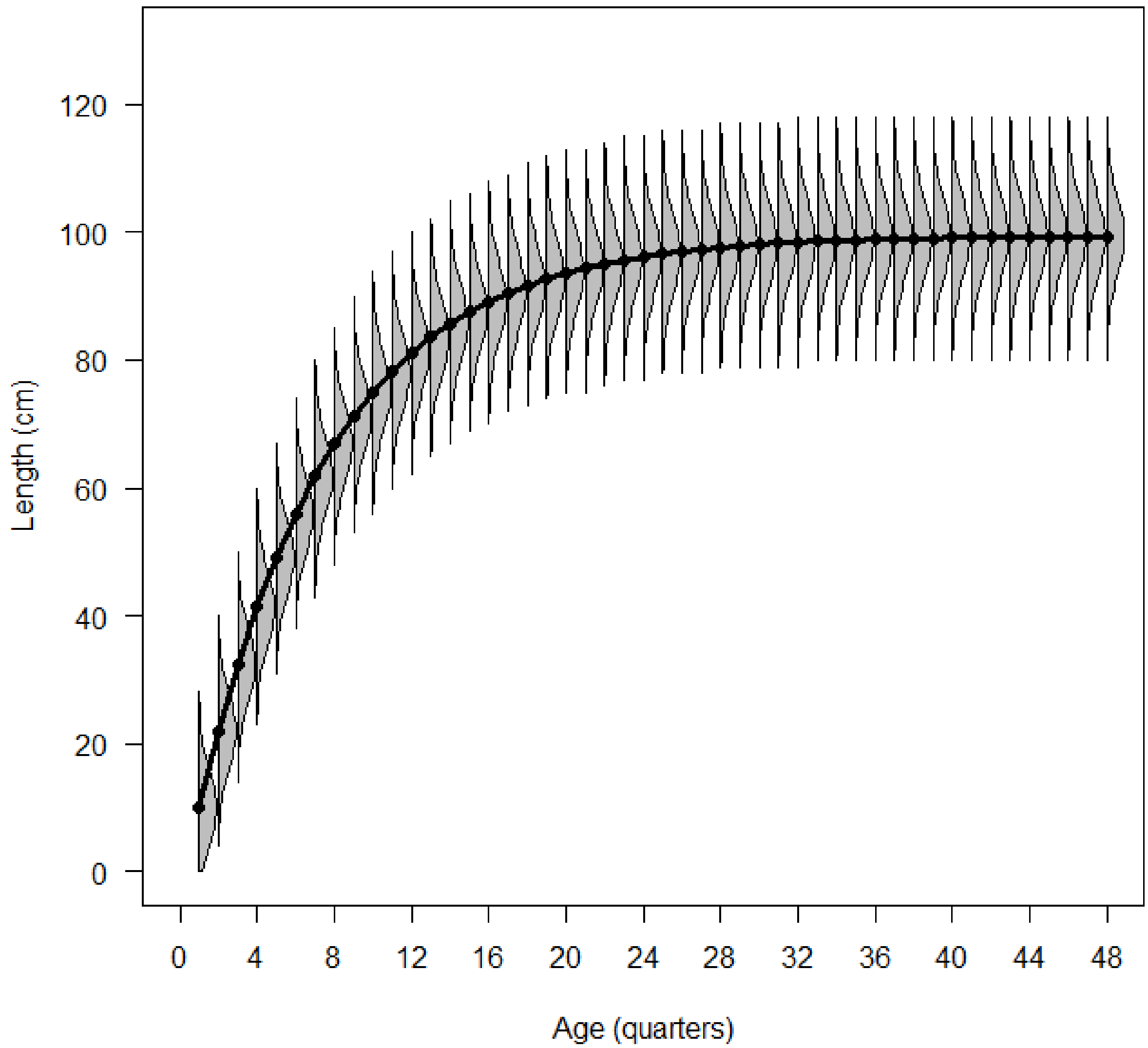


Figure 18: Estimated growth for the reference case. The black line represents the estimated mean fork length (FL, cm) at age and the grey area represents the estimated distribution of length at age.

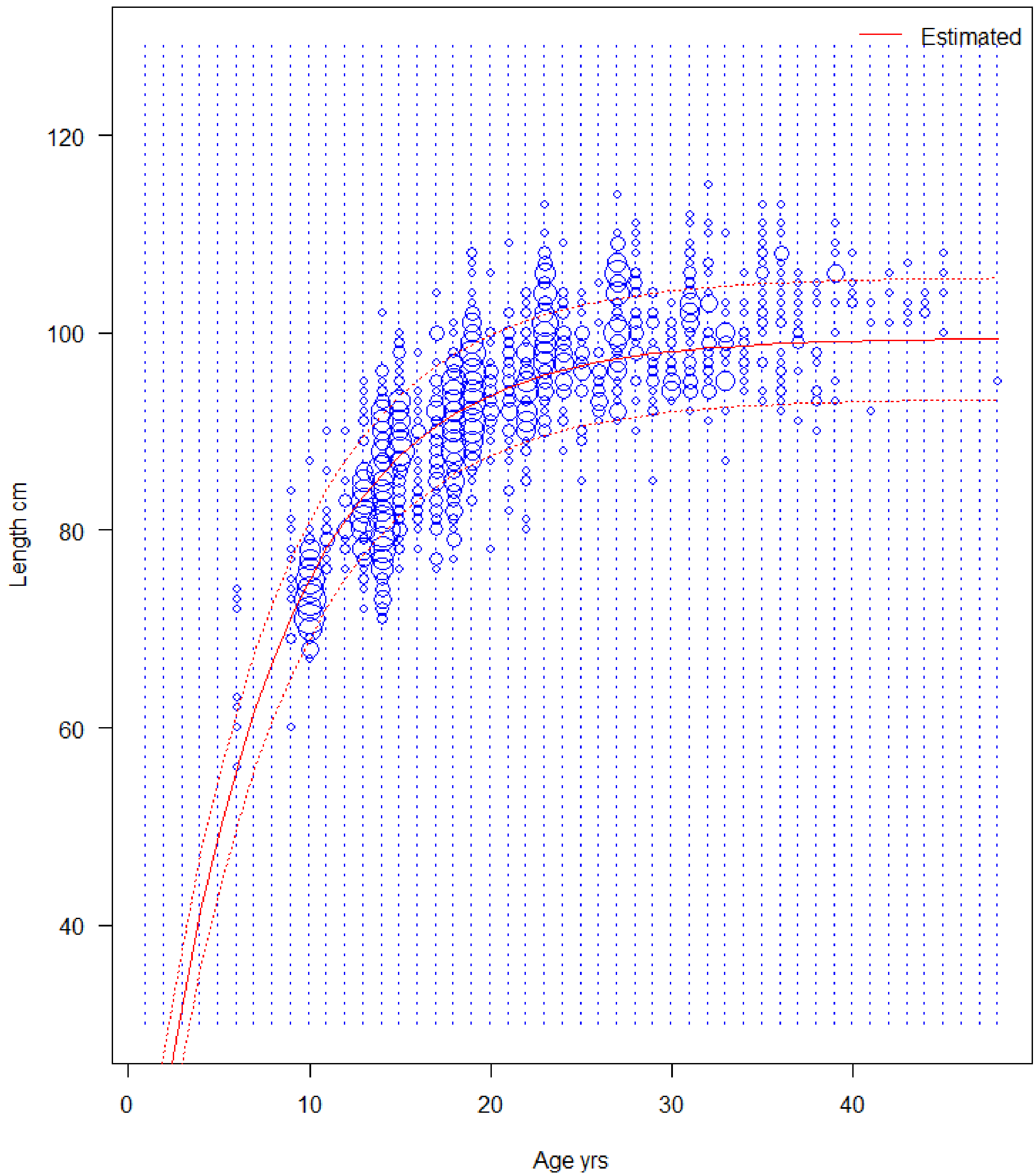


Figure 19: Fit to the age-length observations.

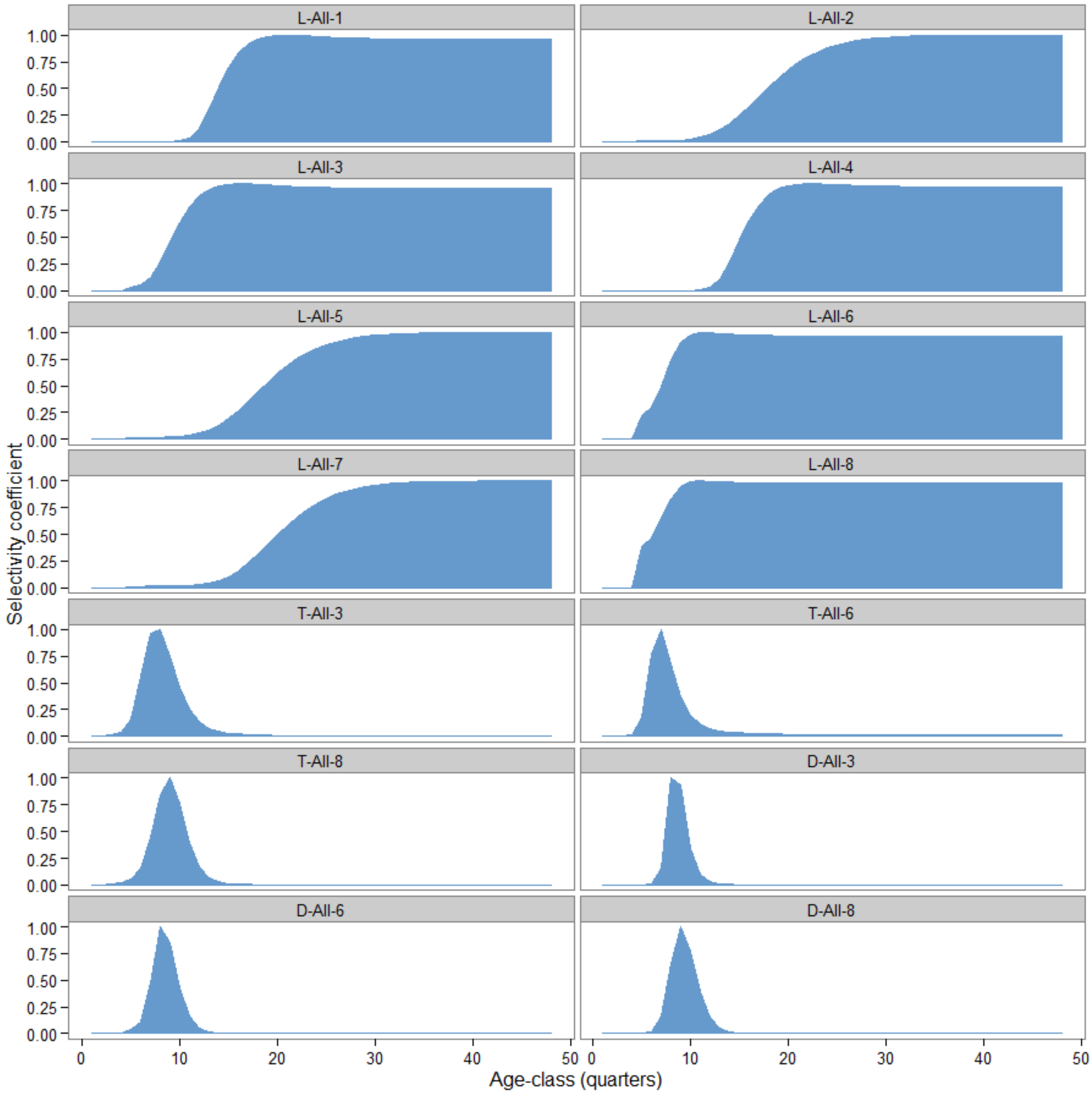


Figure 20: Age-specific selectivity coefficients by fishery.

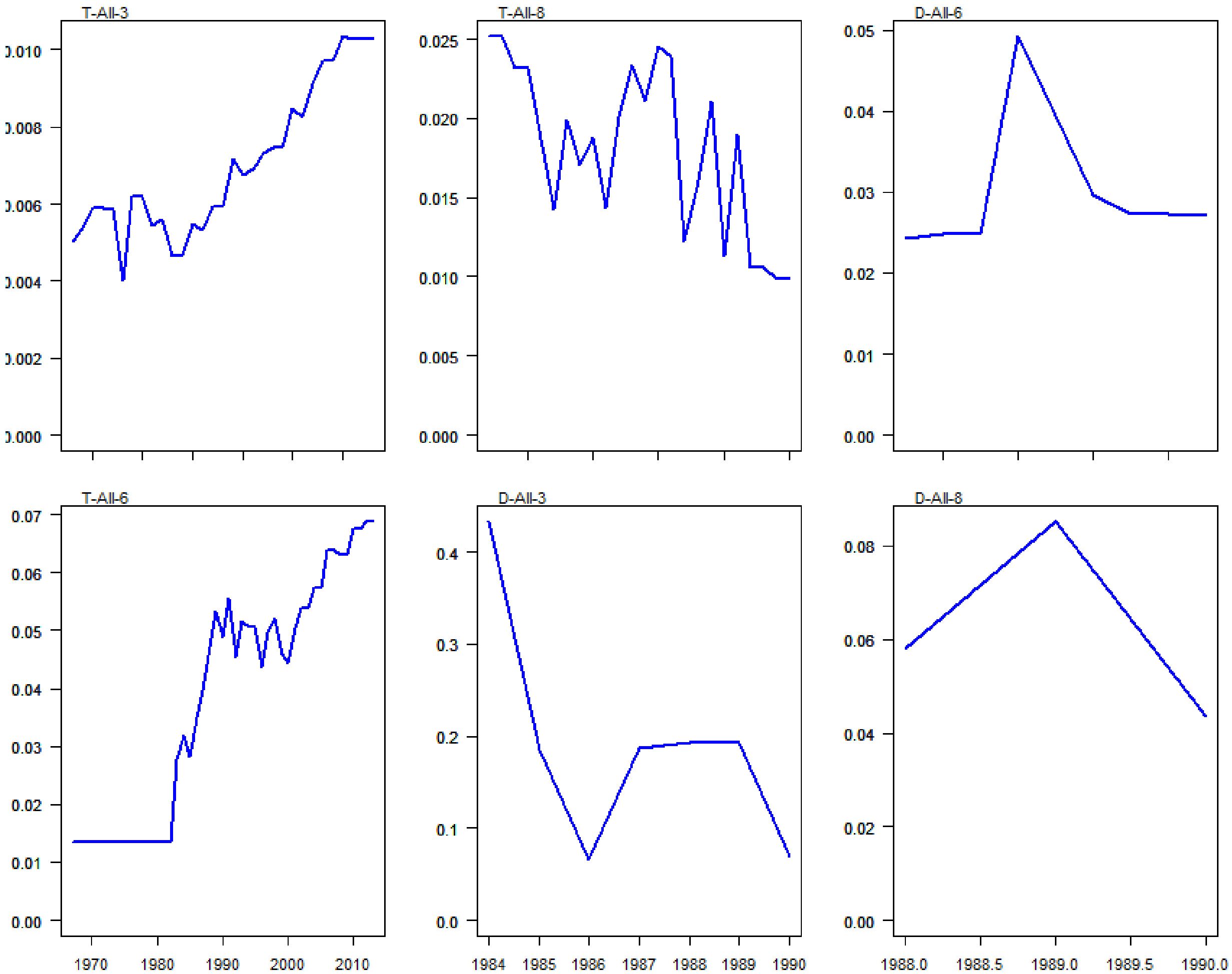


Figure 21: Estimated catchability time-series for those fisheries assumed to have random walk in catchability.

