Western and Central Pacific Fisheries Commission

SCIENTIFIC COMMITTEE
THIRTEENTH REGULAR SESSION

Rarotonga, Cook Islands

9–17 August 2017

Stock assessment of bigeye tuna in the western and central Pacific Ocean

S. McKechnie[1], G. Pilling[1], J. Hampton[1]

[1]Oceanic Fisheries Programme, The Pacific Community

Details of changes made during the revision of the original report

Minor changes have been made to the stock assessment report of bigeye tuna that had previously been uploaded to the WCPFC scientific committee website. These relate to the addition of an extra level to the size data weighting axis in the structural uncertainty grid. Due to time constraints, only the two bounds of the size data weightings (divisors of 10 and 50) that were investigated in the one-off sensitivity model runs were included in the grid, due to the large computational overhead, limited computational resources and time constraints imposed on the assessments. Subsequent to the report deadline, we have completed the model runs for the extra level (divisor of 20; the level used in the diagnostic case model) and so modify the stock assessment report to incorporate summaries that include these extra runs. These, along with other minor corrections can briefly be summarised as:

- An extra 24 models were run as part of the structural uncertainty grid, with all extra models using a divisor of 20 for the size data weighting. These model runs had minimal effects on the summaries of the grid as this level of the size weighting axis generally falls between the more extreme divisors of 10 and 50, but are included for completeness sake. The new figures are shown in Appendix Section 11.4 and the figure numbers are set to be directly comparable to the original figures in the main text (e.g. Figures A39–A41 are equivalent to Figures 39–41). The new tables are also in Appendix Section 11.4 and the same numbering procedure is used (Tables A6–A10 are equivalent to Tables 6–10 in the main text).

- Two further tables are also included at the request of CCMs subsequent to the submission of the original assessment report. These concern subsets of the structural uncertainty grid related to the new and old growth models. The two new tables essentially combine the regional structure levels within the growth axis such that the first table (Table A11) summarises the grid over all models using the new growth (combining models with 2017 and 2014 regional structures) and the second (Table A12) is the same except that it summarises over all models using the old growth models. Note that these summarise the "new" grid that includes the size weighting level with a divisor of 20, and so each table is calculated over 36 models.

- In addition, all references to "spawning biomass" in the tables and figures have been corrected to "spawning potential". These occurred in Table 4, and Figures 13, 31–33 and 41–47.

- In the report previously uploaded to the WCPFC website, the description of the *New growth, fixed external otolith estimate* (Section 5.2.3) - the growth function used for all "new growth" models in the assessment, stated that "by fitting to the annual otolith dataset (supplemented with daily otoliths for young fish from CSIRO and earlier SPC readings; model *Annual-Ot-VB*)". This is the correct model as presented in McKechnie et al. (2017b), although the text has been corrected to read "by fitting to the annual otolith dataset (supplemented with daily otoliths for young fish from CSIRO; model *Annual-Ot-VB*)", as only the CSIRO daily otoliths were included in this model.

- Another request of a CCM subsequent to the submission of the original assessment report was to include figures displaying the recruitment for the old growth sensitivity model *L2-184* to support the statements in the main body of the report that highlighted the recent high recruitment event, even for the those models assuming the old growth function. These are now included in the Appendix Section 11.5. Total biomass and spawning potential for this model are also included to display the recent increase in the former and the absence of an increase in the latter, owing to the slight delay in maturity for this model compared to the new growth models (Figure A42).

Contents

1 Executive Summary

This paper describes the 2017 stock assessment of bigeye tuna (*Thunnus obesus*) in the western and central Pacific Ocean. A further three years data were available since the last stock assessment was conducted in 2014, and the model time period extends to the end of 2015. New developments to the stock assessment include addressing the recommendations of the 2014 stock assessment report (Harley et al., 2014), incorporation of new data such as a recent ageing of otoliths to estimate age-at-length for WCPO fish, investigation of an alternative regional structure, exploration of uncertainties in the assessment model, particularly in response to the inclusion of additional years of data, and improving diagnostic weaknesses of previous assessments.

This assessment is supported by the analysis of recently collected biological data (Farley et al., 2017; McKechnie et al., 2017b), catch-per-unit-effort data for longline fisheries (Tremblay-Boyer and Pilling, 2017a), tagging data (McKechnie et al., 2017b) and the data summaries for fisheries definitions used in the stock assessment (McKechnie et al., 2017b).

Changes made in the progression from the 2014 reference case to 2017 diagnostic case models include:

– Updating all data up to the end of 2015.

– Utilising standardised CPUE indices calculated from the recently collated operational longline CPUE dataset.

– Investigating an alternative spatial structure with the boundaries between the tropical and northern temperate regions shifted from 20° N to 10° N.

– Investigating the use of a new growth curve based on the recently processed otoliths of Farley et al. (2017), which suggest a much lower asymptotic size for old fish.

– Implementation of new features developed in MFCL, including an annual stock recruitment relationship.

In addition to the diagnostic case model, we report the results of one-off sensitivity models to explore the relative impacts of key data and model assumptions for the diagnostic case model on the stock assessment results and conclusions. We also undertook a structural uncertainty analysis (model grid) for consideration in developing management advice where all possible combinations of the most important axes of uncertainty from the one-off models were included. In comparison to previous assessments, little emphasis is placed on the diagnostic case model. Instead it is recommended that management advice is formulated from the results of the structural uncertainty grid.

Across the range of models run in this assessment, the most important factors with

respect to estimates of stock status were the choice of the new (lower asymptotic size) versus old (higher asymptotic size) growth curves. The former estimated considerably more optimistic results than the latter, and this was also the case when compared to the results of the 2014 assessment. The second key axis explored in the structural uncertainty grid was whether the 2014 or 2017 regional structures were assumed. Again, the latter estimated significantly more optimistic stock status (though the effect of this assumption was less than for growth). The models assuming the 2017 regions essentially assign more of the stock to the less exploited temperate regions from the highly exploited equatorial regions where fishing depletion is estimated to be higher.

Based on these results, the main conclusions of the current assessment are more difficult to construct than in previous bigeye assessments. The Scientific Committee will have to assess the plausibility of the different models in the structural uncertainty grid, particularly four groups of models resulting from different combinations of new and old growth/maturity, and the 2017 and 2014 regional structure. To this end, we summarise the general conclusions of this assessment as follows:

1. All models that assume the new growth function estimate significantly more optimistic stock status than the 2014 assessment, with the stock above $20\%SB_{F=0}$ in all cases.

2. All models with the new growth estimate a significant recent recruitment event that has increased spawning potential in the last several years, and it is expected that for the old growth models these recruits will soon progress into the spawning potential and increase stock status, at least in the short-term.

3. Of the four sets of models in the structural uncertainty grid (the combinations of old/new growth and 2017/2014 regions), only the old growth/2014 regions models estimate spawning potential to be below $20\%SB_{F=0}$ for all models in the set. These models estimate $SB_{latest}/SB_{F=0}$ to be between 0.08 and 0.17 which is slightly more pessimistic than the structural uncertainty grid of the 2014 assessment (between 0.1 and 0.2).

4. A substantial decline in bigeye abundance was estimated by all models in the assessment and recent estimates of depletion with respect to estimates earlier in the assessment period, and with respect to estimates in the absence of fishing, are significant and appear to be ongoing, at least on a multi-year scale.

5. The significance of the recent high recruitment events and the progression of these fish to the spawning potential component of the stock are encouraging, although whether this is a result of management measures for the fishery or beneficial environmental conditions is currently unclear. It is noteworthy, however, that recent positive recruitment events have also been estimated for skipjack (McKechnie et al., 2016a) and yellowfin (Tremblay-Boyer et al., 2017) in the WCPO, and bigeye in the EPO

([Aires-da Silva et al., 2017](#)), which may give weight to the favourable environmental conditions hypothesis. Whether these trends are maintained in coming years will help tease these factors apart and will likely provide more certainty about the future trajectories of the stock.

2   Introduction

This paper presents the 2017 stock assessment of bigeye tuna (*Thunnus obesus*; BET) in the western and central Pacific Ocean (WCPO; west of 150° W). Since 1999, the assessment has been conducted regularly and the most recent assessments are documented in Langley et al. (2008); Harley et al. (2009, 2010); Davies et al. (2011) and Harley et al., 2014. Most recently, a sensitivity analysis was conducted that compared assessment outcomes for WCPO and Pacific Ocean model configurations (McKechnie et al., 2015a).

The independent review of the 2011 bigeye tuna assessment (Ianelli et al., 2012) made several recommendations that have improved the subsequent assessments of this stock (Harley et al., 2014). The current assessment continues the implementation of that review's recommendations, facilitated by the ongoing development of the statistical stock assessment software, known as MULTIFAN-CL[2] (MFCL; Fournier et al., 1998; Hampton and Fournier, 2001; Kleiber et al., 2017), that is routinely used by the Pacific Community (SPC). Further developments to the stock assessment have been undertaken to address the recommendations of the 2014 stock assessment report (Harley et al., 2014), to address the recommendations of the 2017 pre-assessment workshop (PAW; Pilling and Brouwer, 2017), to explore uncertainties in the assessment model, particularly in response to the inclusion of additional years of data, to improve diagnostic weaknesses in previous assessments, and to include additional information and data from recent analyses of bigeye biology in the WCPO (Farley et al., 2017; McKechnie et al., 2017b).

The objectives of this assessment are to estimate population parameters, such as time series of recruitment, biomass, biomass depletion and fishing mortality, which indicate the stock status and impacts of fishing. We summarize the stock status in terms of reference points adopted by the Western and Central Pacific Fisheries Commission (WCPFC). The methodology used for the assessment is based on the general approach of integrated modelling (Fournier and Archibald, 1982), which is carried out using MFCL, and 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 (penalties).

This assessment report should not be seen as a standalone document and should be read in conjunction with several supporting papers, specifically the analyses of longline CPUE data (Tremblay-Boyer and Pilling, 2017a), background analyses of biological parameters (Farley et al., 2017; McKechnie et al., 2017b) and definition of the regional and fisheries structures for the updated assessment (McKechnie et al.,

---

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

2017b). Finally, many of these issues were discussed in detail, and recommendations to the assessment approach made, at the PAW held in Noumea over 24–27 April, 2017 (Pilling and Brouwer, 2017).