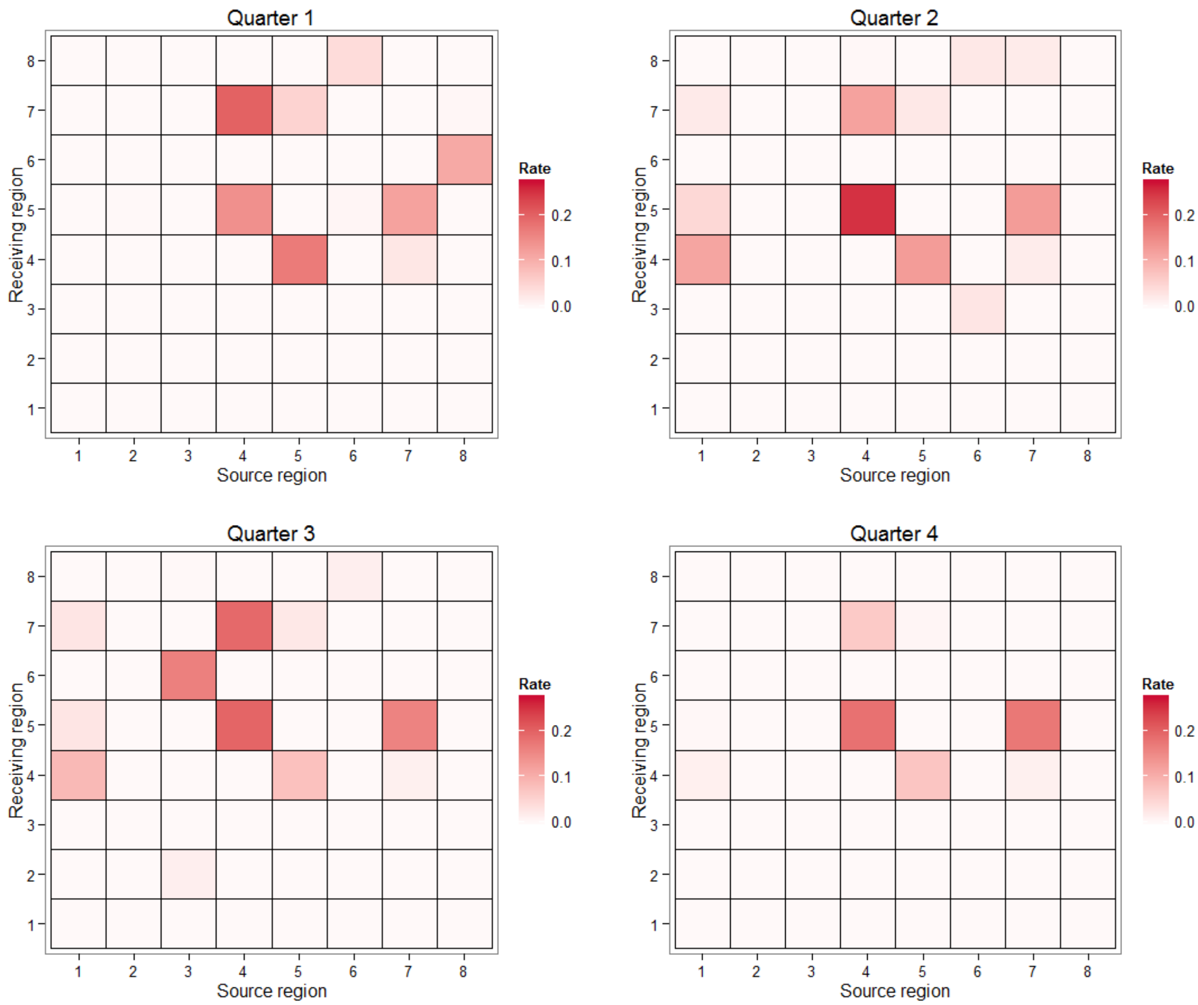


Figure 22: Estimated quarterly movement coefficients for the reference case model. The colour of the tile indicates the magnitude of the movement rate.

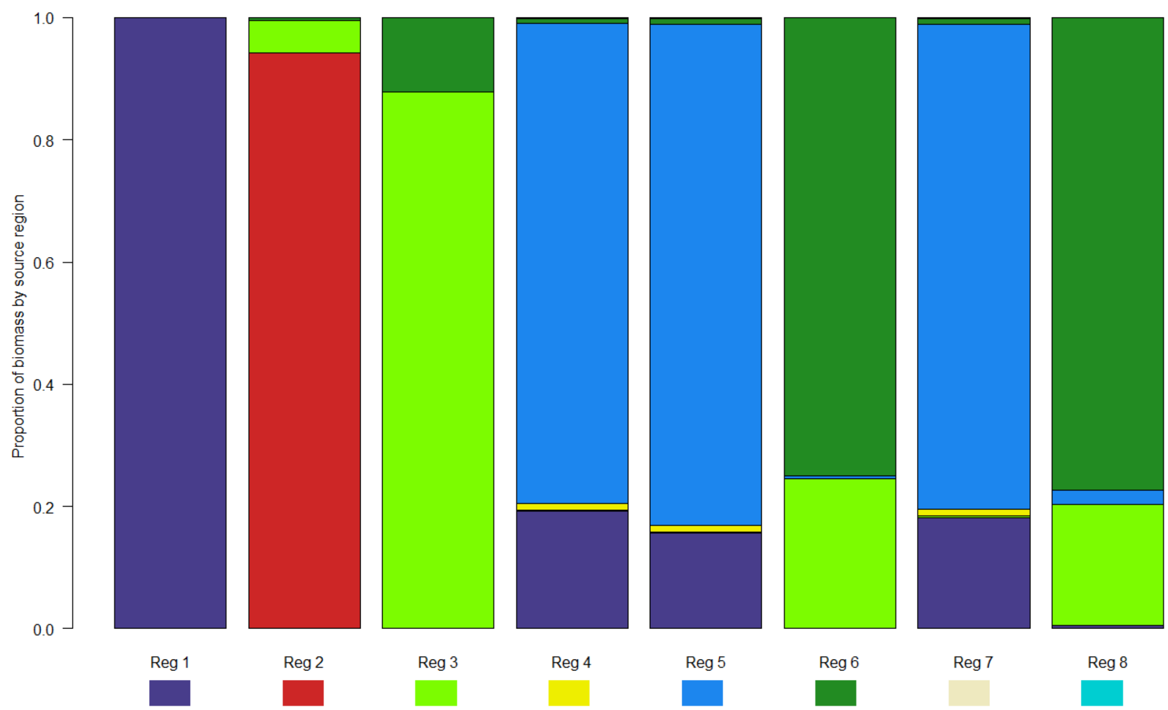


Figure 23: Proportional distribution of total biomass (by weight) in each region apportioned by the source region of the fish for the reference case. The colour of the home region is presented below the corresponding label on the x-axis. The biomass distributions are calculated based on the long-term average distribution of recruitment between regions, estimated movement parameters, and natural mortality. Fishing mortality is not taken into account.

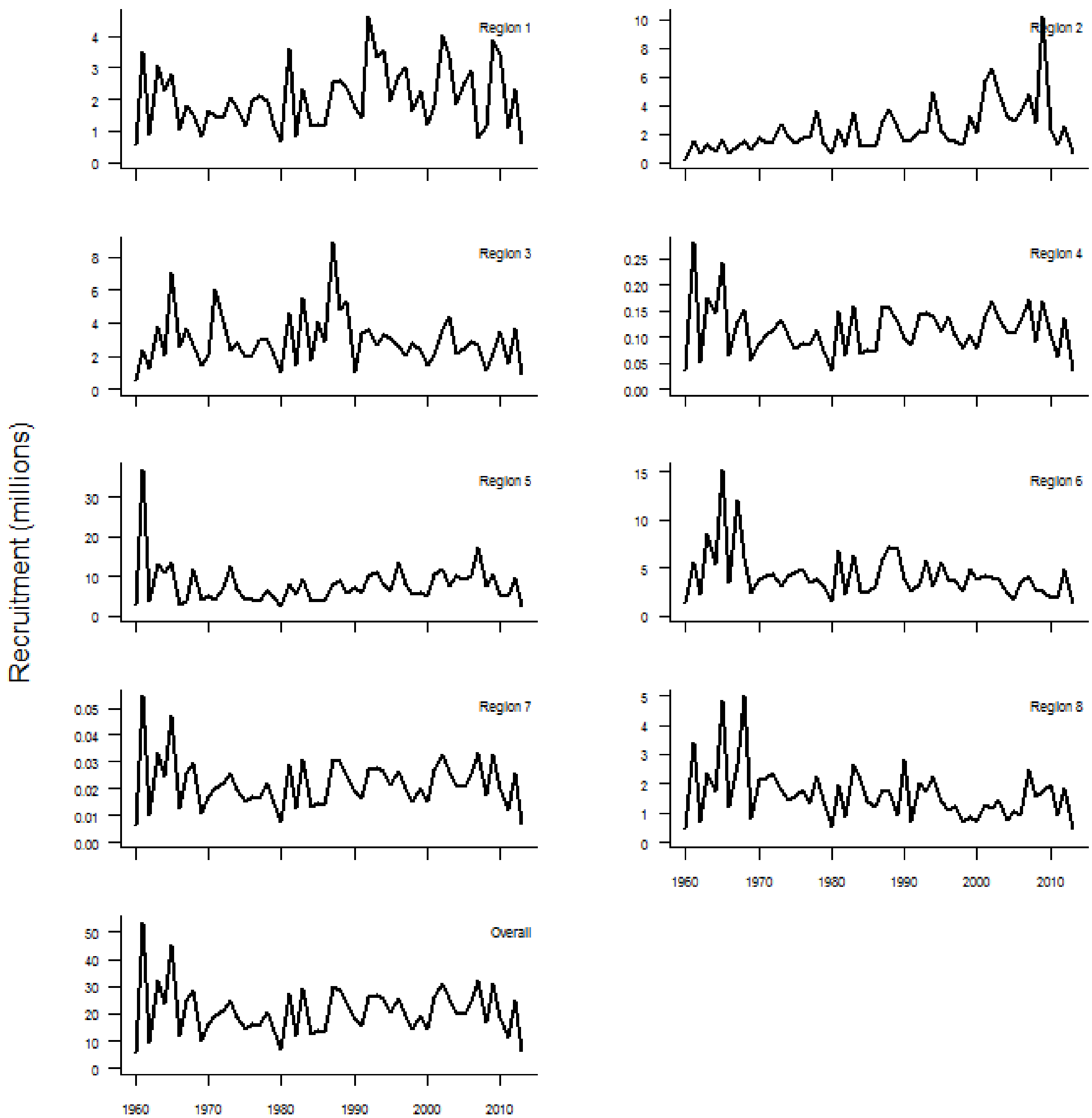


Figure 24: Estimated annual recruitment (millions) by region and overall for the reference case.

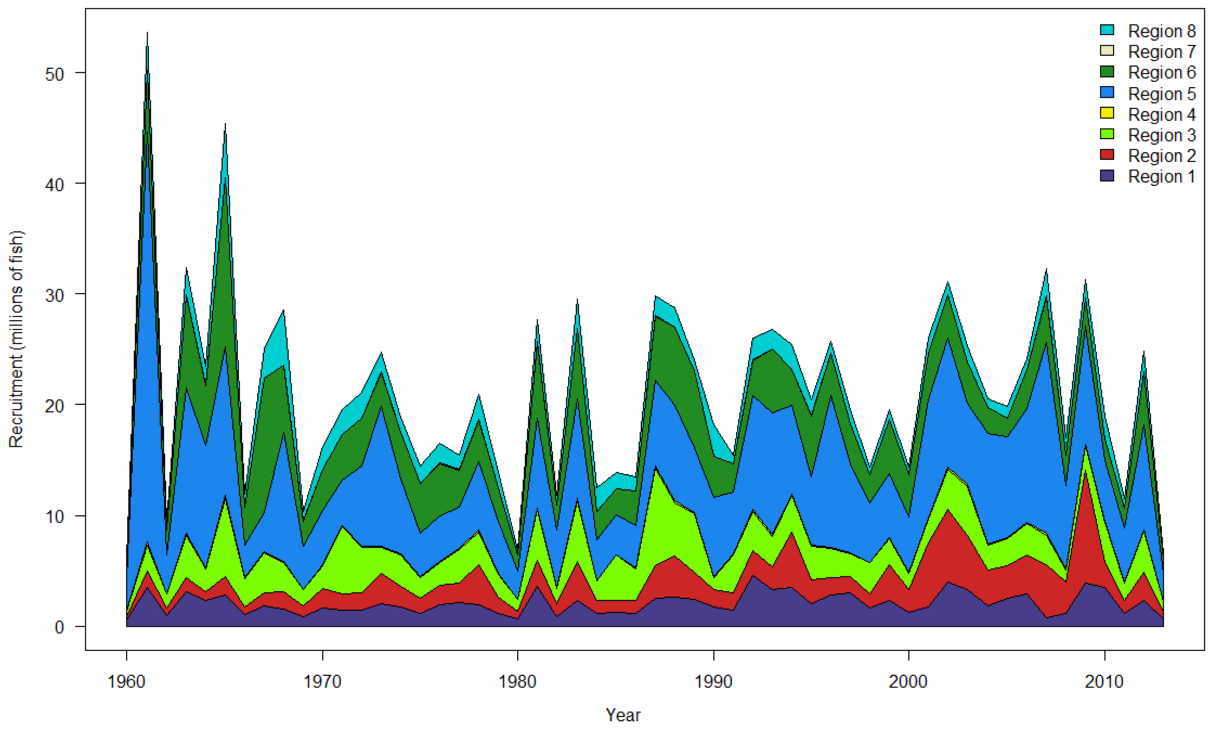


Figure 25: Estimated annual average recruitment by model region for the reference case.