3 Background

3.1 Stock structure

Bigeye are distributed throughout the tropical and sub-tropical waters of the Pacific Ocean. Analysis of mtDNA and DNA microsatellites in nearly 800 bigeye tuna failed to reveal significant evidence of widespread population subdivision in the Pacific Ocean (Grewe and Hampton, 1998). While these results are not conclusive regarding the rate of mixing of bigeye tuna throughout the Pacific, they are broadly consistent with the results of tagging experiments on this stock conducted by SPC and the Inter-American Tropical Tuna Commission (IATTC). Before 2008, most bigeye tuna tagging in the Pacific occurred in the far eastern Pacific (east of about 120° W) and in the western Pacific (west of about 180°). While some of these tagged bigeye were recaptured at distances from release of up to 4,000 nautical miles over periods of one to several years (Figure 1), a large majority of tag returns were recaptured much closer to their release points (Schaefer and Fuller, 2002; Hampton and Williams, 2005).

Since 2008, bigeye tuna tagging by the Pacific Tuna Tagging Programme (PTTP) has been focused in the equatorial central Pacific, between 180° and 140° W. Returns of both conventional and electronic tags from this programme have been suggestive of more extensive longitudinal, particularly west to east, displacements (Schaefer et al., 2015). Few movements between tropical tag release sites and temperate zones have been detected, and this information formed the basis for the PAW recommendation to investigate a new regional structure for the current stock assessment with region boundaries shifted from 20° N to 10° N (McKechnie et al., 2017b). The consequences of this change were examined in detail during this assessment.

It is hypothesised that while bigeye tuna in the far eastern and western Pacific may have relatively little exchange, those in the central part of the Pacific between about 180° and 120° W may mix more readily over distances of 1,000–3,000 nautical miles (Schaefer et al., 2015). In any event, it is clear that there is extensive movement of bigeye across the nominal WCPO/EPO boundary of 150° W (Figure 2). The consequences of this directional movement for estimates of stock status in the WCPO were assessed in the 2015 Pacific-wide bigeye stock assessment (McKechnie et al., 2015a) and the results suggested the current approach of modelling the WCPO was

robust to this behaviour.

3.2 Biological characteristics

Bigeye tuna are relatively fast growing, and have a maximum fork length (FL) of about 200 cm (Aires-da-Silva et al., 2015). Recent integrated analyses of tag recapture and age-at-length data for EPO bigeye (Aires-da-Silva et al., 2015) suggest mean lengths-at-age are larger than those estimated internally in bigeye stock assessments in the WCPO prior to 2014, which were based on fitting to size frequency data over the full range of age-classes. Similarly, McKechnie et al. (2015a) detected significantly smaller lengths-at-age for bigeye in the WCPO than Aires-da-Silva et al. (2015) using similar methods of combining otolith and tag increment data in integrated growth models. This significant divergence in growth between the WCPO and EPO is supported by the new otolith dataset presented by Farley et al. (2017) and will be discussed in detail throughout this paper.

Available data for the WCPO indicate that bigeye tuna begin to be reproductively active from about 80 cm FL, and nearly all individuals >120 cm FL are reproductively mature (Farley et al., 2017). There is some evidence for regional variation in maturity-at-length in the WCPO (Farley et al., 2017), and bigeye tuna appear to be reaching maturity at larger sizes, but at similar ages, in the EPO (Schaefer et al., 2005). As with other tunas, the sex ratio of bigeye tuna changes at around the age/size of reproductive maturity to favour males at larger size (see Figure 7 in McKechnie et al., 2017b). This information is used to define spawning potential (rather than spawning biomass) based on mature female biomass in stock assessments and the recently analysed biological samples have been included in updated estimates of these parameters (McKechnie et al., 2017b).

The natural mortality rate of bigeye tuna is likely to vary with size, with high rates for the youngest age-classes and the lower rates of around 0.5 yr^{-1} for bigeye >40 cm FL (Hampton, 2000). Tag recapture data indicate that significant numbers of bigeye reach at least eight years of age (Hampton and Williams, 2005). The longest period at liberty for a recaptured bigeye tuna tagged in the WCPO is approximately 14 years, for a fish 1-2 years old at release (SPC unpublished data). Natural mortality of female bigeye is hypothesised to increase at around the age of reproductive maturity, due to the physiological stresses of spawning, which, as noted above, results in male-biased sex ratios at larger sizes. This feature of the population dynamics has been incorporated into the fixed natural-mortality-at-age schedules used in recent, and current, bigeye stock assessments. Furthermore, the current assessment also includes estimation of natural mortality-at-age in a sensitivity analysis, using recent

MFCL developments of functional forms for the estimation of natural mortality (e.g. Lorenzen, 1996).

3.3 Fisheries

Bigeye tuna are an important component of tuna fisheries throughout the Pacific Ocean and are taken by both surface gears, mostly as juveniles, and longline gear, as valuable adult fish. They are a principal target species in tropical waters of both the large, distant-water longline fleets of Japan, Korea, China and Chinese Taipei and the smaller, fresh sashimi longline fleets based in several Pacific Island countries and Hawaii. Prices paid for both frozen and fresh product on the Japanese sashimi market are the highest of all the tropical tunas. The longline catch in the WCPFC area had a "delivered" value in 2015 of approximately US$514 million (Williams and Terawasi, 2016).

Bigeye caught by purse seine vessels are taken almost exclusively from sets on natural and artificial floating objects (FADs). Estimation of the bigeye (and to a lesser extent yellowfin) tuna catch from associated sets has been the focus of considerable research over several years. Section 4.5.2 and references within provide details of this work. The purse seine fishery mostly targets skipjack, and to a lesser extent yellowfin, though significant incidental catches of small bigeye occur. This fishery expanded rapidly from the early 1980's and the estimated annual catch for this gear has recently been around 50,000-75,000 mt.

A small purse seine fishery also operates in the coastal waters off Japan with an annual bigeye catch of approximately 1,000 mt. A slightly higher level of bigeye catch is taken by the coastal Japanese pole-and-line fishery.

In recent years, collaborative work between SPC, WCPFC, CSIRO (primarily in Indonesia), and fisheries agencies in Indonesia, Philippines, and Vietnam have yielded improved catch statistics for their fleets. In some instances data are available at the individual fisheries level (e.g., longline or large-fish handline), but often statistics are aggregated across a variety of gears that typically catch small bigeye tuna, e.g., ring-net, handline, and troll. Data for these fisheries have been included in the assessment, as described below.

4 Data compilation

4.1 General notes

Data used in the stock assessment of bigeye using MFCL consist of catch, effort and length-frequency data for the fisheries defined in the analysis, and tag-recapture data. Conditional age-at-length data may also be used directly as data in the assessment model; however, for reasons discussed later, we opted to estimate growth externally and use those estimates as fixed parameters in the main assessment runs. Improvements in these data inputs are ongoing and more detailed summaries of the analyses and methods of producing the necessary input files are given by Farley et al. (2017) (developments to biological data), Tremblay-Boyer and Pilling (2017a) (CPUE standardisations) and McKechnie et al. (2017b) (tagging, size composition, CPUE standardisations and biological parameter analyses). In addition, a more detailed description of the fisheries definitions, their data summaries and changes in fisheries structures since the 2014 stock assessment, are provided in McKechnie et al. (2017b). The full details of these analyses are not repeated here, rather, a brief overview of the key features is provided and readers are directed to the relevant papers referenced throughout this section.

4.2 Spatial stratification

The geographical area considered in the 2014 stock assessment (Harley et al., 2014) corresponded to the WCPO from 50° N to 40° S, and from oceanic waters adjacent to the east Asian coast (110° E between 20° N and 10° S; 120° E north of 20° N) to 150° W. This regional structure comprises nine regions (Figure 3), with two regions north of 20° N (Regions 1 and 2), and four equatorial regions between 10° S to 20° N. The western equatorial region covers the area from 110° E to 140° E (Region 7) and was established in 2014 to reduce the impact of uncertainty in the catch time-series from Indonesia, Philippines, and Vietnam. The eastern equatorial region (region 4) is defined from 170° E to 150° W. The central equatorial regions (regions 3 and 8) together comprise the area between 140° E and 170° E. Region 8 is designed to approximate the archipelagic waters of PNG and Solomon Islands, where considerable tagging effort has occurred and analyses show more persistent residence compared to the wider western equatorial region.

The southern regions extend from 10° S to 40° S, and from 140° E to 150° W, with the boundary between the eastern (Region 5) and western (Region 6) regions established at 170° E. Within region 5 a small region (Region 9) is present and was established in 2014 to better model the tagging data from the Coral Sea (Harley et al., 2014). The

eastern boundary of the assessment is 150° W, and as such, excludes the WCPFC Convention area component that overlaps with the IATTC area.

The initial stepwise progression of the 2017 stock assessment utilised these region definitions. However, the northern boundaries between regions 1 and 3, and between regions 2 and 4 were subsequently shifted from 20° N to 10° N for many of the 2017 models (Figure 4). The rationale for this modification and the consequences for fisheries definitions are detailed in McKechnie et al. (2017b). In brief summary, this was undertaken based on recent analyses of tagging data from the Central Pacific (Schaefer et al., 2015) which indicated very limited movement between the equatorial zone and more temperate waters, and also more closely reflects the distribution of purse seine fishing in the Pacific, which is largely restricted to the zone between 10° S and 10° N.

4.3 Temporal stratification

The time period covered by the assessment is 1952–2015 which includes all significant post-war tuna fishing in the WCPO. Within this period, data were compiled into quarters (1; Jan–Mar, 2; Apr–Jun, 3; Jul–Sep, 4; Oct–Dec). As agreed at SC12, the assessment does not include data from the most recent calendar year. This is because these data are only finalized very late and often subject to significant revision post-SC, in particular the longline data on which this assessment greatly depends.

4.4 Definition of fisheries

MFCL requires "fisheries" to be defined that consist of relatively homogeneous fishing units. Ideally, the defined fisheries will have selectivity and catchability characteristics that do not vary greatly over time and space, although in the case of catchability some allowance can be made for time series variation. For most pelagic fisheries assessments, fisheries are defined according to gear type, fishing method and region, and sometimes also by vessel flag or fleet. The fisheries definitions (32 fisheries) used in the diagnostic case model are presented in Table 1 and consist of longline, purse seine, pole and line and various miscellaneous small-fish fisheries in Indonesia and the Philippines.

The fisheries for models utilising the 2014 regional structure remained the same as those used in the 2014 assessment, with a total of 33 fisheries defined. The consequences of modifying the region boundaries for fisheries definitions in the 2017 regional structures are presented in detail by McKechnie et al. (2017b) and led to a reduction to 32 fisheries (there is no longer a US LL fishery in region 4). A graphical

summary of the availability of data for each fishery used in the diagnostic case model is provided in Figure 5.

Equatorial purse seine fishing activity was aggregated over all nationalities, but stratified by region and set type, in order to sufficiently capture the variability in fishing operations. Set types were grouped into associated (log, FAD, whale, dolphin, and unknown set types) and unassociated (school) sets. Further fisheries were defined for pole-and-line fisheries and miscellaneous fisheries (gillnets, ringnets, handlines etc.) in the western equatorial area. At least one longline fishery was defined in each region, although in regions 3 and 7 longline fishing was separated into distant water and offshore components to account for the apparent differences in fishing practices (including captured fish sizes) for these fleets in these regions.

4.5   Catch and effort data

4.5.1   General characteristics

Catch and effort data were compiled according to the fisheries defined above. Catches by the longline fisheries were expressed in numbers of fish, and catches for all other fisheries expressed in weight. This is consistent with the form in which the catch data are recorded for these fisheries.

Total annual catches by major gear categories for the WCPO are shown in Figure 6 and a regional breakdown is provided in Figure 7. The spatial distribution of catches over the past ten years is provided in Figure 8. Discarded catches are estimated to be minor and were not included in the analysis. Catches in the northern region are highly seasonal and the annual catch has been relatively stable over much of the assessment period. Most of the catch occurs in the tropical regions (3, 4, 7, and 8). As noted above, only data through 2015 were used in the current assessment to overcome the delays and data issues that commonly occur in the most recent year.

Within the model, effort for each fishery was normalised to an average of 1 to assist numerical stability. Some longline fisheries were grouped to share common catchability parameters in the various analyses. For such grouped fisheries, the normalisation occurred over the group rather than for the individual fisheries so as to preserve the relative levels of effort between the fisheries. For some fisheries no effort is used - this is typically in cases where effort data is either considered unreliable or the fishery aggregates different "other" fishing gears such that effort units are not compatible.

A number of significant trends in the fisheries have occurred over the model period, specifically:

16

– The steady increases in bigeye catch by longline vessels over most of the assessment period in the equatorial and southern regions.

– The relatively stable catches of bigeye in the northern temperate region by longline vessels, and to a much lesser degree, Japanese pole and line and purse seine vessels in region 1.

– The development of the equatorial purse-seine fisheries from the mid-1970s and the widespread use of FADs since the mid-1990s, allowing an expansion of the purse-seine fishery, and corresponding increases in catch of bigeye, particularly in equatorial regions 3, 4 and 8.

– Large changes in the purse seine fleet composition and increasing size and efficiency of the fleet.

– The steady increase in catch for the domestic fisheries of Indonesia and the Philippines since 1970.

– The apparent stabilisation of catches of bigeye for most gears after the mid 2000's, due to limits under the WCPFC conservation management measures.

4.5.2 Purse seine

Previous assessments have considered two sets of purse-seine input catch data, but the problems surrounding logbook reporting of bigeye catches and grab-sample bias have been clearly demonstrated and only a single set of purse seine catch estimates have been included in the 2014 and current assessments. Details of the analyses, including the independent review and response are provided in Lawson (2013); Cordue (2013); Powers (2013), and McArdle (2013). The procedure now used routinely for estimating purse seine catch by species is documented in Hampton and Williams (2017) as Method 3.

As in previous assessments, effort data units for purse seine fisheries are defined as days fishing and/or searching, and are allocated to set-types based on the proportion of total sets attributed to a specified set type (associated or unassociated sets) in logbook data. Recently it has been discovered that some fleets have changed their reporting practices (SPC-OFP, 2013), such that far fewer searching days are reported and these are instead reported as non-fishing transit days. This practice essentially represents effort creep and we have not yet specifically corrected recent data to ensure consistency of reporting. Therefore the impact of this is not known, but it will be minimized by the practice of estimating frequent time-based changes in catchability.

4.5.3 Longline fisheries

The LL CPUE indices for the main longline fisheries in each region are one of the most important inputs to the assessment as they provide information on trends in abundance over time for each region, and differences in abundance among the regions. For the current assessment, several sources of standardized CPUE series were used in various stages of the assessments.

2014 assessment approach and update

The first set of indices was constructed to closely match those used in the 2014 reference case model and were used early in the stepwise progression from the 2014 reference case model to the 2017 diagnostic case model. These were originally derived from Japanese operational-level longline data using generalized linear models (GLMs) with a delta-lognormal approach (Hoyle and Okamoto, 2011; McKechnie et al., 2014b). These were only available for the old (2011 assessment regional structure) regions 1-6 and through to 2009 and for some areas the indices for 2009 were very uncertain. In order to have time series that went through until 2012 (for the 2014 assessment) it was necessary to use Japanese aggregate catch and effort data and then "splice" these together. The procedures for this are described in McKechnie et al. (2014b). These indices were used in regions 1 and 2 of the 2014 assessment, and this process was updated for the 2017 assessment (for the early models in the stepwise development) by extending the indices to the end of 2015.

For the remaining regions of the model, the indices used in 2014 were calculated using SPC-held operational LL logsheet data for a range of LL fleets, including both distant-water and domestic-based vessels. The procedures used to standardise these data are again outlined in McKechnie et al. (2014b). The development of these methods was undertaken in response to the independent review of the bigeye assessment which suggested the spatial contraction of the Japanese fleet (creating uncertainty in the indices based on it) and targeting changes as the two major issues to address for standardisation of longline CPUE data (Ianelli et al., 2012). The indices developed for the 2014 assessment attempted to address these issues in two ways: 1) by using data across multiple fleets in order to minimize the spatial/temporal gaps in longline CPUE coverage; and 2) using operational data which allows consideration of vessel effects, clustering based on species composition of the catch, and other operational details to better account for targeting changes. For these regions, the indices used in the 2017 assessment were recalculated using identical methods on datasets updated to the end of 2015.

In 2014, the indices in regions 4 and 6 were calculated from datasets supplemented with Chinese Taipei logsheet data that were not held in SPC databases. This allowed

estimation of these indices in earlier years before more widespread data become available in the SPC-held operational database. To update these indices for the stepwise progression of model development it was necessary to again "splice" the 2014 indices with indices calculated on the SPC-held data updated up to the end of 2015.

Full longline operational dataset indices

Subsequent to the 2014 assessment, an extensive operational LL dataset has been collated from the SPC-held data together with the data held by all the important distant water fishing nations (see McKechnie et al., 2015b for details of this development). For the 2017 assessment it was possible to calculate standardised indices for all regions for most of the assessment time-period. Several types of models were fitted to these data and these are outlined in several ancillary papers (McKechnie et al., 2017b; Tremblay-Boyer and Pilling, 2017a,b). The indices used in the diagnostic case model were estimated with models that were extremely similar to the 2014 standardisation models and produced indices that were in general congruent with the previous estimates, although in some regions the indices now cover the early time-period which was often missing in the 2014 assessment. These indices were estimated for both the 2014 and 2017 regional structures and these are utilised in the stepwise model progression process (Figure 9).

Alternative indices for these fisheries were utilised in sensitivity analyses that, 1) attempted to overcome the missing vessel identifiers for Japanese vessels in the early time-period (Tremblay-Boyer and Pilling, 2017b), and 2) utilise newly developed geostatistical techniques to address some of the issues of the changing spatial coverage of fishing effort in the dataset over time (Tremblay-Boyer and Pilling, 2017a).

Coefficients of variation (CVs) for region-specific standardised effort were averaged to 0.2 over the period 1980–1990. GLMs fitted to the operational dataset indices produce CVs which were comparable across all regions so it was decided that a similar mean CV be used for all regions. By implementing these methods, MFCL is able to account for the time-varying nature of the CVs such that the CPUE data is given more weighting in time-steps with more precise estimates of abundance (typically late in the assessment period when sample sizes in the standardisations are larger).

Another important input for the standardized indices is regional scaling factors which are incorporated to estimate the relative level of exploitable longline biomass among regions (see McKechnie et al., 2014a). While used as the basis for a sensitivity analysis in the assessment, the geostatistical models (Tremblay-Boyer and Pilling, 2017a) also produce an abundance surface over the entire stock assessment region, and so it is trivial to compute relative estimates of abundance among regions which are used to scale the regions-specific standardised indices, even for alternative regional

structures. We therefore used the geostatistical model outputs aggregated for the period 1980-1990 (during which the spatial coverage of fishing effort was high) to estimate the regional scaling factors. This was consistent with the approach adopted in 2014.

Catch in numbers were used for all LL fisheries in the model, and for the other longline fisheries that do not receive standardised CPUE indices the effort units were defined as the total number of hooks set divided by 100.

Recent research (MRAG Asia Pacific, 2016) has indicated the potential for systematic under reporting of bigeye catch for some fleets. An alternative catch history for the LL fisheries in the assessment was constructed at the request of the PAW and was included as a one-off sensitivity model. Details of this catch history are provided by (McKechnie et al., 2017b) and are presented in Section 7.5.1. Note that the CPUE indices are maintained in these models, with only the catches modified (and the effort adjusted to preserve the CPUE ratio). We note that inclusion of the additional catch within the CPUE calculation would somewhat reduce the declining trends generally estimate.

4.5.4   Other fisheries

There has been continual improvement in the catch estimates from Indonesia and the Philippines through the GEF-WPEA project and since the 2014 assessment catch data from fisheries in Vietnam have also been included. These improved catch estimates have been incorporated into the current assessment.