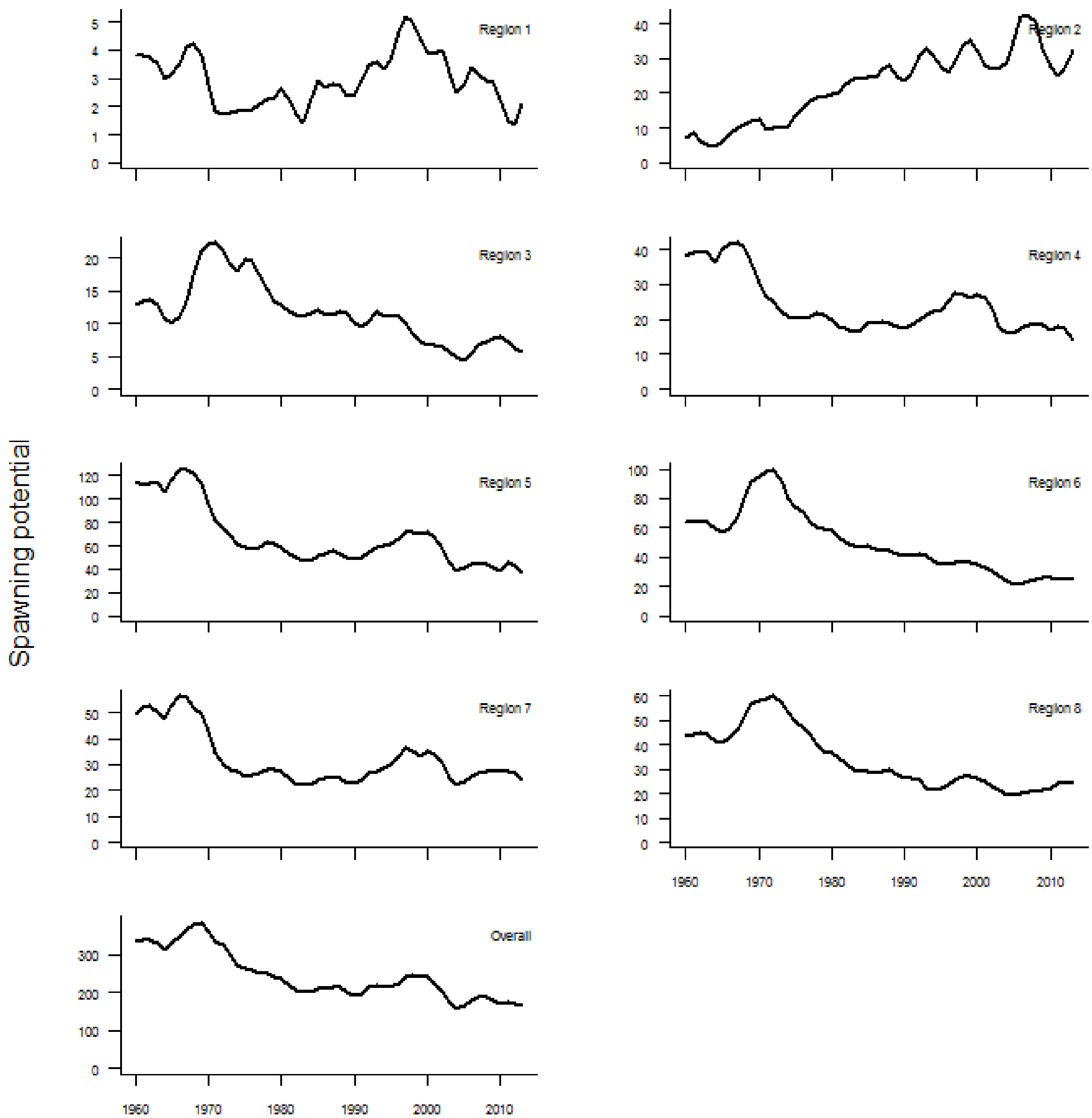


Figure 26: Estimated annual average spawning potential by region and overall for the reference case.

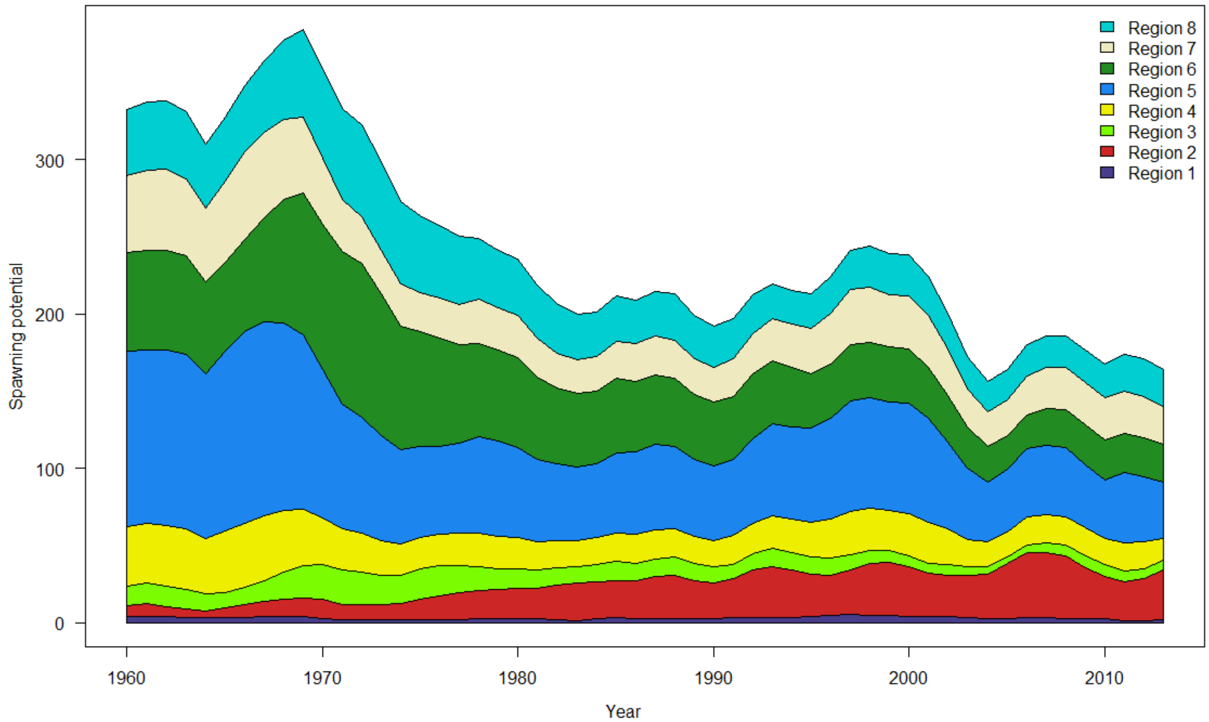


Figure 27: Estimated annual average spawning potential by model region for the reference case.

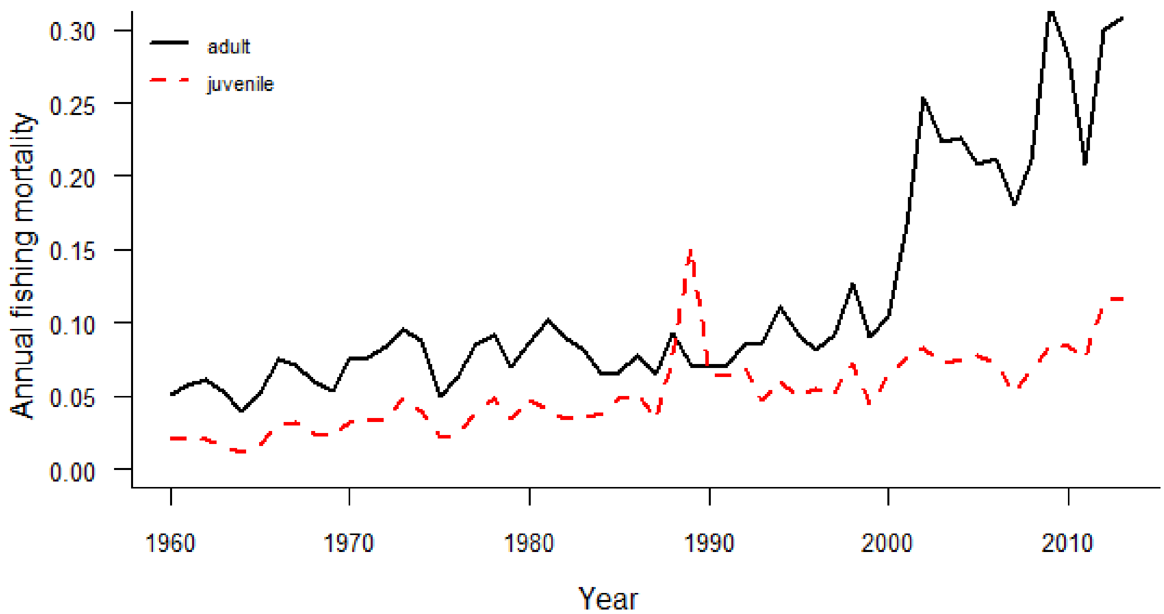


Figure 28: Estimated annual average juvenile and adult fishing mortality for the reference case.

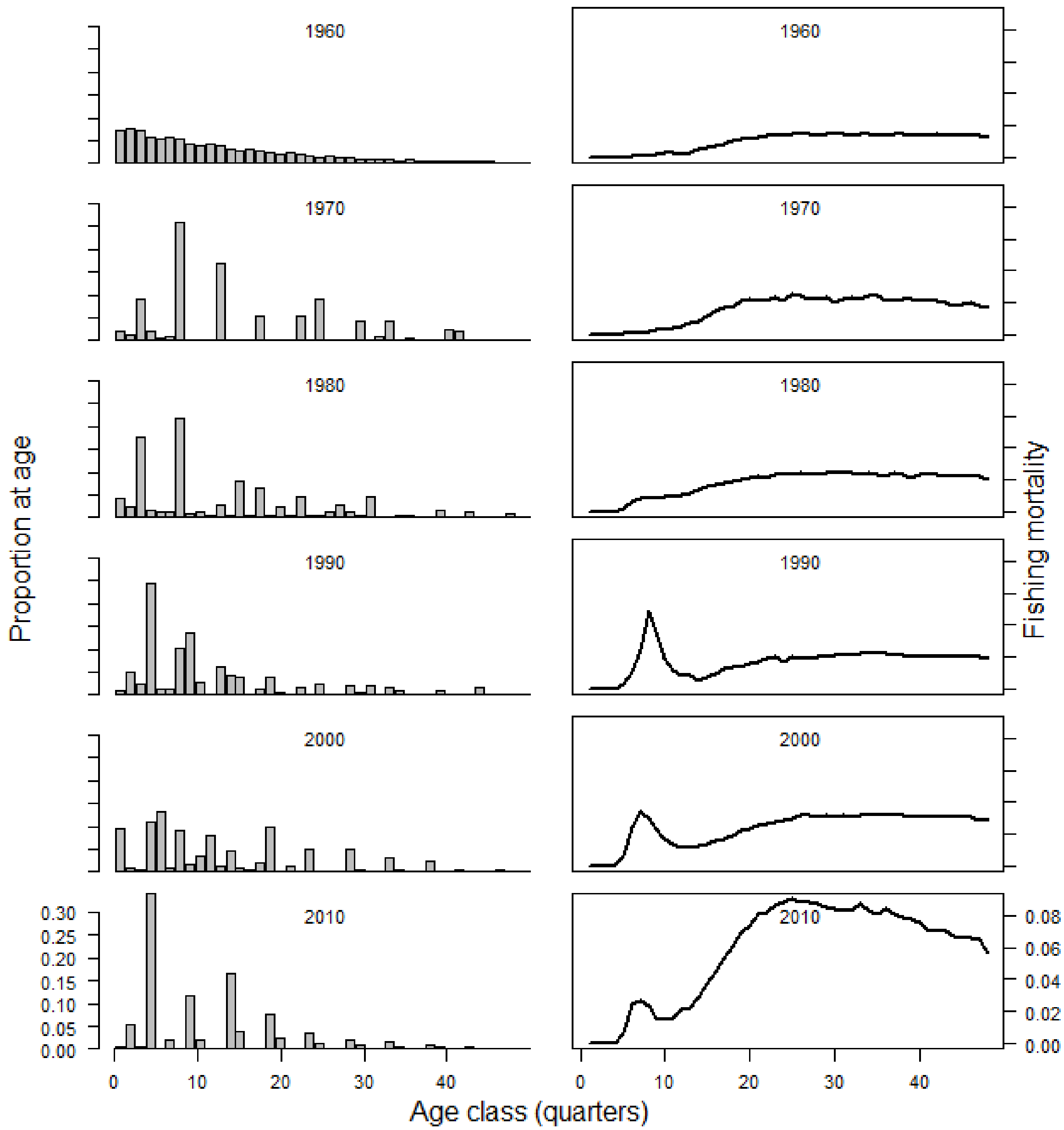


Figure 29: Estimated proportion at age (quarters) for the south Pacific albacore tuna population (left) and fishing mortality at age (right) by year at decade intervals for the reference case.