For these other fisheries effort is either included as days fished, or more often set to missing. Effort was set to missing for five fisheries, the three small-fish miscellaneous fisheries, the combined Indonesia and Philippines handline and ex-EEZ purse seine fisheries. A nominal effort of one was included for the final year of the model to allow the estimation of a catchability coefficient to assist with projection analyses.

4.6   Size data

Available length-frequency data for each of the fisheries were compiled into 95 2cm size classes (10–12 cm to 198–200 cm). Weight data were compiled into 200 1kg size classes (0–1 kg to 199–200 kg). Most weight data were recorded as processed weights (usually recorded to the nearest kilogram). Processing methods varied between fleets requiring the application of fishery-specific conversion factors to convert the available weight data to whole fish equivalents. Details of the conversion to whole weight are described in Langley et al. (2006). Data were either collected onboard by fishers,

through observer programmes, or through port sampling. Davies et al. (2011) provides more details on the source of the size data. Each length-frequency record in the model consisted of the actual number of bigeye tuna measured and Figure 5 provide details of the temporal availability of length and weight (for longline) frequency data. Note that a maximum observed sample size of 1,000 was implemented in the assessment and the effective sample size was further downweighted, with the sensitivity to the magnitude of the divisor investigated in the sensitivity and structural uncertainty analyses.

4.6.1 Purse seine

Only length frequency samples are used in the bigeye assessment for purse seine fisheries, and assessments prior to 2014 used only observer samples which had been corrected for grab-sample bias. As observer coverage had been very low and unrepresentative in early years, there were many gaps and the time series of size data did not show evidence of modal progression. Two major changes were made for the 2014 assessment and are also adopted here. These are described in detail in Abascal et al. (2014): first the long time series of port sampling data from Pago Pago was included, and second all samples were weighted by the catch - both at the set and strata level, with thresholds applied to ensure that small samples from important catch strata did not get too much weight (consistent with the approach taken for the longline fishery - see below).

4.6.2 Longline

A detailed review of all available LL length and weight frequency data for bigeye has previously been undertaken by McKechnie (2014) and that paper provides the details of the analytical approaches for constructing data inputs for the 2014 assessment, which also apply to the current assessment. The key principle used in constructing the data inputs was not to use weight and length data at the same time - even if it was available - as it would either introduce conflict (if data were in disagreement) or over-weight the model fit (if they were in agreement). Therefore, we considered the coverage and size of samples and typically chose to use weight frequency data where it was available. Japanese weight data were not available for regions 4, 5, and 6 in recent years and had to be supplemented by "all flags" length data.

The general approach used by McKechnie (2014) was that Japanese size data were weighted with respect to the spatial distribution of catch within the region, and the size data from all-fleets data were weighted by the flag-specific catch. For the catch weighting, a moving 11 quarter time window was used to calculate the relative

importance of each stratum. The exact methods used in 2014 were adopted in the current assessment, with the addition of three extra years of data. During model development the length data for LL fisheries in regions 5 and 6 were removed due to conflict with the weight data which was first encountered in 2014, and further details of this modification are provided in Section 6.1.

4.6.3 Other fisheries

Size data were either missing or poor for the Indonesian and Vietnamese small-fish fisheries and the Indonesian-Philippines ex-EEZ purse seine fishery. In the case of the first two, selectivity was assumed to be shared with the Philippines small-fish fishery and in the last case it was shared with the associated purse seine fishery, also in region 7.

Size composition data for the Philippines domestic fisheries (both small-fish fisheries and large-fish handline fisheries) were derived from a number of sampling programmes conducted in the Philippines since the 1980s. In more recent years, size-sampling data have been substantially augmented by the work of the GEF-WPEA project.

As for the 2011 and 2014 assessments the length frequency samples from the small fish miscellaneous fishery (MISC PH 7) were adjusted to exclude all reported fish lengths greater than 90 cm from the current assessment. This was done on the basis that it is suspected that the presence of these large fish may be due to mis-reporting of the fishing gear in some of the regional sampling programmes.

Length data from the Japanese coastal purse-seine and pole-and-line fleets were provided by the National Research Institute of Far Seas Fisheries (NRIFSF). For the equatorial pole-and-line fishery, length data were available from the Japanese distant-water fleet (sourced from NRIFSF) and from the domestic fleets (Solomon Islands and PNG). Since the late 1990s, most of the length data were collected by observers covering the Solomon Islands pole-and-line fleet.

4.7 Tagging data

A moderate amount of tagging data were available for incorporation into the assessment and a summary of its characteristics and the process of constructing the MFCL tagging file are presented in detail by McKechnie et al. (2017b). The data were available from the Regional Tuna Tagging Project (RTTP) during 1989–92 (including affiliated in-country projects in the Solomon Islands, Kiribati, Fiji and the Philippines), more recent (1995, 1999-2001) releases and returns from the Coral Sea Tagging Programme (CSTP) by CSIRO (Evans et al., 2008), and the Pacific Tuna

22

Tagging Programme (PTTP) carried out during the period 2006 until the 3rd quarter of 2014. Additional data have become available since the 2014 assessment for the Japanese Tagging Programme (JPTP), conducted by NRIFSF and the Ajinomoto Co. Inc, over the period 2000–2014, which were included in a sensitivity analysis. The amount of JPTP tagging data is low compared to the SPC-held (357 usable recaptures) but these come from a large number of small release events. Therefore the model runs (because MFCL effectively creates extra "populations" for each release event) are considerably slower with their inclusion, with little extra information added to the model. This precludes using these data in the structural uncertainty grid, and so we restrict our investigation of this data to sensitivity analyses.

Tags were released using standard tuna tagging equipment and techniques by trained scientists and technicians. Tags have been returned from a range of fisheries, having been recovered onboard or via processing and unloading facilities throughout the Asia-Pacific region.

In the current assessment, the numbers of tag releases input to the assessment model were adjusted for a number of sources of tag loss - unusable recaptures due to lack of adequately resolved recapture data, estimates of tag loss (shedding and initial mortality) due to variable skill of taggers, and estimates of base levels of tag shedding/tag mortality. An additional problem for the bigeye assessment is that there are a considerable number of tag returns that were released within the WCPO but recaptured to the east, outside the assessment region. The procedures used in re-scaling the releases to account for these losses are described in detail in Berger et al. (2014) and McKechnie et al. (2016b), but essentially the re-scaling preserves the recovery rates of tags from the individual tag groups that would otherwise be biased low when an often significant proportion of recaptures cannot be assigned to a recapture category in the assessment. The same processing methods were adopted for the current assessment and a more detailed summary of the process and the resulting dataset can be found in McKechnie et al. (2017b).

There is a delay between tagged fish being caught, the tag being reported and the data being entered into tagging databases. If this delay is significant then reported recapture rates for very recent release events will be biased low and will impact estimates of fishing mortality in the terminal time periods of the assessment. For this reason, any release events occurring after the second quarter of 2014 were excluded from the assessment (McKechnie et al., 2016b), which is consistent with the approach taken in 2014.

For incorporation into the assessment, tag releases were stratified by release region, time period of release (quarter) and the same size classes used to stratify the length-frequency data. A total of 17,886 effective releases were classified into 62 tag release

groups (Table 2). The returns from each size-class of each tag release group (6,344 effective, usable tag returns in total) were then classified by recapture fishery and recapture time period (quarter). Because tag returns by purse seiners were often not accompanied by information concerning the set type, tag return data were aggregated across set types for the purse seine fisheries in each region. The population dynamics model was in turn configured to predict equivalent estimated tag recaptures by these grouped fisheries.

5 Model description

5.1 General characteristics

The model can be considered to consist of several components, (i) the dynamics of the fish population; (ii) the fishery dynamics; (iii) the dynamics of tagged fish; (iv) the observation models for the data; (v) the parameter estimation procedure; and (vi) stock assessment interpretations. Detailed technical descriptions of components (i)–(iv) are given in Hampton and Fournier (2001) and Kleiber et al. (2017). In addition, we describe the procedures followed for estimating the parameters of the model and the way in which stock assessment conclusions are drawn using a series of reference points.

5.2 Population dynamics

The model partitions the population into nine spatial regions and 40 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 1952–2015. The main population dynamics processes are as follows.

5.2.1 Recruitment

Recruitment is defined as the appearance of age-class 1 fish (i.e. fish averaging ~ 20–30 cm given current growth curves) in the population. Tropical tuna spawning does not always follow a clear seasonal pattern but occurs sporadically when food supplies are plentiful (Itano, 2000). It was assumed that recruitment occurs instantaneously at the beginning of each quarter. This is a discrete approximation to continuous recruitment, but provides sufficient flexibility to allow a range of variability to be incorporated into the estimates as appropriate.

Spatially-aggregated (over all model regions) recruitment was assumed to have a weak relationship with spawning potential via a Beverton and Holt stock-recruitment relationship (SRR) with a fixed value of steepness (h). Steepness is 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 (Francis, 1992; Harley, 2011). Typically, fisheries data are not very informative about the steepness parameter of the SRR parameters (ISSF, 2011); hence, the steepness parameter was fixed at a moderate value (0.80) and the sensitivity of the model results to the value of steepness was explored by setting it to lower (0.65) and higher (0.95) values.

In the diagnostic case model, it was assumed that annual recruitment was related to annual mean spawning potential, which was recommended by the 2011 Bigeye Tuna Peer Review (Ianelli et al., 2012) and was previously assumed for the south Pacific albacore and WCPO skipjack assessments (Harley et al., 2015; McKechnie et al., 2016a). An alternative model run exploring the former assumption of estimating the SRR at the quarterly-scale was also undertaken.

The SRR was incorporated mainly so that yield analysis and population projections could be undertaken for stock assessment purposes, particularly the determination of equilibrium- and depletion-based reference points. We therefore applied a weak penalty (equivalent to a CV of 2.2) for deviation from the SRR so that it would have negligible effect on recruitment and other model estimates (Hampton and Fournier, 2001), but still allow the estimation of asymptotic recruitment. This approach was recommended (recommendation 20) by the 2011 bigeye assessment review (Ianelli et al., 2012). The SRR was calculated over the period from 1962–mid-2014 to prevent the early recruitments (which appear to be part of a different "regime" to subsequent estimates and may not be well estimated in any case), and the terminal recruitments (which are not freely estimated), from influencing the relationship, which is consistent with the approach of the 2014 assessment.