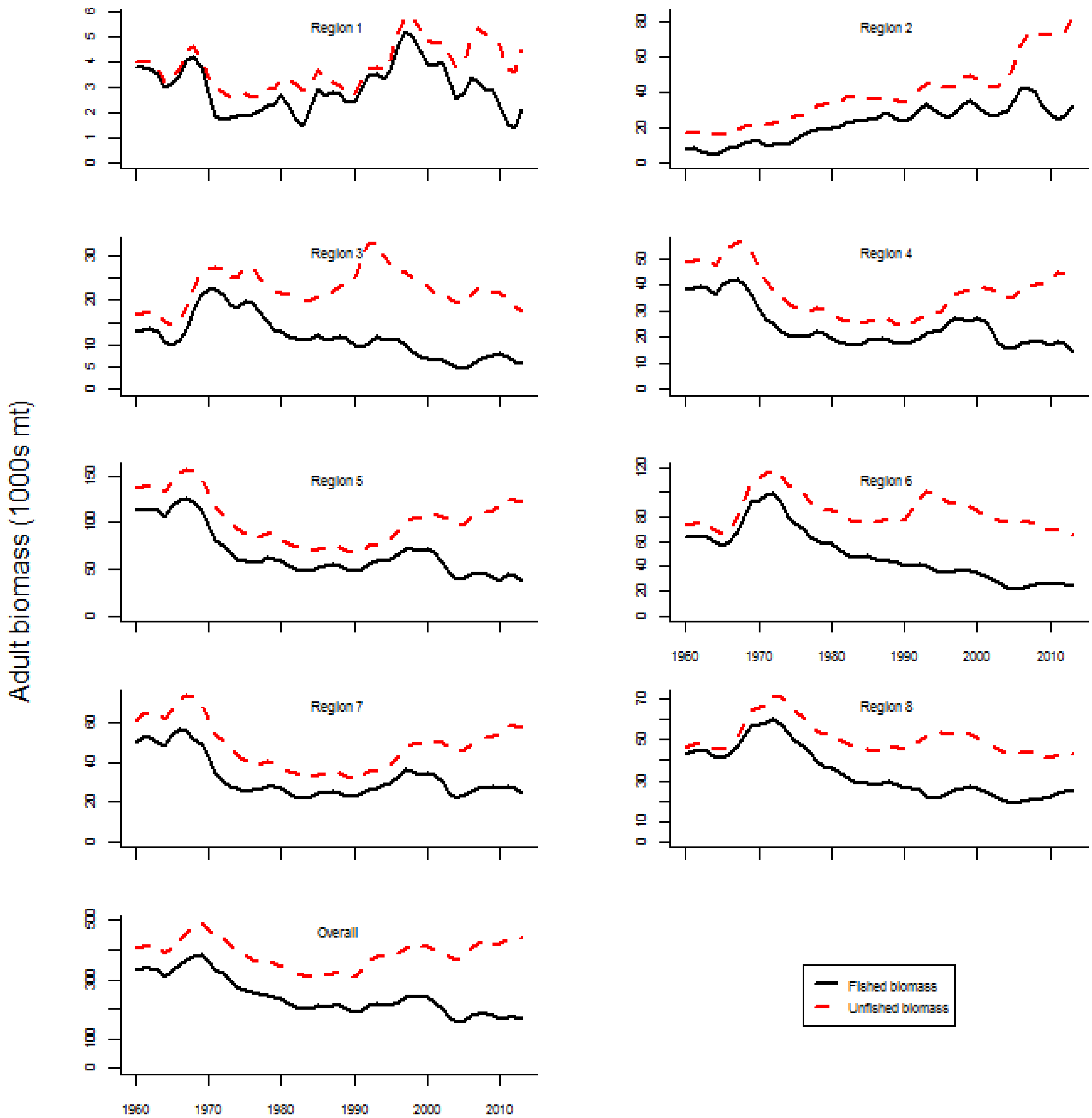


Figure 30: Comparison of the estimated spawning potential trajectories (lower solid black lines) with those trajectories that would have occurred in the absence of fishing (upper dashed red lines) for each region and overall for the reference case.

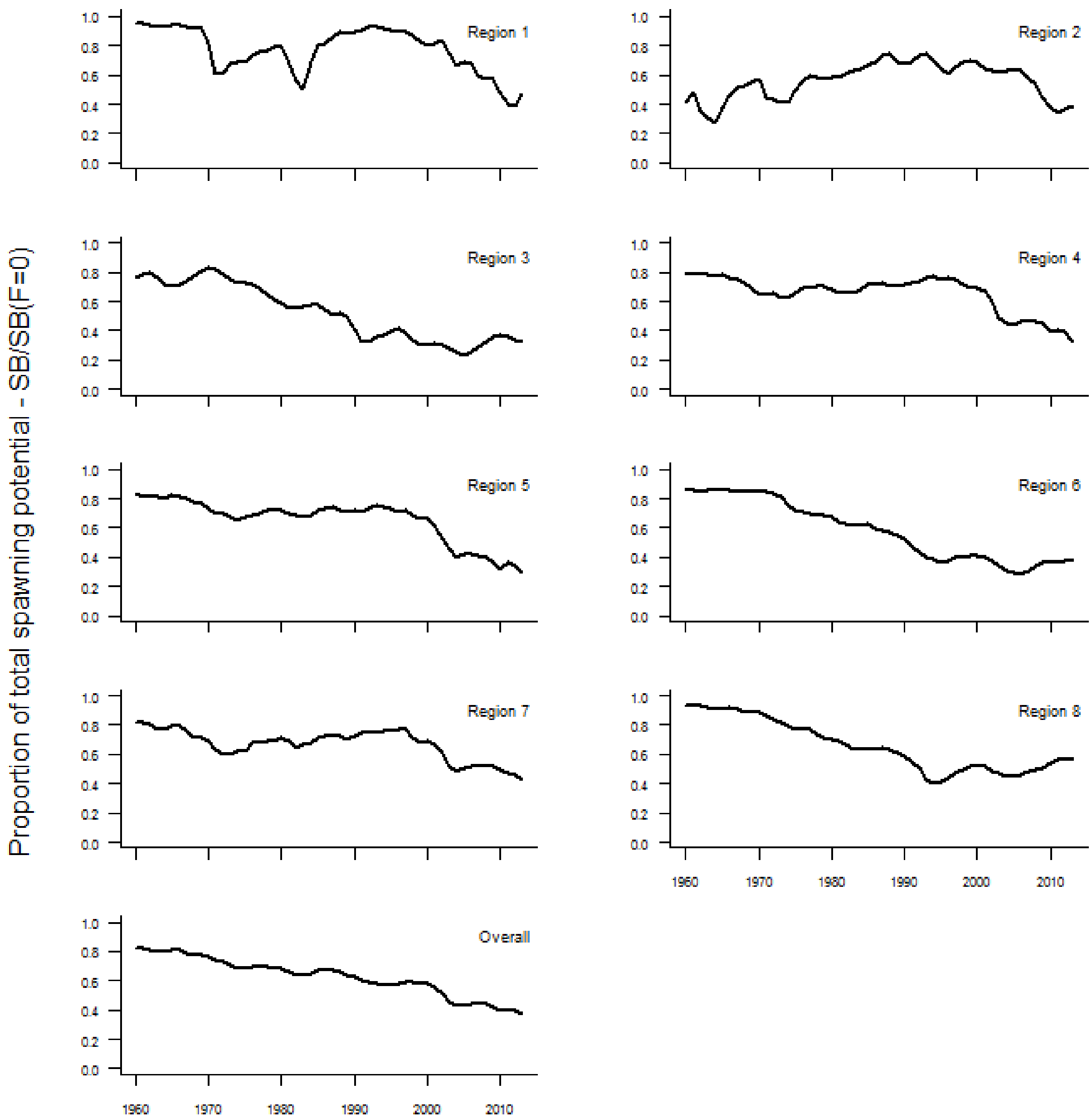


Figure 31: Ratio of exploited to unexploited spawning potential, $SB_{latest}/SB_{F=0}$, for each region and overall for the reference case.

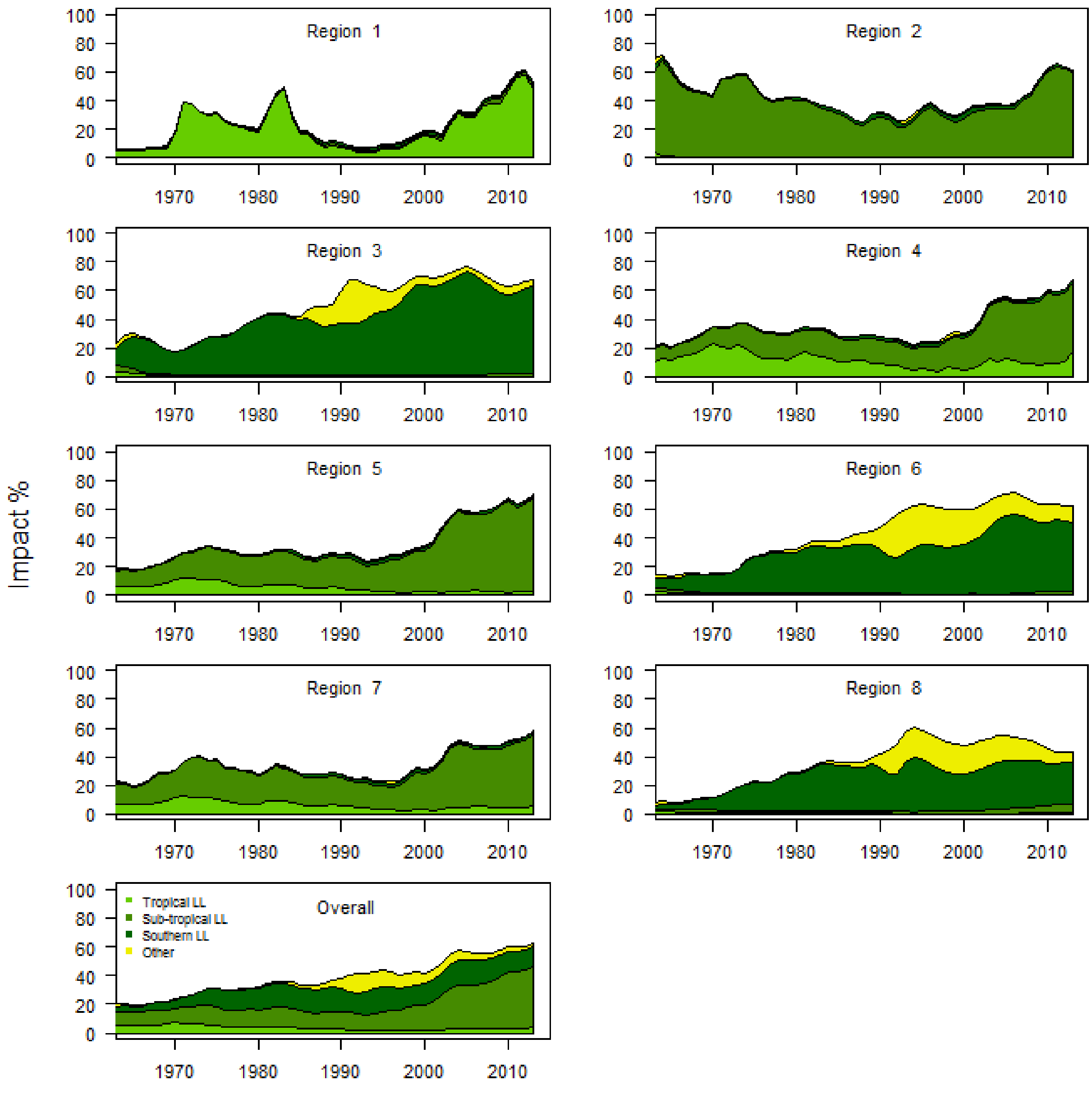


Figure 32: Estimates of reduction in spawning potential due to fishing (fishery impact = $1 - SB_{latest}/SB_{F=0}$) by region and overall attributed to various fishery groups for the reference case.

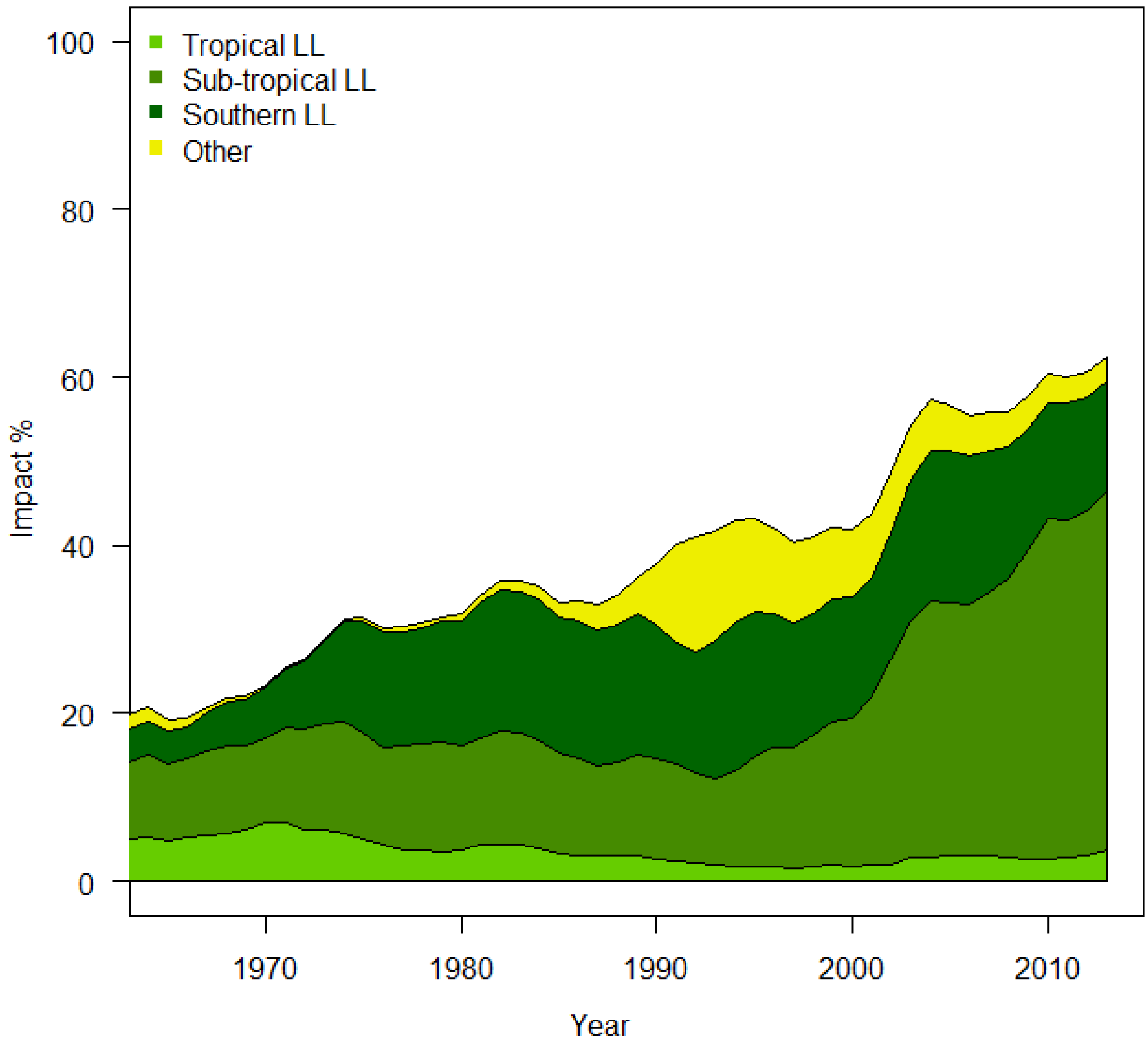


Figure 33: Estimates of reduction in spawning potential due to fishing (fishery impact = $1-SB_{latest}/SB_{F=0}$) overall attributed to various fishery groups for the reference case.

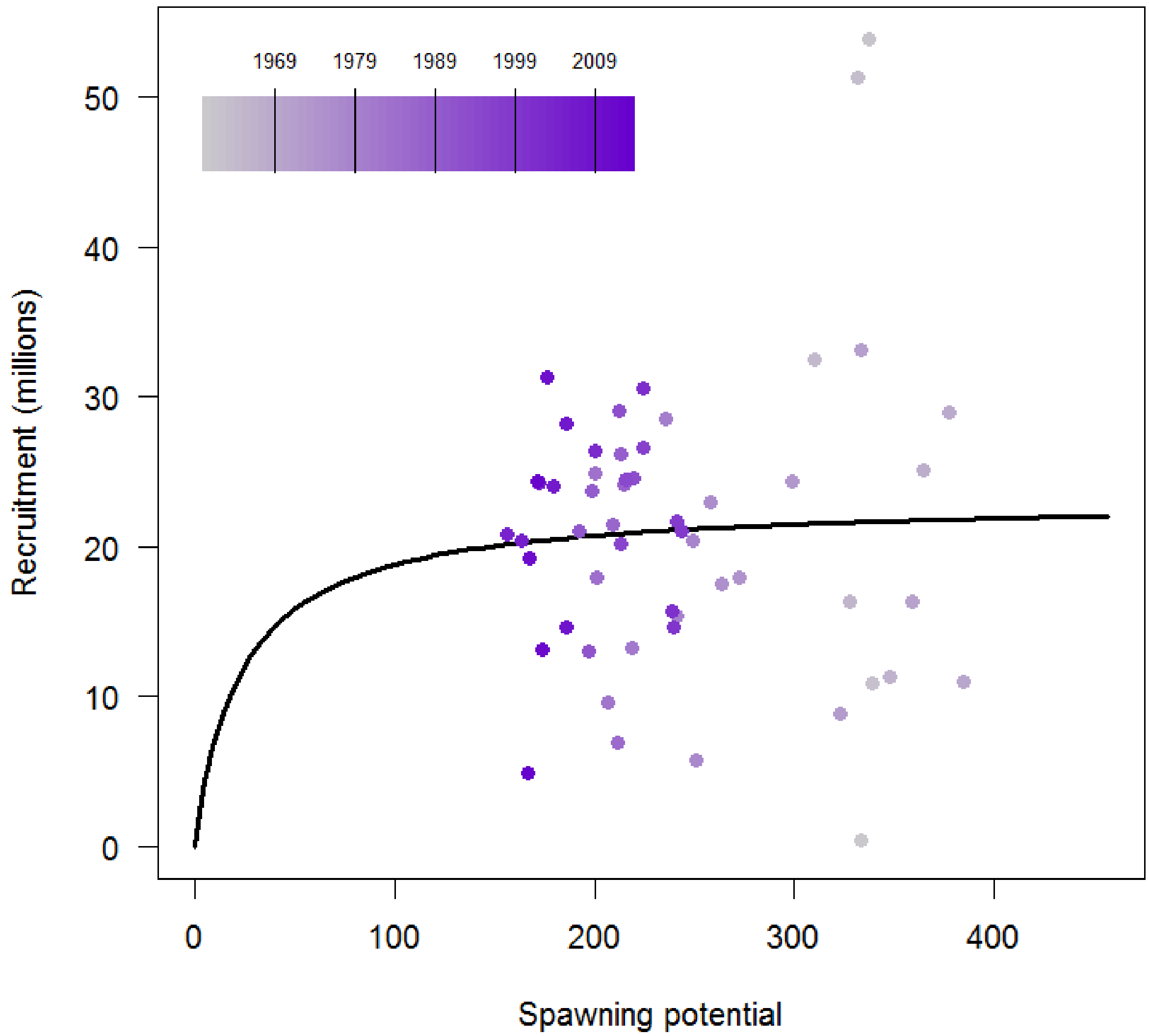


Figure 34: Estimated relationship between recruitment and spawning potential based on annual values for the reference case.The darkness of the circles changes from light to dark through time.

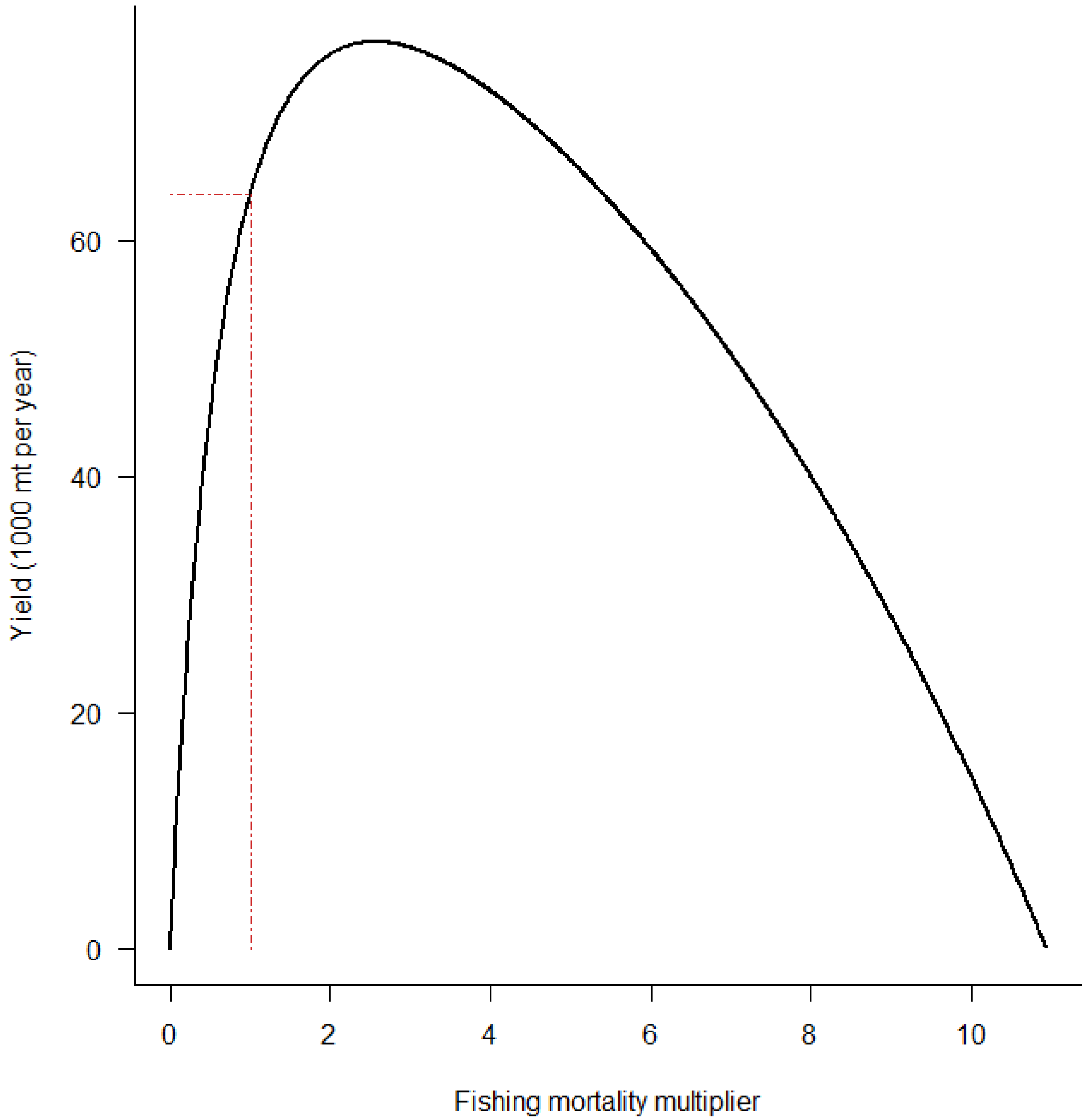


Figure 35: Estimated yield as a function of fishing mortality multiplier for the reference case. The red dashed line indicates the equilibrium yield at current fishing mortality.

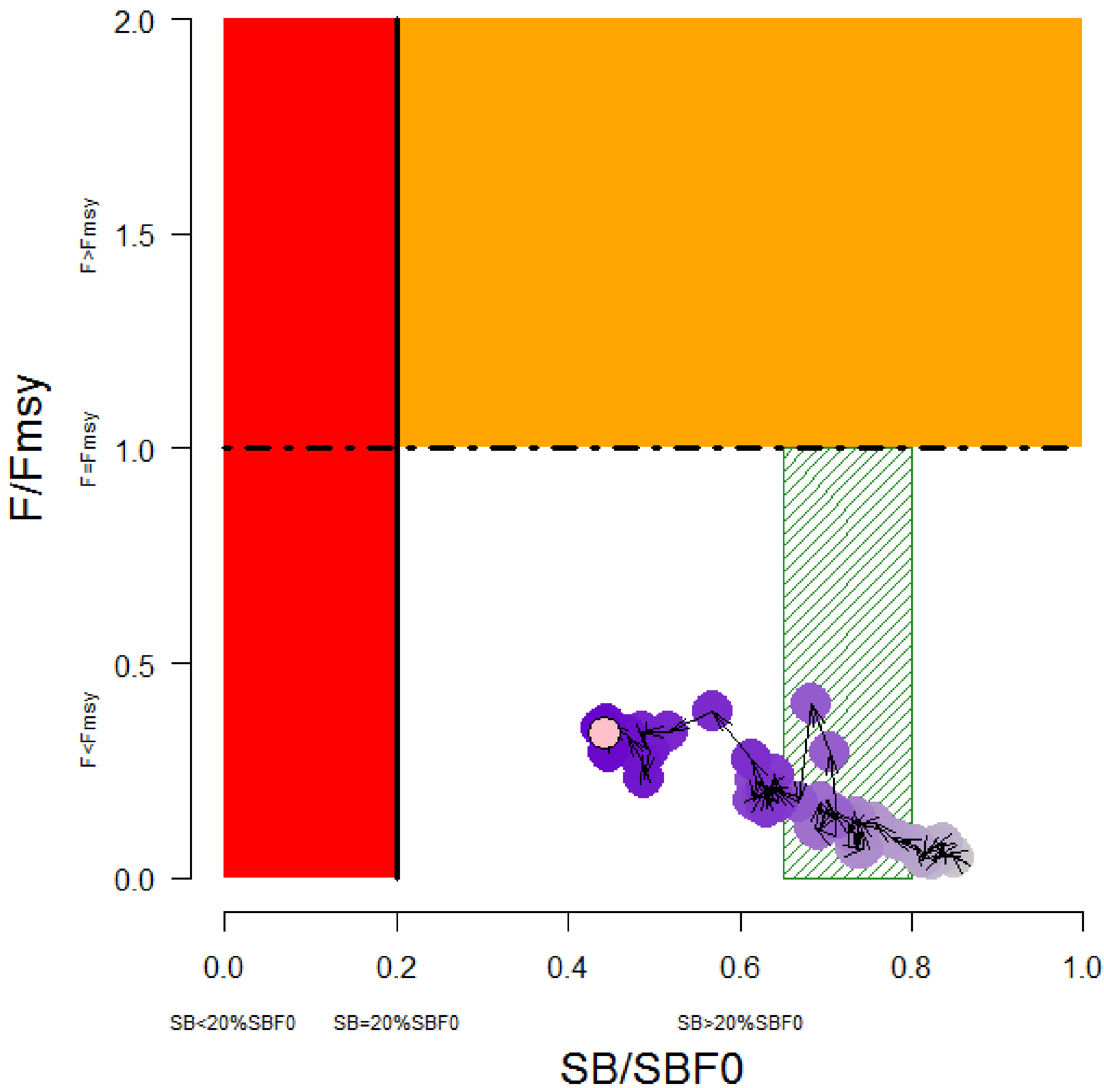


Figure 36: Majuro plot: representing stock status in terms of spawning potential depletion and fishing mortality. The red zone represents spawning potential levels lower than the agreed limit reference point which is marked with the solid black line. The orange region is for fishing mortality greater than F_{MSY} (F_{MSY} is marked with the black dashed line). The lightly shaded green rectangle covering 0.65–0.80$SB_{F=0}$ is the 'space' consistent with the candidate economic-based Target Reference Points provided in Pilling et al. (2015). The pink circle the latest period as defined in Table 2.

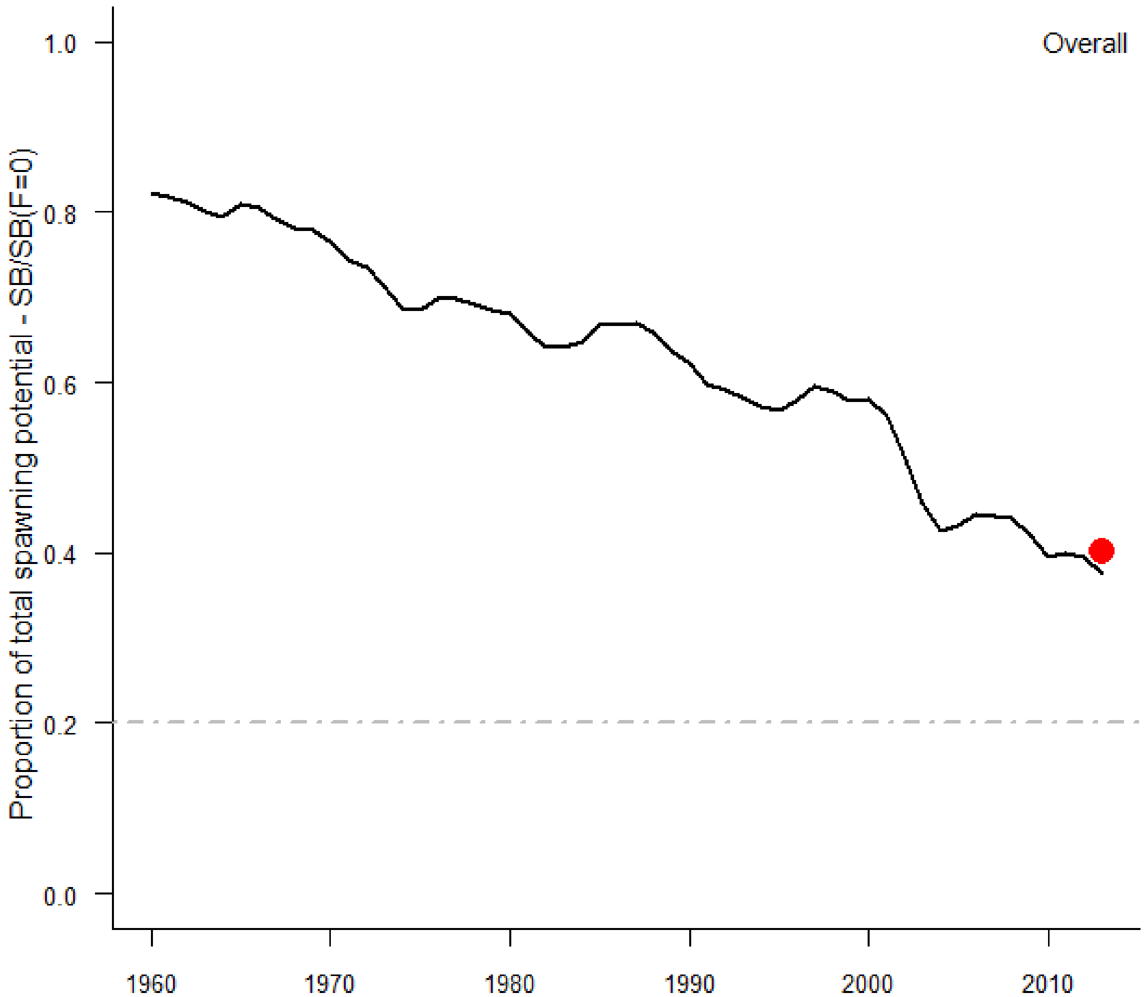


Figure 37: Ratio of exploited to unexploited spawning potential, $SB_{latest}/SB_{F=0}$, for the reference case. The current WCPFC limit reference point of $20\%SB_{F=0}$ is provided for reference as the grey dashed line and the red circle represents the level of spawning potential depletion based on the agreed method of calculating $SB_{F=0}$ over the last ten years of the model (excluding the last year).