In recent assessments of tuna in the WCPO the terminal recruitments have often been fixed at the mean recruitment of the rest of the model period. This acknowledges that these estimates are poorly supported by data and if unconstrained can vary widely, with potentially large consequences for stock projections. This approach has been continued here by fixing the six terminal recruitments at the mean of the recruitments over the rest of the assessment period.

The distribution of recruitment among the model regions was estimated within the model and allowed to vary over time in a relatively unconstrained fashion.

5.2.2 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.

5.2.3 Growth

The standard assumptions for WCPO assessments fitted in MFCL were made concerning age and growth: 1) the lengths-at-age are normally distributed for each age-class; 2) the mean lengths-at-age follow a von Bertalanffy (VB) growth curve; 3) the standard deviations of length for each age-class are a log-linear function of the mean lengths-at-age; and 4) 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 and temporally 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.

The growth curve is a major focus in the current assessment due to new otolith aging data that suggest a significant departure from the growth curve used in the 2014 assessment (Farley et al., 2017). McKechnie et al. (2017b) explore fitting a range of models to these new otolith data external to MFCL with the aim of providing potential growth curves that could be used in the assessment. Most of the growth estimates are very similar and display an L2 parameter (asymptotic size of an average old fish) of around 150–160cm, which is considerably smaller than the value of 184 cm used in the 2014 reference case model.

Three alternative approaches to modelling growth were investigated in the current assessment, all assuming a VB growth function:

– *Old growth, 2014 methodology:* The exact methodology used in the 2014 reference case model. The L2 parameter is fixed at 184 cm and the L1 and K parameters are freely estimated by MFCL as it fits growth on the basis of modal progression in the size frequency data. Small deviations from the VB for ages 2–8 were allowed using penalisation, as was undertaken in 2014. The two length-at-age standard deviation (SD) parameters, the "generic" SD and the "age-dependent" SD (Kleiber et al., 2017) were also freely estimated.

– *New growth, fixed external otolith estimate:* The VB growth function estimated by McKechnie et al. (2017b), by fitting to the annual otolith dataset (supplemented with daily otoliths for young fish from CSIRO; model *Annual-Ot-VB*). Again, the two length-at-age SD parameters were freely estimated given that variation in asymptotic length for the otolith dataset is likely to be underestimated, and to allow the model to fit the size data adequately, especially for large fish caught in LL fisheries. The L1 parameter was also estimated, to allow it to estimate a slightly lower value that improved the fit to the length frequencies of the small-fish fisheries.

– *New growth, internal estimate with conditional age-length data:* A conditional age-length dataset was constructed from the combined daily and annual otolith dataset (McKechnie et al., 2017b) and was input to MFCL, with all growth parameters freely estimated from these data, in addition to the other data components, most notably the size data displaying modal progression. The L1, L2, K and two SD parameters were freely estimated. Small deviations from the VB for young age classes were allowed using penalisation, in the same manner as the old growth model. Preliminary model fits suggested that these deviates were minor after age-class 5 and so only deviates up to this age were estimated.

5.2.4 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., 2017 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.region boundaries (13)×4 quarters. 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 low penalty is applied to deviations from the prior.

5.2.5 Natural mortality

Natural mortality (M) may be held fixed at pre-determined, age-specific values or estimated as age-specific parameters. M-at-age was recalculated for previous assessments using an approach applied to other tunas in the WCPO and EPO (Harley and Maunder, 2003; Hoyle, 2008; Hoyle and Nicol, 2008). The generally increasing proportion of males in the catch with increasing size is assumed to be due to an increase in the natural mortality of females, associated with sexual maturity and the onset of reproduction. The externally-estimated M-at-age function was input to MFCL as fixed values and is shown in Figure 10 for the diagnostic case model. Because the fixed values of M-at-age are initially calculated at-length and then back transformed to at-age using a growth curve, it is important to calculate a specific M-at-age for each growth curve used in the modelling. McKechnie et al. (2017b) detail this process for a number of growth curves, and provide the details of the data and models used in the estimation process. For the purposes of this paper it should be noted that we utilise a specific M-at-age for each of the growth curves used in the various models in the assessment (these are presented in Figure 10).

An alternative to fixed M-at-age is to allow it to be estimated internally in MFCL. This was investigated in the 2014 assessment using a penalisation approach in sensitivity analyses and in the structural uncertainty grid. MFCL has recently been developed to include functional forms for M estimation subsequent to the 2014 assessment, including a Lorenzen-type (Lorenzen, 1996) M-at-age function, and a parameterisation based on cubic splines. We investigate the Lorenzen parameterisation in sensitivity analyses and present these in more detail in Section 6.2.

5.2.6 Sexual maturity, or reproductive potential-at-age

Reproductive output at age, which is used to derive spawning biomass, attempts to provide a measure of the relative contribution of fish at different ages to the next generation. The maturity-at-age was calculated from data collected in the WCPO, and was based on relative reproductive potential rather than the relative biomass of both sexes above the age of female maturity. This approach was previously applied to albacore (Hoyle, 2008) and bigeye (Hoyle and Nicol, 2008) tunas in the WCPO. The reproductive potential of each age class was assumed to be the product of the proportion female at age, the proportion of females mature at age, the spawning frequency at age of mature females, and the fecundity at age per spawning of mature females (Figure 11). For the diagnostic case model, the growth function, estimated maturity-at-age (from the new data of Farley et al., 2017) and updated proportion of females at-age (from Section 5.2.5) were all updated from the 2014 assessment. The

relative index of spawning potential for the diagnostic case model and for the other growth functions explored during the assessment are shown in this figure. The new growth function, combined with the new maturity function (see Farley et al., 2017 and McKechnie et al., 2017b for more details) results in a slight shift in the age of first maturity towards younger fish, and less reduction in the reproductive potential for older age classes compared to the old growth/old maturity estimates (Figure 11).

The process of estimating reproductive potential-at-age shares some similarities with calculation of M-at-age (Section 5.2.5) in that the function is calculated at-length and then backtransformed to at-age using an assumed growth function. Again, a detailed description of this process is provided by McKechnie et al. (2017b). In short, the approach taken for the 2017 assessment follows exactly the methods conducted in previous bigeye stock assessments except that updated observer sex ratio data for longline fisheries and a maturity function estimated from new data from Farley et al. (2017) were incorporated into the estimation. Again, specific reproductive potential-at-age functions were calculated for each growth curve used in the model development and uncertainty analyses.

5.3 Fishery dynamics

The interaction of the fisheries with the population occurs through fishing mortality. Fishing mortality is assumed to be a composite of several separable processes - selectivity, which describes the age-specific pattern of fishing mortality; catchability, which scales fishing effort to fishing mortality; and effort deviations, which are a random effect in the fishing effort - fishing mortality relationship.

5.3.1 Selectivity

In many stock assessment models, selectivity is modelled as a functional relationship with age, e.g. using a logistic curve to model monotonically increasing selectivity and various dome-shaped curves to model fisheries that select neither the youngest nor oldest fish. Modelling selectivity with separate age-specific coefficients (with a range of 0-1), constrained with smoothing penalties, allows more flexibility but has the disadvantage of requiring a large number of parameters. In most cases we have instead used the same methods as the 2014 assessment which was based on cubic spline interpolation techniques. This is a form of smoothing, but the number of parameters for each fishery is the number of cubic spline "nodes" that are deemed to be sufficient to characterise selectivity over the age range. We use five nodes, which seems to be sufficient to allow for reasonably complex selectivity patterns.

For particular fisheries, alternative functions were employed, including logistic selectivity (Table 3). In all cases, selectivity is assumed to be fishery-specific and time-invariant. However, it is possible for a single selectivity function to be "shared" among a group of fisheries that have similar operational characteristics and/or exist in similar areas and with similar length compositions. This grouping facilitates a reduction in the number parameters being estimated and the groupings used are provided in Table 3.

While full length-based selectivity is not currently permitted in MFCL, the age-based selectivity functions are penalised such that selectivity of age-classes that are similar in size will have similar selectivities for a given fishery or group of fisheries.

5.3.2 Catchability

Constant (time-invariant) catchability was assumed for all fisheries that received standardised indices of relative abundance (the LL ALL fisheries in each region; Table 3). This assumption is similar to assuming that the CPUE for these fisheries indexes the exploitable abundance over time. The LL ALL fisheries were grouped for the purpose of initial catchability, and to maintain the relativity of catch rates among regions. This provides the model with information on the relative population sizes among regions, which previous experience suggests is difficult to estimate without this assumption.

For all other fisheries, catchability was allowed to vary slowly over time (akin to a random walk) using a structural time-series approach. Random walk steps were taken every two years, and the deviations were constrained by prior distributions of mean zero and variance specified for the different fisheries according to our prior belief regarding the extent to which catchability may have changed. For fisheries having no available effort estimates (e.g. the Philippines and Indonesian surface fisheries), partial fishing mortalities were estimated consistent with the observed catches using a Newton-Raphson procedure. Therefore, catchability deviations (and effort deviations) are not estimated for these fisheries.

For the other fisheries with time-series variability in catchability, the catchability deviation priors were assigned a variance approximating a CV of 0.10. Apart from those fisheries for which the data were based on annual estimates, the catchabilities of all other fisheries were allowed to vary seasonally, including those fisheries that received standardised CPUE indices (Table 3).

5.3.3 Effort deviations

Effort deviations were used to model the random variation in the effort - fishing mortality relationship, and are constrained by pre-specified prior distributions (on the log-scale). There were several categories of fisheries with respect to the effort deviation penalties applied and these are outlined in Table 3 and presented in Figure 12. The region-specific CPUE indices represent the principal indices of stock abundance, and the extent to which the model can deviate from the indices is moderated by the penalty weights assigned to the standardised effort series. For these fisheries the prior was set to have a mean of zero and the CV was allowed to be time-variant and based on the variance estimates (using the canonical variance method of Francis, 1999) from the GLMs fitted to each fishery (McKechnie et al., 2017b). As occurred in the 2014 assessment and as explained in Section 4.5, the regional differences in the estimated CVs were sufficiently small that we assumed the same average CV for all indices; the average CV for the period 1980-90 was set to 0.2. The resulting scaled CVs were transformed to an effort deviate penalty for each CPUE observation in MFCL. Penalties are inversely related to variance, such that lower effort penalties are associated with indices having high variance, consequently these indices are less influential in fitting the model.