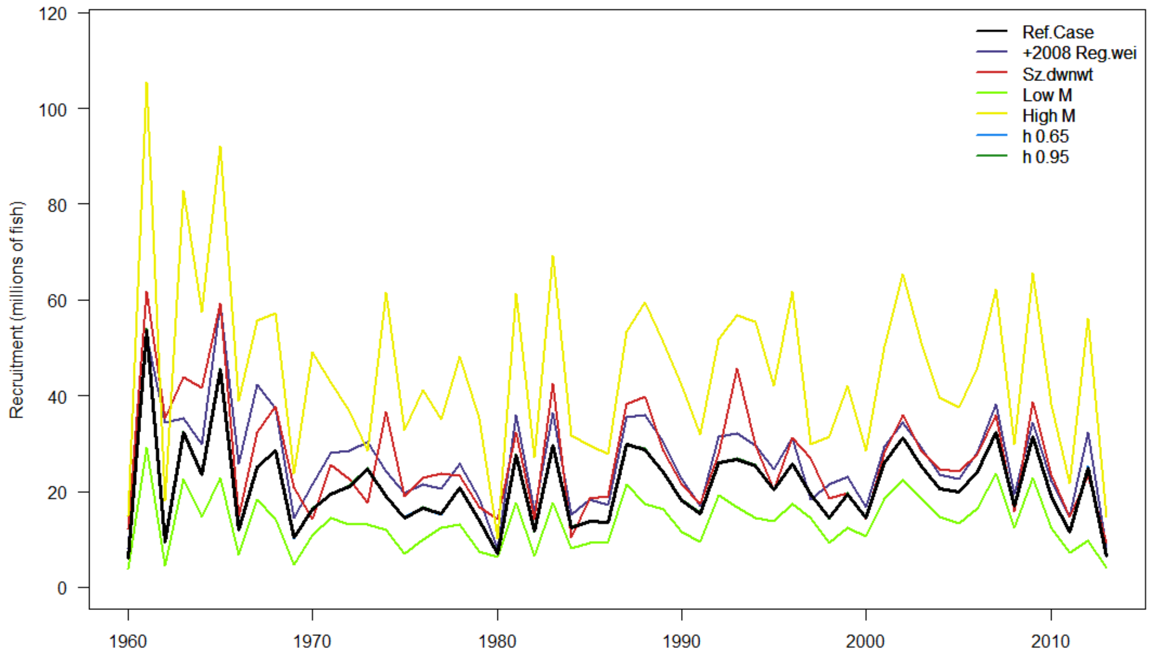


Figure 38: Estimated annual recruitment (millions) for the reference case and one-off sensitivity runs.

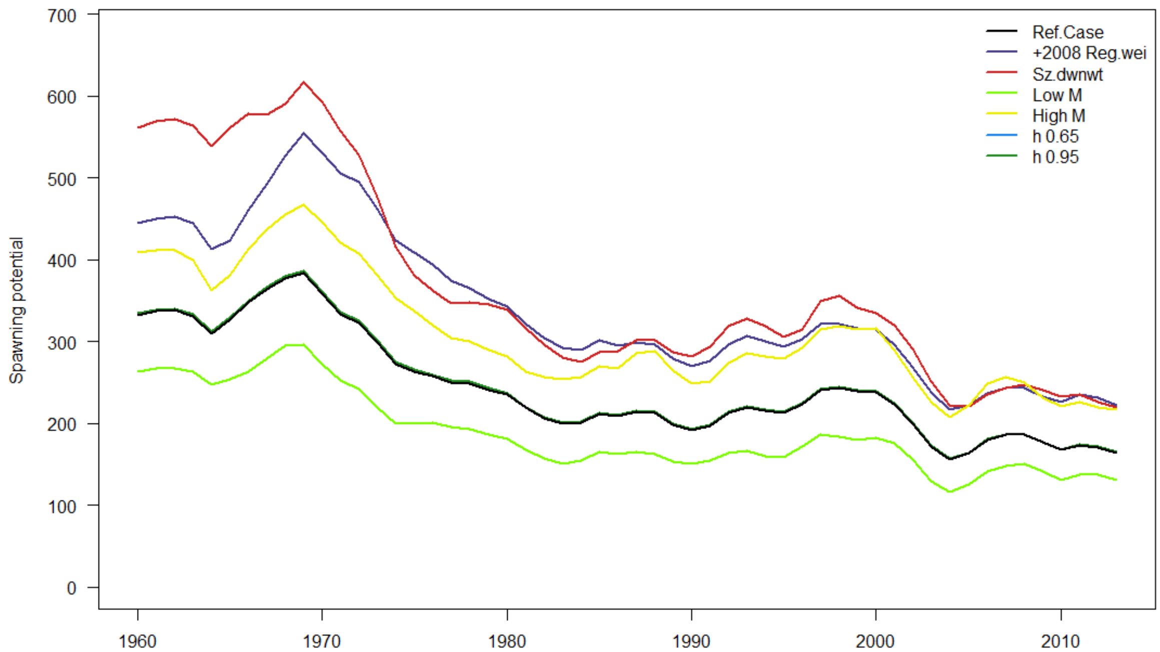


Figure 39: Estimated annual average spawning potential for the reference case and one-off sensitivity runs.

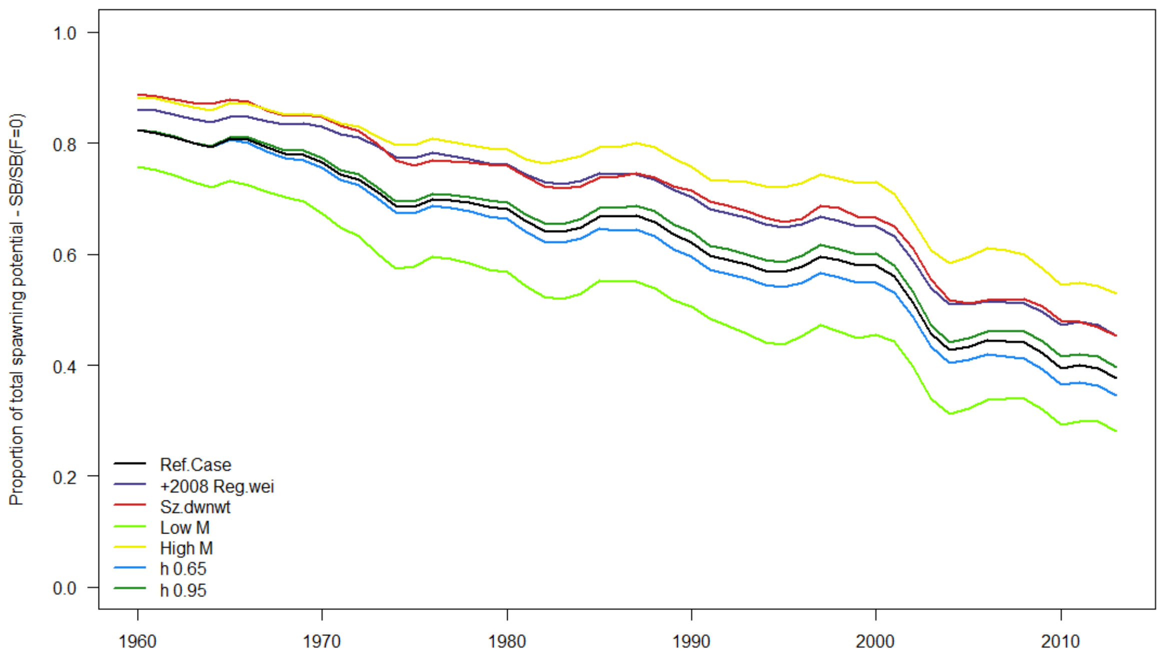
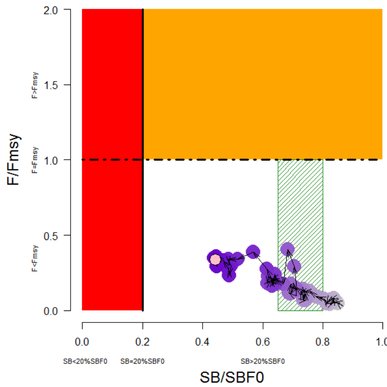
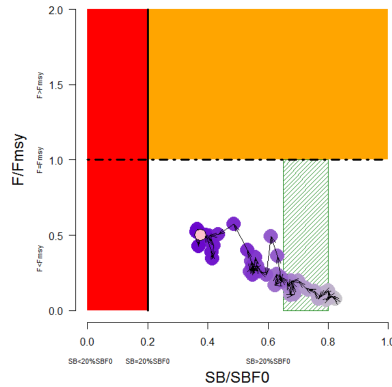


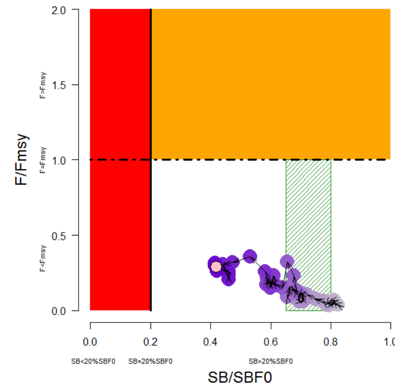
Figure 40: Ratio of exploited to unexploited spawning potential, $SB_{latest}/SB_{F=0}$, for the reference case and one-off sensitivity runs.



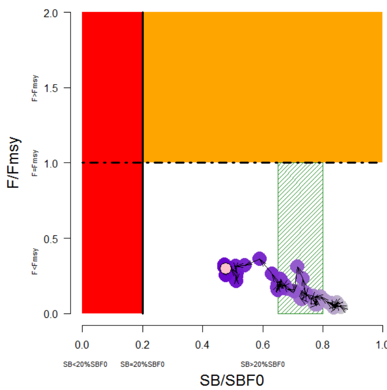
(a) Reference case



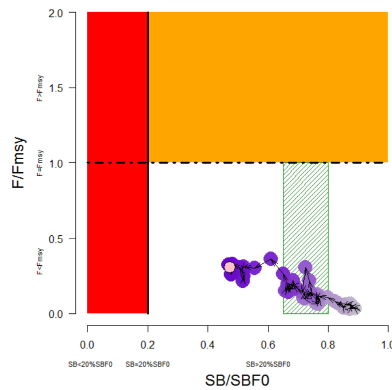
(b) Low steepness ($h = 0.65$)



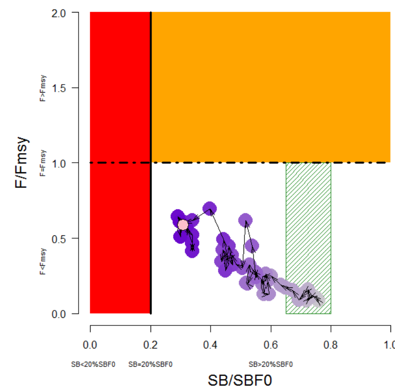
(c) High steepness ($h = 0.95$)



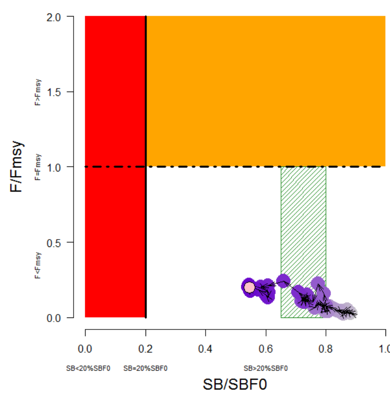
(d) Regional weights from 2008+



(e) Downweight size data

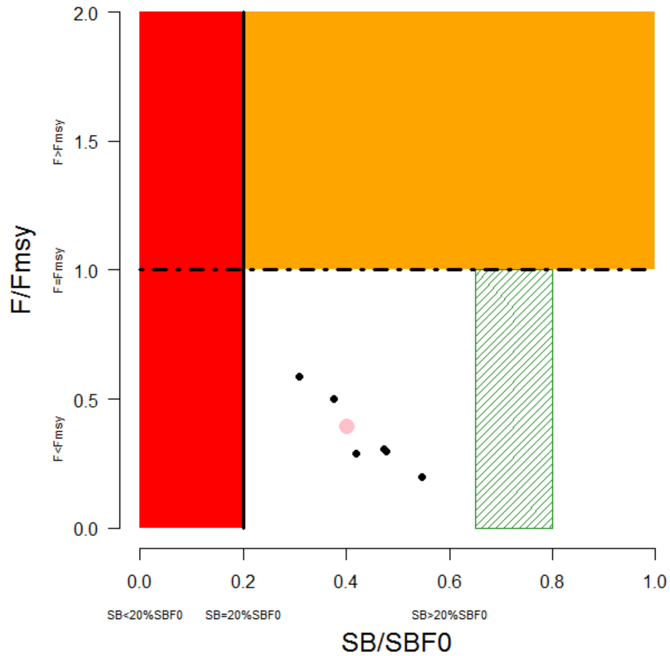


(f) Low M (0.25)

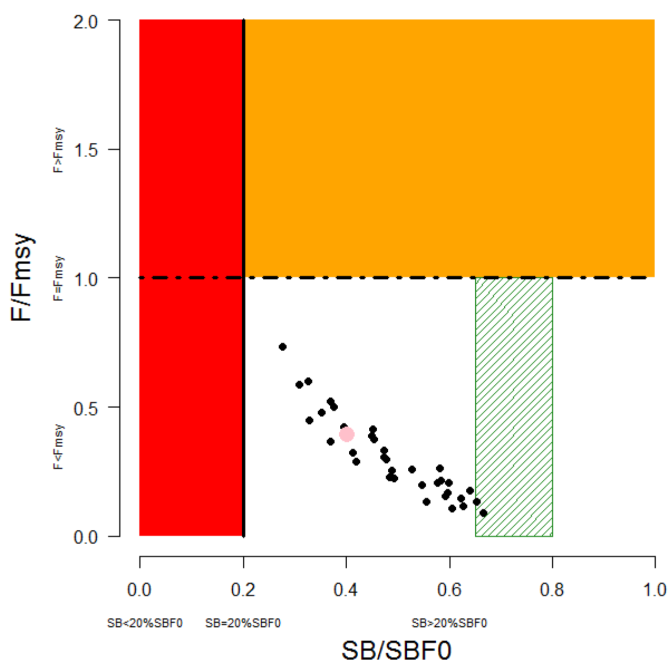


(g) High M (0.4)

Figure 41: Majuro plot for the reference case and one-off sensitivity runs

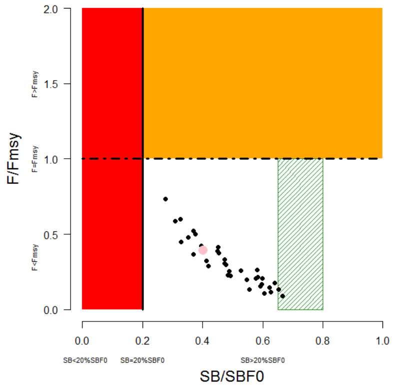


(a) One off model runs

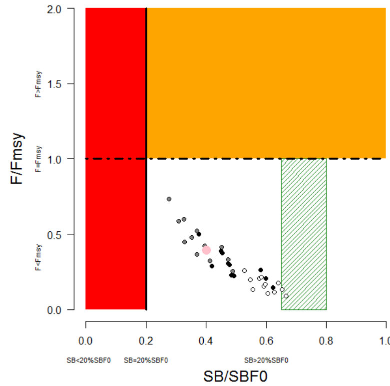


(b) Entire structural sensitivity grid

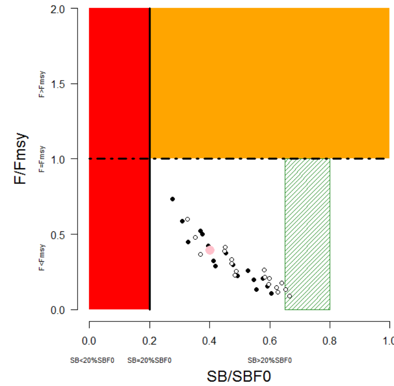
Figure 42: Majuro plot for one-off sensitivity runs and the entire structural sensitivity grid. The reference case model result is denoted by the pink circle.



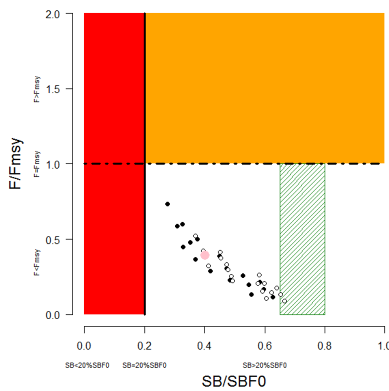
(a) Entire grid



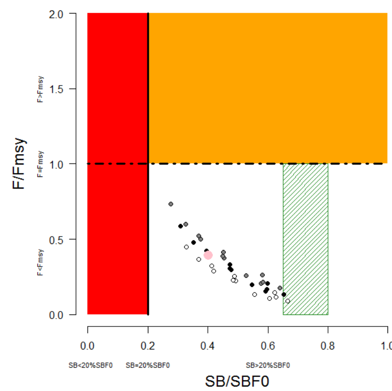
(b) Natural mortality



(c) LF data weighting



(d) Regional weights



(e) Steepness

Figure 43: Majuro plots for the entire structural sensitivity grid presenting the results for the different uncertainty axes. In each panel runs with the reference case assumption for that axes are presented with black circles. If there are only two options for an axis then the alterantive is presented as the white circle with the black outline. When there are three options (e.g., M and h) the lower value is presented with the grey circle and the higher value with the white circle with the black outline. The reference case model result is denoted by the pink circle.

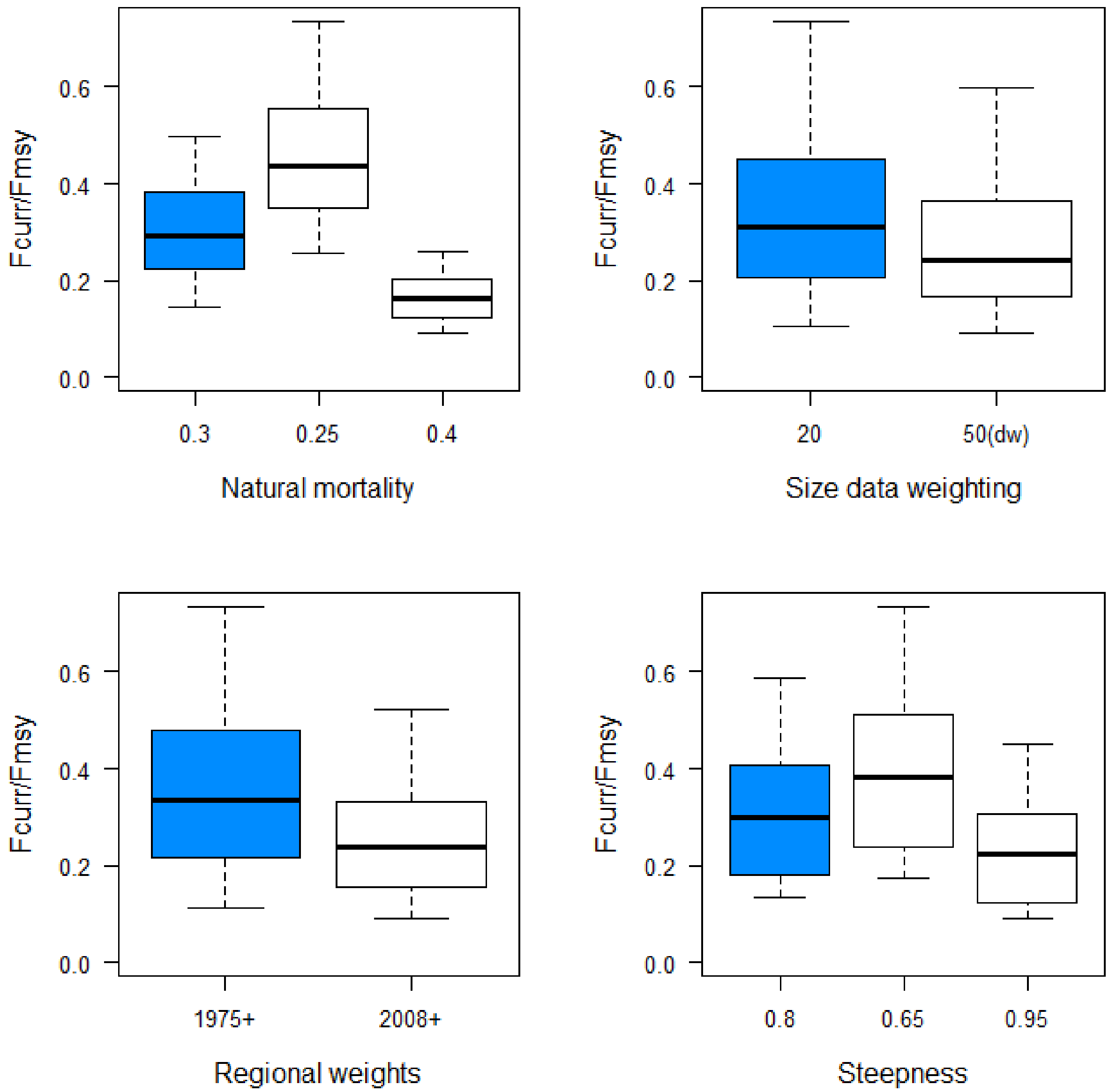


Figure 44: $F_{current}/F_{MSY}$ for different factors in the grid

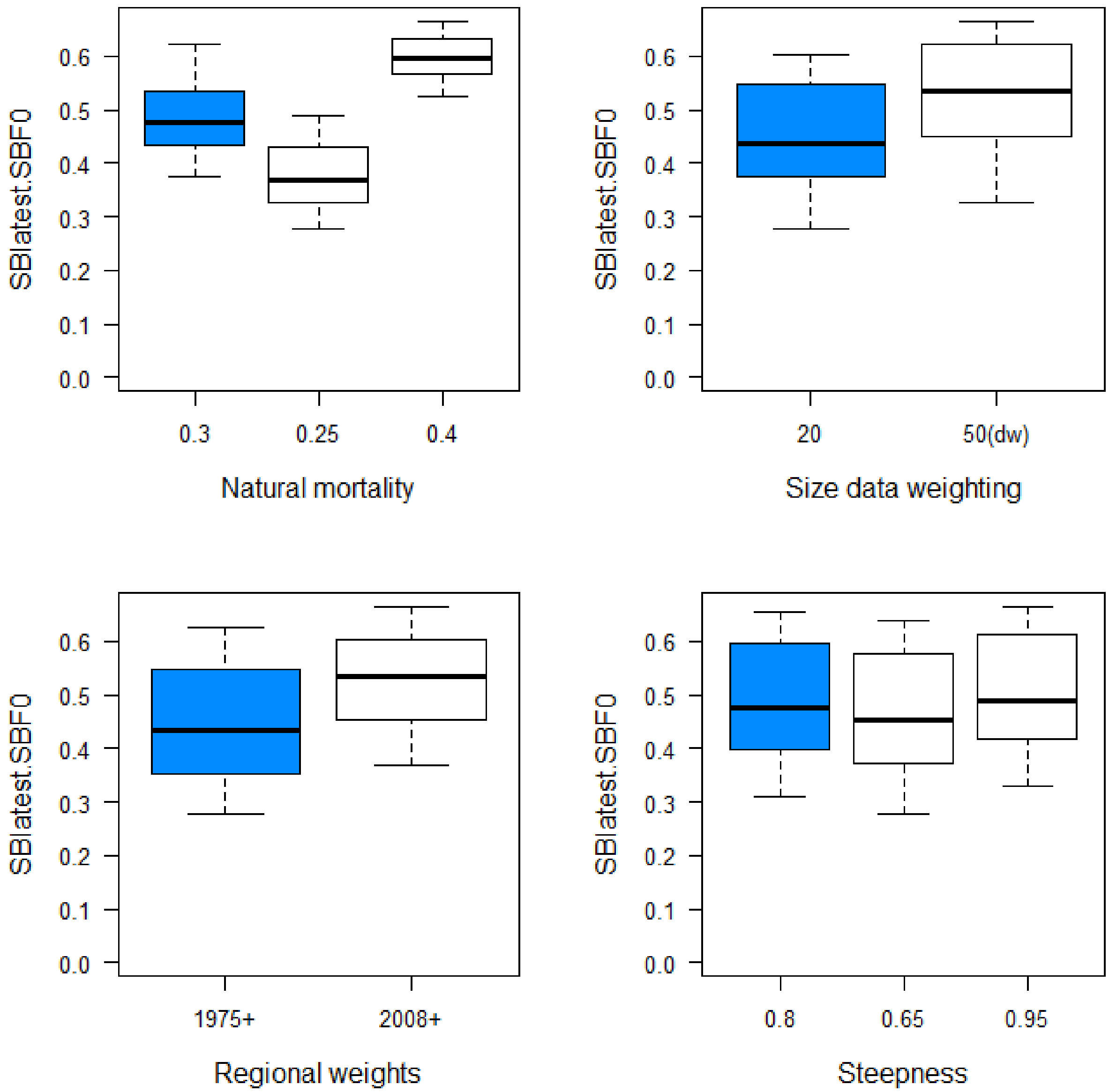


Figure 45: $SB_{latest}/SB_{F=0}$ for different factors in the grid

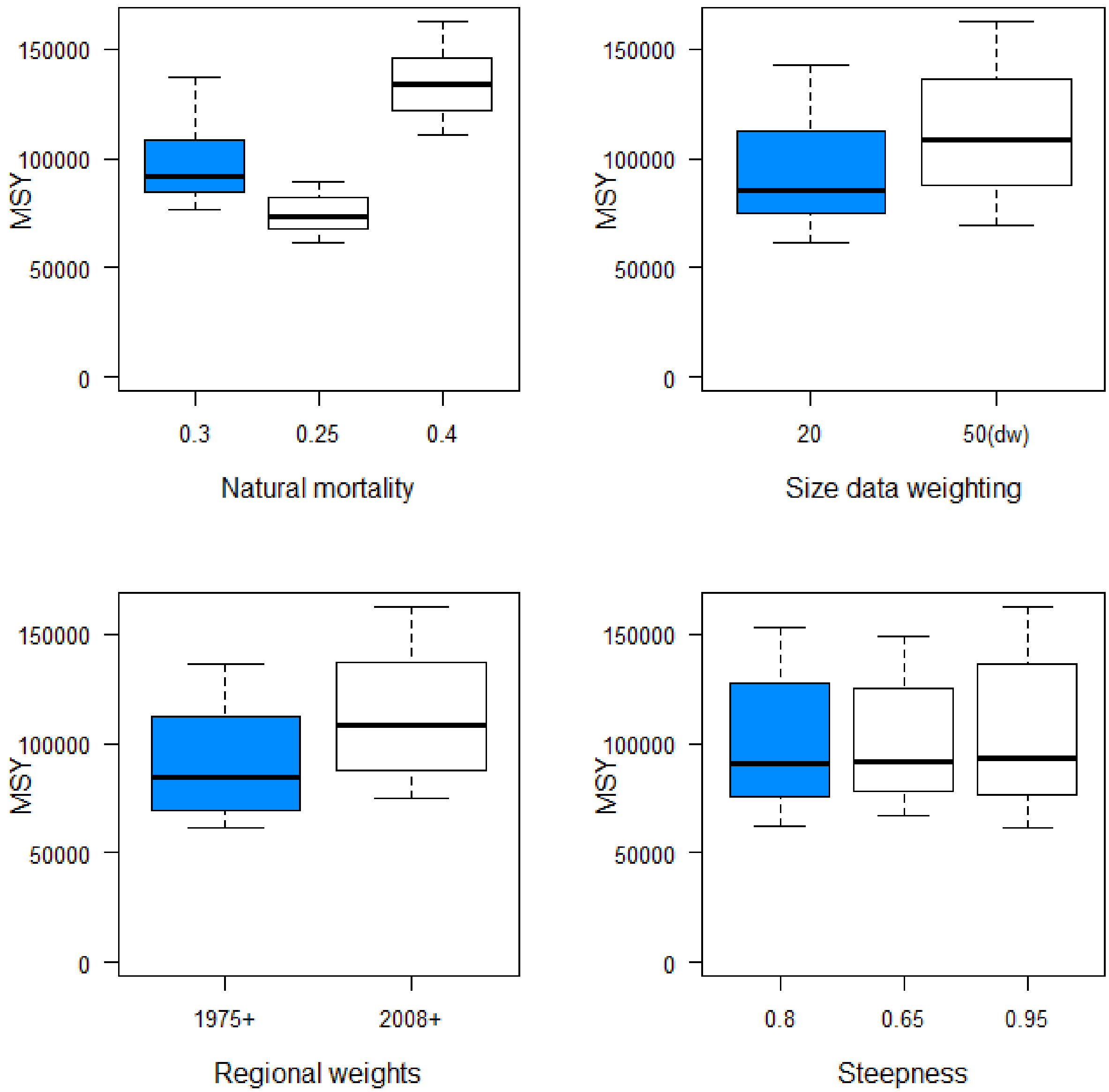


Figure 46: MSY for different factors in the grid

8 Annex

8.1 Likelihood profile

To evaluate the information available in the observation data component on the model's estimate of scale, a maximum likelihood profile was calculated over a global scaling parameter estimated by the model ("totpop"). The profile reflected the loss of fit over all the data, i.e. the overall objective function value, caused by changing the population scale from that of the maximum likelihood estimated value. The total population scaling parameter (totpop) of MULTIFAN-CL was used to explore the range of population scale because it directly determines the level of recruitment and, hence, absolute biomass. The profile entailed fitting a set of models over a range of fixed totpop values above and below the maximum likelihood estimate.