The miscellaneous fisheries have very unreliable estimates of effort and to prevent them from influencing population dynamics they are given missing effort at all time-steps except the last four quarters, when they are given an arbitrary effort of one, and receive effort deviation penalties equivalent to a CV of about 0.22. This is done simply to provide a basis for scaling effort so that the effort deviates converge around 0 to allow effort-based population projections.

Finally, for all other fisheries the nominal effort was used, but to prevent the CPUE of these fisheries from influencing population dynamics they received effort deviation penalties equivalent to a CV of about 0.7 for the average effort (Table 3). The penalities for many of the fisheries are assumed to be caled according to the square root of the observed effort such that low penalites are applied for low observed effort and higher penalites are applied for high effort.

5.4 Dynamics of tagged fish

5.4.1 Tag reporting

In principle, tag-reporting rates can be estimated internally within the model. In practice, experience has shown that independent information on tag-reporting rates for at least some fisheries tends to be required for reasonable model behaviour to

be obtained. In addition to varying by fishery, we allowed reporting rates to also vary among tagging "programmes" implemented at different times in the history of the fishery, or conducted by different agencies. We provided reporting rate priors for all fishery/tagging programme groups that reflect independent estimates of the reporting rates and their variances. These were derived from analyses of tag seeding experiments (Peatman et al., 2016). For the RTTP and PTTP, relatively informative priors were formulated for the equatorial purse seine fisheries given that tag seeding experiments were focused on purse seiners. All reporting rates within a tagging programme were assumed to be time-invariant. Tag recapture and reporting rate groupings are provided in Table 3.

5.4.2 Tag mixing

The population dynamics of the fully recruited tagged and untagged 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 distribution 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 in that model region. It is therefore desirable to designate one or more time periods (quarters) after release as "pre-mixed" and compute fishing mortality for the tagged fish based on the actual recaptures, corrected for tag reporting, 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 removing fish from the tagged population for the recaptures that occurred. We assume that tagged bigeye gradually mix with the untagged population at the region-level and that this mixing process is complete by the end of the second quarter after release. We investigate the robustness to this assumption in sensitivity analyses.

Tagged fish are modelled as discrete cohorts based on the region, year, quarter

and age at release for the first 30 quarters after release. Subsequently, the tagged fish are pooled into a common group. This is to limit memory and computational requirements.

5.5 Likelihood components

There are four data components that contribute to the log-likelihood function for this assessment - the total catch data, the length-frequency data, the weight-frequency data and the tagging data. For the model in which the conditional age-at-length data were included, this constituted a fifth data component to the log-likelihood.

The observed total catch data are assumed to be unbiased and relatively precise, with the SD of residuals on the log scale being 0.002. Note that this is close to the "catch conditioned" approach often used in other integrated assessments.

The probability distributions for the length- and weight-frequency proportions are assumed to be approximated by robust normal distributions, with the variance determined by the effective sample size (ESS) and the observed length-frequency proportion. Size frequency samples are assigned ESS lower than the number of fish measured. Lower ESS recognise that (i) length- and weight-frequency samples are not truly random (because of non-independence 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. We divided the observed sample sizes by 20, and then applied a maximum ESS of 50 to each length and weight sample for a fishery. Alternative divisors for specifying ESS were explored in sensitivity analyses.

We also examined in a sensitivity analysis the application of a new size-likelihood based on the Dirichlet-Multinomial (Thorson et al., 2017). This method allows the estimation of ESS from the data, conditional on the specification of maximum ESS, thus substantially removing the need for arbitrary assumptions about ESS. As this is a new feature of MFCL (Davies et al., 2017), we did not apply it to all of the models developed in this assessment. However, once testing and documentation of the method is complete we expect it will become the standard likelihood function for compositional data in MFCL in the future.

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 overdispersion parameter such that as it approaches 1, 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 recognise 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. A complete derivation and description of the negative binomial likelihood function for tagging data is provided in Kleiber et al. (2017).

A further log-likelihood component is introduced for models that include the conditional age-at-length dataset (McKechnie et al., 2017b). These data can be included in the assessment to assist in estimating growth parameters because they provide direct observations of the distribution of fish ages within length classes. These data are assigned to fisheries and model time-steps (see McKechnie et al., 2017b for complete details of file construction) and the model fits the observed data based on the selectivity of the capture fishery and the predicted lengths and standard deviations of the modelled population of fish defined by the growth function. The observed age composition within each length interval is assumed to be multinomially distributed, and this forms the basis of the likelihood component for this data source.

5.6 Parameter estimation and uncertainty

The parameters of the model were estimated by maximizing the log-likelihood of all data components 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 relatively 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 bet.ini input file.

In this assessment two approaches were used to describe the uncertainty in key model outputs. The first estimates the statistical uncertainty within a given assessment model, while the second focuses on the structural uncertainty in the assessment by considering the variation among a suite of models.

For the first approach, the Hessian was calculated for the diagnostic case model run to obtain estimates of the covariance matrix, which is used in combination with the delta method to compute approximate confidence intervals for parameters of interest (for example, the biomass and recruitment trajectories). For the second approach, a factorial grid of model runs was undertaken which incorporated many of the options

of uncertainty explored in 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. Therefore, a likelihood profile analysis was conducted for the marginal posterior likelihood in respect of the total average population biomass as a measure of population scaling (Lee et al., 2014, with the definition of this parameter detailed in Kleiber et al., 2017). Previous assessments had conducted likelihood profiling for the total population scaling parameter in MFCL. While this was convenient in the sense that this was a model parameter that could be fixed to obtain points along the profile, we realised that there was potential confounding of this parameter with the mean of the recruitment deviations, such that an increase or decrease in the scaling parameter could be compensated at least in part by a shift in the distribution of recruitment deviations in the opposite direction. We therefore opted to use the population biomass averaged over the entire model period as a basis for generating the likelihood profile. This has the slight complication that average biomass is not a model parameter, but is a derived quantity determined by all model parameters. It was therefore necessary to take a penalised approach, whereby the model is penalised for deviations from the "target" biomass for each point on the profile. Initially, the penalties are set to be small, and progressively increased to be very large to obtain a final model fit conditioned on the specified average biomass. Reasonable contrast in the profile obtained using this method is taken to indicate that sufficient information existed in the data for estimating absolute abundance, and also offered confirmation that the maximum likelihood estimate obtained represented a global solution, at least with respect to total population scaling. This procedure is presented in the Appendix (Section 11.1), including examination of the profiles for the individual data components.

Retrospective analyses are also undertaken as a general test of the stability of the model, as a robust model should produce similar output when rerun with data for the terminal year/s sequentially excluded (Cadigan and Farrell, 2005). The retrospective analyses for the 2017 diagnostic case model are presented in the Appendix (Section 11.2).

5.7   Stock assessment interpretation methods

Several ancillary analyses using the fitted model/suite of models were conducted in order to interpret the results for stock assessment purposes. The methods involved are summarized below and further details can be found in Kleiber et al. (2017).

5.7.1 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 using the formulae given in Kleiber et al. (2017), and is equivalent to the MSY. Similarly the spawning potential at MSY (SB_{MSY}) can also be determined. The ratios of the current (or recent average) levels of fishing mortality and biomass to their respective levels at MSY are determined for all models of interest, including those in the structural uncertainty grid, and so alternative values of steepness were assumed for the SRR in many of them.

Fishing mortality-at-age (F_a) for the yield analysis was determined as the mean over a recent period of time (2011–2014). We do not include 2015 in the average as fishing mortality tends to have high uncertainty for the terminal data year of the analysis and the terminal recruitments in this year are constrained to be the average over the full time-series, which affects F for the youngest age-classes.

MSY was also computed using the average annual F_a from each year included in the model (1952–2015). This enabled temporal trends in MSY to be assessed and a consideration of the differences in MSY levels under historical patterns of age-specific exploitation. More details of this approach are provided in Section 5.7.4.

5.7.2 Depletion and fishery impact

Many assessments estimate the ratio of recent to initial biomass (usually spawning biomass) as an index of fishery depletion. The problem with this approach is that recruitment may vary considerably over the time series, and if either the initial or recent 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 potential time series (at the region level) using the estimated model parameters, but assuming that fishing mortality was zero. Because both the estimated spawning potential SB_t (with fishing), and the unexploited spawning potential $SB_{F=0[t]}$, incorporate recruitment variability, their ratio at each quarterly time step (t) of the analysis, $SB_t/SB_{F=0[t]}$, 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. To achieve this the estimated recruitment deviations are multiplied by a scalar based on the difference in the SRR between the estimated fished and unfished spawning potential estimates.

A similar approach can be used to estimate depletion associated with specific fisheries or groups of fisheries. Here, fishery groups of interest - LL, PS associated sets, PS unassociated sets, pole and line and "other" fisheries, are removed in-turn in separate simulations. The changes in depletion observed in these runs are then indicative of the depletion caused by the removed fisheries.

5.7.3 Reference points

The unfished spawning potential ($SB_{F=0}$) in each time period was calculated given the estimated recruitments and the Beverton-Holt SRR as outlined in Section 5.7.2. This offers a basis for comparing the exploited population relative to the population subject to natural mortality only. The WCPFC adopted $20\%SB_{F=0}$ as a limit reference point (LRP) for the bigeye stock, where $SB_{F=0}$ is calculated as the average over the period 2005–2014. Stock status was referenced against these points by calculating $SB_{recent}/SB_{F=0}$ and $SB_{latest}/SB_{F=0}$, where SB_{latest} and SB_{recent} are the estimated spawning potential in 2015, and the mean over 2011–2014, respectively (Table 4).

The other key reference point, $F_{recent}/F_{\mathrm{MSY}}$ (Table 4), is the estimated average fishing mortality over the full assessment area over a recent period of time (F_{recent}; 2011–2014 for this stock assessment) divided by the fishing mortality producing MSY which is produced by the yield analysis and has been detailed in Section 5.7.1.