This analysis was undertaken with a model very similar to the reference case. Unlike the bigeye tuna assessment in 2014 (Harley et al., 2014), the analysis for albacore tuna showed no evidence that the reference case solution was a local minima. Higher values of the scaling parameter gave clearly worse fit to the data, as did lower values, but there was some evidence of a lack of convergence for some of the runs with lower totpop values. Future profiling exercises should examine the fit to individual likelihood components to better understand potential data conflict. This was done in the earlier stages of model development.

8.2 Retrospective analyses

8.2.1 Removal of recent years data

Retrospective analysis involves rerunning the model by consecutively removing successive years of data to estimate model bias (Cadrin and Vaughan, 1997; Cadigan and Farrell, 2005). Note, the retrospective analyses used a different, but very similar model to the reference case with terminal recruitment estimated.

A series of models were fitted starting with the full data-set (through 2013), followed by models with the retrospective removal of all input data for the years 2013, 2012 and 2011. The models are named below by the final year of data included (e.g., 2010–2012). A comparison of the recruitment, spawning potential, and depletion trajectories is shown in Figure 48.

Aside from the model using data only through 2012, all models gave very similar results for the model outputs examined. It is possible that the model using data through 2010 either did not fully converge, or converged to a local minima, but this was not examined further.

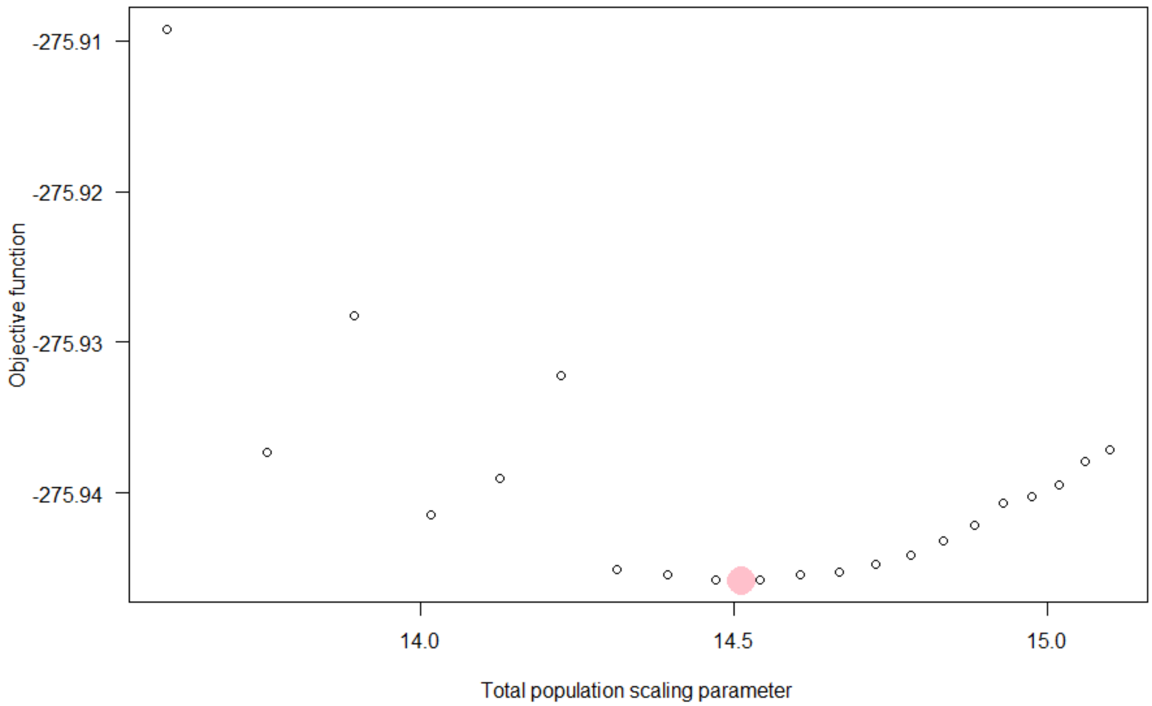
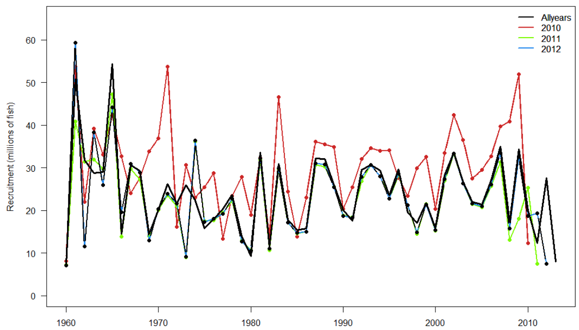


Figure 47: Profile of the marginal total negative log-likelihood in respect of the population scaling parameter. The large pink circle represents the estimated value from the model used in the profiling exercise.

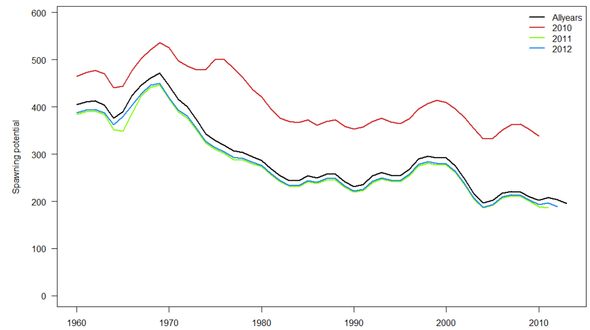
8.2.2 Comparison to previous assessments

The reference case model for the current (2015) assessment was compared retrospectively to those for the past two assessments done in 2012 and 2011. There were many changes to the 2015 assessment in terms of the addition of spatial structure, moving to a quarterly time step, and different values for natural mortality and maturity.

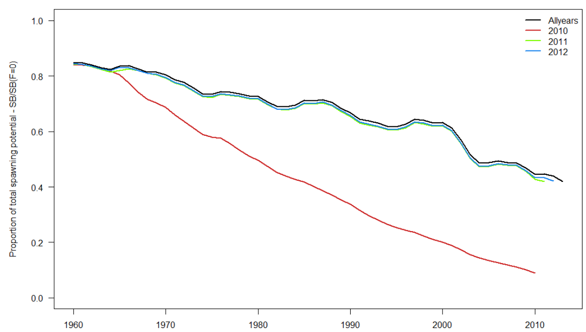
While there are clear differences in the absolute values for recruitment and spawning potential, the key management quantities are more similar. The current assessment gives results most similar to those from the 2011 assessment (Table 8 and Figure 50).



(a) Recruitment



(b) Spawning potential

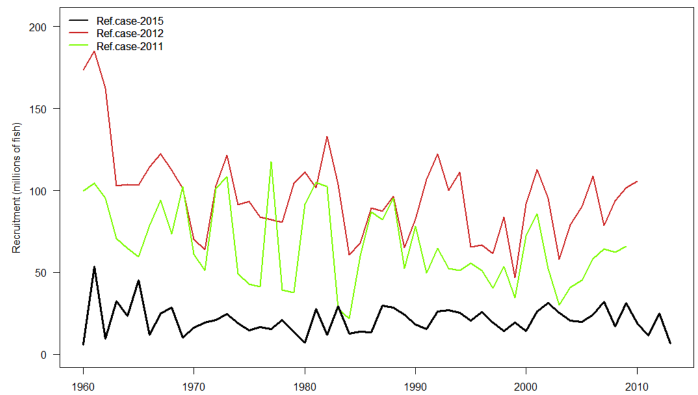


(c) Spawning potential depletion

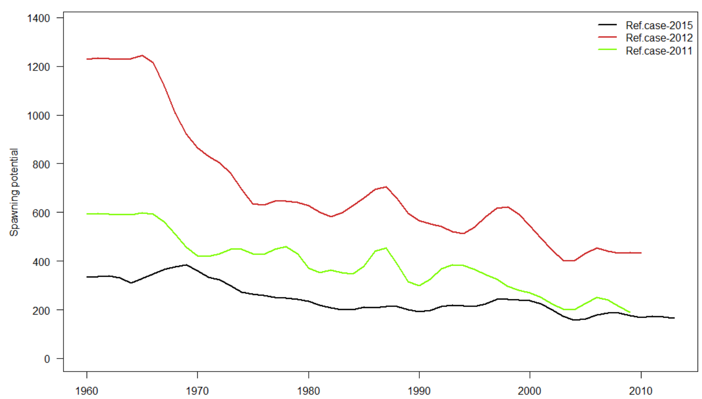
Figure 48: Recruitment estimates, (top-left) spawning potential (top-right), and depletion (bottom) from the penultimate version of the reference case, and for retrospective analyses for the successive removal of data from the end of the observation time series from 2012 to 2010. Model runs are denoted by the final year of data.

Table 8: Key management quantities for the reference case models used for the south Pacific albacore stock assessments in 2011, 2012, and the current assessment (2015).

	2011	2012	2015
MSY	$85,130$	$133,200$	$76,800$
$F_{current}/F_{MSY}$	0.259	0.143	0.392
$SB_{latest}/SB_{F=0}$	0.539	0.706	0.403

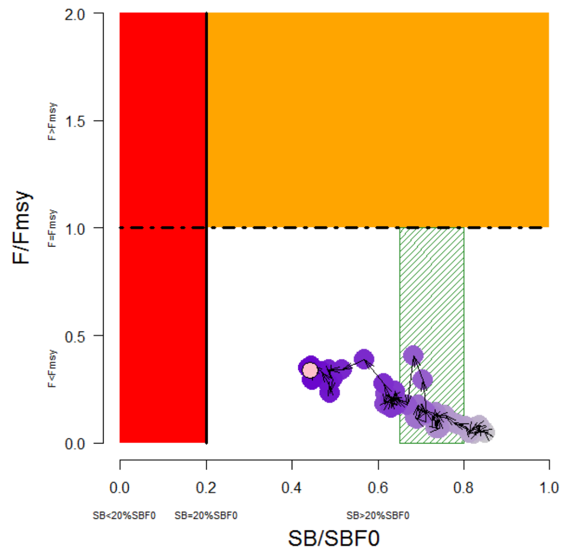


(a) Recruitment

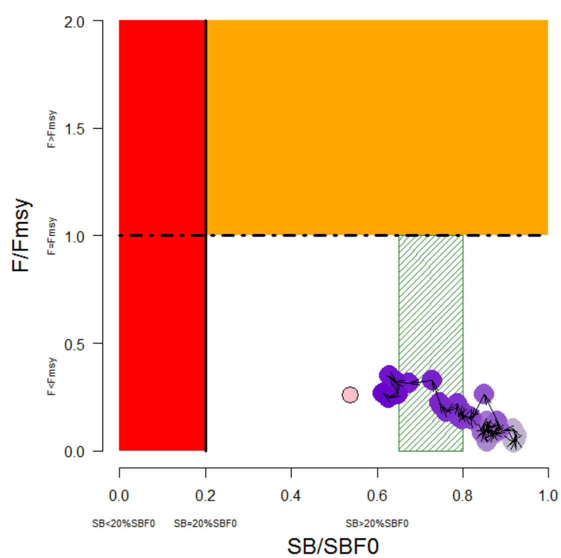


(b) Spawning potential

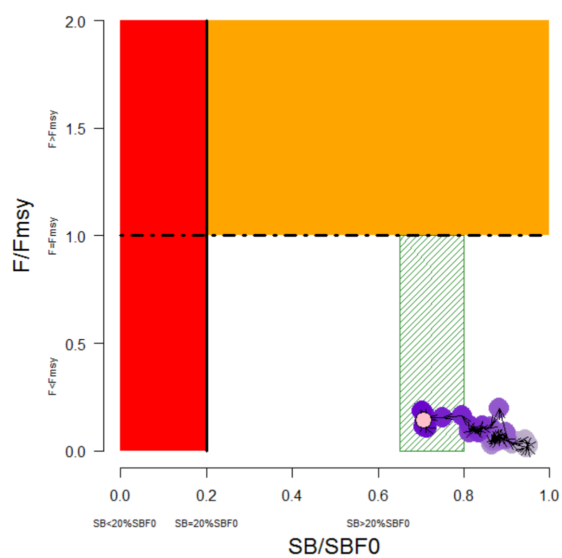
Figure 49: Annual recruitment (top) and spawning biomass (bottom) estimates from the reference case models for the south Pacific albacore tuna assessments from 2011, 2012 and the current assessment (2015).



(a) 2015 assessment



(b) 2011 assessment



(c) 2012 assessment

Figure 50: Comparison of the estimates of stock status in respect of spawning potential relative to $SB_{F=0}$ and $F_{current}/F_{MSY}$ for the 2011, 2012, and 2015 assessments. The pink circle the latest period ($SB_{latest}/SB_{F=0}$) in each assessment.