5.7.4 Majuro and Kobe plots

For the standard yield analysis (Section 5.7.1), the fishing mortality-at-age, F_a, is determined as the average over some recent period of time (2011–2014 herein). In addition to this approach the MSY-based reference points (F_t/F_{MSY}, and SB_t/SB_{MSY}) and the depletion-based reference point ($SB_t/SB_{F=0[t]}$) were also computed using the average annual F_a from each year included in the model (1952–2014, with no value calculated for the terminal year) by repeating the yield analysis for each year in turn. This enabled temporal trends in the reference point variables to be estimated taking account of the differences in MSY levels under varying historical patterns of

age-specific exploitation. This analysis is presented in the form of dynamic Kobe plots and "Majuro plots", which have been presented for all WCPO stock assessments in recent years.

6 Model runs

6.1 Developments from the last assessment

The progression of model development from the 2014 reference case to the model proposed as the diagnostic case in 2017 was incremental, with stepwise changes to the model made in-turn to ensure that the consequences of each modification could be ascertained. Changes made to the previous assessment model include additional input data for the years 2013–2015, modified methods in producing the input files (e.g. McKechnie et al., 2017b; Tremblay-Boyer and Pilling, 2017a), new regional structures (McKechnie et al., 2016b), updated biological information and data (Farley et al., 2017), and implementation of additional, or new features of MFCL (Davies et al., 2015). An outline, and basis for, the progression through the models is as follows:

(a) The 2014 reference case model [*Ref14*].

(b) The 2014 reference case model with the new MFCL executable [*Ref14-NewEx*].

(c) A complete update of the 2014 reference case model - all inputs extended from 2012 to 2015 using identical methodology for CPUE, tagging, size frequencies etc, and the same MFCL model settings [*Ref14Update*].

(d) The previous model with the same structure and MFCL settings but CPUE indices using the new GLM approaches on the operational LL database (Tremblay-Boyer and Pilling, 2017a) [*Update-CPUE*].

(e) The previous model with the same MFCL settings but with the new regional structure and consequently all fisheries, and input data (including CPUE standardisations), reconfigured based on these new regional definitions [*NewAreas*].

(f) The previous model but with (i) a fixed growth curve (which has a considerably lower asymptotic length than assumed in previous models) estimated externally to MFCL based on the latest otolith data set (McKechnie et al., 2017b); and (ii) the new maturity-at-length schedule estimated by Farley et al. (2017) converted to maturity-at-age using the new growth model. Note that the fixed M-at-age, which is a function of growth and maturity, it also adjusted at this step. [*NewGrthMat*].

(g) The previous model with two modifications to the recruitment estimates; the change from quarterly to annual recruitments when estimating the spawner-recruit relationship, and the fixed terminal six recruits set at the arithmetic rather than geometric mean of recruitments for the remaining period [*RecMods*].

(h) The previous model with length frequency data for LL fisheries 11 (LL-ALL-5) and 12 (LL-ALL-6) removed, leaving only weight frequency data, as there was considerable conflict between the two in the 2014, and current, assessment. [*Diagnostic*]

6.2 Sensitivity analyses

Several hundred runs were undertaken in conducting the 2017 bigeye assessment, but in terms of presenting information on the bounds of plausible model sensitivity we have focused on a small set of uncertainty axes which are described in further detail below. These axes were used for "one-off" changes from the diagnostic case model and several of these sensitivity models were used in the structural sensitivity analyses (after Hoyle et al., 2008). The latter process involves constructing a grid of model runs where all-possible combinations of the assumptions are explored (see Section 6.3).

The recommendations of the 2017 PAW formed the basis for several of the one-off sensitivity analyses undertaken from the diagnostic case, but several other runs were undertaken in order to provide a better understanding of the impact of some of the changes in modelling assumptions. Each of the one-off sensitivity runs was carried out by making a single change to the diagnostic case, and it should again be reinforced that these model runs are not carried out to provide absolute estimates of management quantities but to demonstrate the *relative* changes that result from the various changed assumptions, and for that reason the reference points are presented in the Appendix (Tables 11–16).

6.2.1 Steepness [*h0.65, h0.95*]

Steepness is a particularly difficult parameter to estimate in stock assessment models, but if it is fixed in the model, the choice of value may have significant influence on most reference points used in management. As was the case in other tropical tuna and albacore tuna assessments, we assumed a value of 0.8 for the diagnostic case, but examined values of 0.65 (*h0.65*) and 0.95 (*h0.95*) in sensitivity runs. This choice of values is consistent with the results of the meta-analysis conducted on

tuna stock-recruitment data and has been well established in previous Scientific Committees.

6.2.2 Tag mixing period [*Mix1, Mix-CS*]

The tag mixing period is imposed to allow tagged fish to distribute themselves throughout the region of tagging, although it is somewhat difficult to ascertain how long this period should be. In the diagnostic case model the mixing period was set at two quarters and an alternative model was run assuming a mixing period of one quarter (*Mix1*). A further sensitivity model was run with the tags for the Coral Sea releases set at 28 quarters (*Mix-CS*) to effectively remove the influence of these tags on the model likelihood. These releases were very influential in assessments before 2014 and this sensitivity model has been previously used to assess their impact in subsequent assessments.

6.2.3 Relative weighting of length and weight frequency data [*Size10, Size50*]

The difficulties in assigning weighting to the length frequency data were discussed in Section 5.5. To assess the sensitivity of model results to the weighting of these data, two alternative models were considered; a model where frequency data were up-weighted (corresponding to a maximum effective sample size of 100 fish) relative to the diagnostic case model (*Size10*); and one where those data were down-weighted (corresponding to a maximum effective sample size of 20 fish) relative to the diagnostic case (*Size50*).

6.2.4 Dirichlet multinomial likelihood for the size frequency data [*Dch*]

The Dirichlet multinomial likelihood is a new development in MFCL (Davies et al., 2017), and offers the potential to estimate the effective sample size of frequency samples used in the model, thus obviating the need for largely arbitrary assumptions that determine the overall weighting of the size data in the model likelihood. We have applied the Dirichlet multinomial in a sensitivity analysis in this assessment, assuming a maximum ESS of 100, pending further testing and evaluation of its performance.

6.2.5 Alternative growth functions [*AL, L2-184*]

The importance of the growth function and the details of the alternative methods used in the assessment are detailed in Section 5.2.3, with two methods additional

to the diagnostic case being implemented as sensitivity models. The first sensitivity fits growth by including the conditional age-length data, constructed from daily and annual otolith readings, within the model-fitting procedure of MFCL. This data set therefore contributes a likelihood component to the model. In this sensitivity, we simply assumed a multinomial ESS for each of the fishery- and time-specific age-length samples equal to the observed sample size. Future implementations in MFCL will use the Dirichlet multinomial, to recognise aging errors as well as unaccounted-for process error impacting the observations. The second sensitivity reverted to the growth specifications used in the 2014 reference case model, i.e. fixing the L2 parameter at 184 cm and estimating all other growth parameters from the data internally in MFCL. This sensitivity is similar to the model (*NewAreas*) in the step-wise development, but is applied to the diagnostic case where all model developments have been incorporated.

6.2.6   Old growth with new maturity/natural mortality [*L2-184-NewMat*]

In the step-wise model development, we implemented both new growth and new maturity in the same step. In this sensitivity, we implemented just the new maturity, and its impacts on the M-at-age specification, in combination with the old growth specification used in the 2014 assessment. This was done so as to observe the separate effects of the new growth and new maturity on the model results of main interest.

6.2.7   Overdispersion in the tagging likelihood [*OD2*]

The negative binomial likelihood for the tagging data includes a parameter (τ) that determines the variance of the negative binomial distribution in relation to the Poisson distribution (where the variance is equal to the mean). Therefore, the τ parameter effectively determines the weighting received by the tagging data in the overall model likelihood. Initial attempts to estimate τ (assumed common for all fisheries and tag release groups) saw it converge to a value close to 1, i.e. Poisson. This setting was therefore used in the Diagnostic Case. It is likely, however, that there is overdispersion in the tagging data resulting from non-independent tagged fish, lack of mixing, and other unaccounted-for process error. Therefore, this sensitivity analysis assumed a negative binomial variance of twice the Poisson, i.e. $\tau = 2$, in order to determine the impact of down-weighting the tagging data in the overall likelihood.

6.2.8   Quarterly stock recruitment relationship [*SRR-qtr*]

To assess the sensitivity of the model to the new procedure that assumes deviations in total recruitment from the SRR are computed annually (rather than quarterly in the

2014 assessment), an alternative model was run reverting to the previous quarterly procedure for computing recruitment deviations.

6.2.9 Estimate natural mortality M-at-age [*Lorenzen*]

While the M-at-age function in the diagnostic case model was fixed at the values estimated externally from MFCL (McKechnie et al., 2017b), a sensitivity model was constructed to estimate natural mortality using recent developments in MFCL that allow biologically reasonable functional forms for M-at-age (Kleiber et al., 2017). We utilise the Lorenzen approach (Lorenzen, 1996) which assumes a monotonically declining relationship between M and the mean length of fish in successively older age classes a, l_a, such that $M_a = c(l_a)^b$, with two estimable parameters, b and c.

6.2.10 Alternative standardised CPUE indices [*CPUE-Proxy, CPUE-Geostat*]

The model CPUE-Proxy assesses the effects of using the alternative GLM-standardised CPUE indices calculated by Tremblay-Boyer and Pilling (2017b). These indices were an attempt to account for some of the problems associated with missing identification of JP vessels in the operational dataset, which makes standardising out vessel effects difficult.

A further sensitivity model (*CPUE-Geostat*) was fitted that included the geostatistical CPUE indices estimated by Tremblay-Boyer and Pilling (2017a). These standardisations attempted to account for some of the issues that are encountered when the spatial distribution of fishing effort in the CPUE dataset changes over the model period.

6.2.11 Inclusion of tags from the Japanese Tagging Programme [*JP*]

The Japanese Tagging Programme provided additional tagging data for the 2017 assessments of bigeye and yellowfin. However in the case of bigeye, the additional data required a large number of additional tag-release groups to be defined, which considerably slowed the parameter estimation procedure. It was therefore not feasible to include these data in the main model runs. Instead, this sensitivity model included these data to the diagnostic case model to assess their impact on model output.

6.2.12 Relaxation of reporting rate parameter bounds [*RR0.99*]

There has been concern that estimates of reporting rate parameters at their bounds (typically 0.9 by default) have affected previous assessments of bigeye and other species by acting as a limit on biomass estimates. This sensitivity relaxes the bound such that it occurs at a value of 0.99 rather than 0.9 in the diagnostic case model.

6.2.13 Uncertainty in longline catch reporting [*C-Uncert*]

An alternative catch history for LL fisheries was constructed based on evidence for under-reporting of bigeye catch (MRAG Asia Pacific, 2016), and is outlined in detail in McKechnie et al. (2017b). This catch history assumes that under-reporting of LL catch of bigeye was at a low base rate throughout the assessment period until 2005, when tropical tuna CMMs were implemented. From 2005 until 2012 the rate of under-reporting was assumed to increase linearly to the under-reporting level estimated by MRAG Asia Pacific (2016), and was maintained at this level for the remaining years in the assessment. This scaling of the catch history was applied to all longline fisheries in the assessment as a pessimistic scenario for the accuracy of LL catch estimates compared to the diagnostic case model.

6.2.14 Regional structure [*2014Reg*]

The progression from the 2014 reference case model to the 2017 diagnostic case model (Section 6.1) identified significant changes in biomass, depletion and stock status estimates during the progression from the 2014 to 2017 regional structures. A sensitivity model was therefore constructed where the diagnostic case model settings were maintained but reverting to the previous 2014 regional structure and fisheries configuration.

6.3 Structural uncertainty

Stock assessments of pelagic species in the WCPO in recent years have utilised an approach to assess the structural uncertainty in the assessment model by running a "grid" of models to explore the interactions among selected "axes" of uncertainty. The grid contains all combinations of two or more parameter settings or assumptions for each uncertainty axis. The axes are generally selected from those factors explored in the one-off sensitivities with the aim of providing an approximate understanding of variability in model estimates due to assumptions in model structure not accounted

for by statistical uncertainty estimated in a single model run, or over a set of one-off sensitivities.

The structural uncertainty grid for the 2017 assessment was constructed from 5 axes – steepness (3 settings), growth/maturity (2), tagging data overdispersion (2), size data weighting (2) and regional structure (2), with the settings used directly comparable to those presented in Section 6.2 through identical notation. The final grid thus consisted of 48 models (Table 5). Note that due to the very large computational load of running the grid, with each model requiring significant run time, the size frequency data weighting level for the diagnostic case model (divisor of 20) was not included. Instead the upper (divisor of 10 *Size10*) and lower (divisor of 50 *Size50*) weightings of the sensitivities were used to bound the uncertainty from this axis, thus also reducing the number of models that had to be run.

7   Results

7.1   Consequences of key model developments

The progression of model development from the 2014 reference case model to the 2017 diagnostic case model is outlined in Section 6.1 and results are displayed in Figure 13. A summary of the consequences of this progression through the models is as follows:

– The reference case model refitted with the latest version of MFCL (red line; Figure 13) produced very similar results to the 2014 version, with only a small deviation from the reference case model in the early years.

– The model with fully updated datasets (catch/effort, size frequencies, tagging) scaled both spawning potential and fisheries depletion down moderately (gold line; Figure 13), although temporal changes in both model outputs remained very similar to the previous model.

– The model that included the CPUE indices calculated for the operational LL dataset estimated a slightly flatter trajectory of spawning potential and less fisheries depletion from about 1970 onwards (dark green line; Figure 13).

– The model that introduced the 2017 regional structure estimated a higher spawning potential time-series than the previous model (purple line; Figure 13) although the relative trends in abundance were very similar. The trajectory of fisheries depletion was moderately more optimistic with the terminal depletion estimate about 10 percentage points above the previous model.

– The next stepwise model introduced the new (low) fixed growth rate estimated for the new otolith data external to MFCL. This had the most substantial impact on model output with much higher estimates of spawning potential and a significant upturn in in the last several years of the time-series (dark blue line; Figure 13). This model also estimated much more optimistic stock status with respect to fisheries depletion, with the terminal estimate of spawning potential at over 0.4 of the unfished estimate.

– The final two stepwise progressions (modifying the recruitment assumptions - annual SRR and setting the fixed terminal recruitments at the arithmetic mean of the other recruitment; and removing the conflicting LL length frequency data) had a negligible impact on estimates of spawning potential and fisheries depletion (light green and black lines; Figure 13), and the latter model is considered the diagnostic case model, whereby it is used to present fits to the various data components and other model diagnostics. We do not place any particular emphasis on this single model when assessing stock status and formulating management advice. Instead, we recommend that focus should be placed on the full range of models presented in the structural uncertainty analysis.

7.2 Model fit for the diagnostic case model

This section discusses the diagnostics for the diagnostic case model, defined by the final step in the stepwise model development (model *Diagnostic*; Figure 13), the final step described in Section 7.1.

7.2.1 Catch data

Very high penalties were applied to the catch data for all fisheries and so the residuals of the observed and model-predicted catches were very small (Figure 14).

7.2.2 Standardised CPUE

There was substantial temporal variability in the standardised CPUE indices used in the assessment, but despite this, the model-predicted CPUE fitted the indices very well(Figure 15). In general, the model captures most of the seasonal variation in the more temperate regions and the more stable dynamics in the equatorial regions. Of particular importance is an upturn in observed CPUE in many of the regions in the final few quarters, which are also predicted by the model (with the exception of region 4 where the model under-predicts CPUE in the final year).

There is some minor lack of fit early in the time series in regions 2 and 4 in the very early period which can be partly attributed to the extreme variability in the CPUE indices in these years. That variability cannot be explained by seasonal effects alone and probably results from low sample sizes in the operational dataset. Consequently, these values have very high CV's which allows the model to deviate away from them somewhat (Figure 9). These patterns are reflected in the time-series of effort deviations (Figure 16) where there is a tendency for positive and negative deviations early in the time series in regions 2 and 4, respectively. In several other regions the observed data displayed more pronounced variation than the model predictions although in some cases the observed indices likely reflect the poor quality of the data with small samples sizes (e.g. region 7; Figure 15)

The CPUE of fisheries not receiving standardised CPUE indices have far less influence in estimating temporal trends in biomass as much lower effort deviation penalties are applied and catchability is allowed to vary over time in these fisheries. Examination of the estimated effort deviations for these fisheries did not reveal any cause for concern, with deviations centred around zero and no temporal trends (that might indicate misspecification of temporal changes in catchability etc.; Figure 17).

7.2.3 Size frequency data

There was a generally reasonable fit to the length frequency data for fisheries with adequate sample sizes (Figure 18), particularly the LL fishery in region 4 (L-ALL-4), the important PS fisheries and the miscellaneous PH fishery in region 7 (Z-PH-7). There is some modality in the observed frequencies for some of these fisheries that could not be completely fitted by the model and presumably relates to the diverse sources of some of these data. Several fisheries display some lack of fit but in most cases this reflects a lack of data over most of the time-series and/or the sharing of selectivity as sample sizes are inadequate to fit separate selectivities in many cases (e.g. many of the PL fisheries). These are also often low volume fisheries.

The weight frequency data for the longline fisheries were generally well fitted by the model (Figure 19). There is a slight overestimation of large fish in several fisheries (e.g. the LL fisheries in regions 3, 7 and 8) although to a much lesser degree than in past assessments. As was the case in the 2014 assessment, not all modality in the observed data was predicted by the model (e.g. L-ALL-1) and this may be due to the difficulty of attributing size data to different fishing operations (distant water vs offshore). The fit to sizes for the longline fisheries in regions 5 and 6 (L-ALL-5 and L-ALL-6) has been improved with the removal of the conflicting length data that was previously included in the assessment.

These patterns were largely reflected in the comparison of temporal variation in observed and model-predicted median fish lengths (Figure 20) and weights (Figure 21). Some temporal lack of fit can be seen in several fisheries where high time-series variation in observed sizes is evident and is unlikely to reflect variation in sizes caught by the overall fishery or changes in the sizes of the underlying population (e.g. L-ALL-4, P-ALL-4). More and better sampling of size frequencies of the catch for these fisheries would be beneficial for the assessment. The temporal variation in the weight frequencies was again better fitted than the lengths. Exceptions to this were the model predictions diverging from the observed sizes in the offshore longline fishery in region 7 (L-OS-7) and the apparent conflict between the longline fisheries in region 2 where the sizes for L-ALL-2 decline but for L-US-2 remain stable. The model is more consistent with the size data for the latter fishery.

7.2.4 Tagging data

The model appeared to fit the overall tagging data very well (Figure 22), and also at the finer scale of the individual tag recapture groupings (Figure 23). However, that good fit is likely exaggerated by the inclusion of returns in the mixing period, which are fitted exactly. This period will be removed for the diagnostic plots of future tuna assessments. The observed and model-predicted recapture numbers show large increases during the three tagging programmes that released the most fish (RTTP; early 1990's, PTTP; 2006–present; Figure 22). There is close concordance between the observed and modelled values for most time-periods with some lack of fit during periods of particularly high recaptures late in the assessment period. The model appears to have some trouble fitting the very high temporal variation in the Coral Sea recaptures which was also a feature of the previous assessment (Group 23; Figure 23). However this is unlikely to impact on the model output given that when these data are excluded from the likelihood (sensitivity model *Mix-CS*), model results are extremely close to the diagnostic case model.

The observed and model-predicted tag returns by time-at-liberty across all tag release events for the diagnostic case model were very similar (Figure 24).

7.3 Model parameter estimates (diagnostic case)

7.3.1 Catchability

Time-series changes in catchability were estimated for most fisheries that did not receive a standardised CPUE index. There is substantial variation in the dynamics of catchability among fisheries (Figure 25), although in many cases this reflects the

sparse nature of the catch and effort time-series for fisheries that were previously, or recently (in the case of fisheries that have substantially contracted their effort, such as several PL fisheries), very low volume fisheries. Perhaps the most important fisheries to examine are the PS fisheries where effort creep is potentially much more likely than some other gear-types. In general, catchability for the high volume PS fisheries has increased over time, however in recent years there is substantial among-fishery variation in trajectories with several fisheries (e.g. those in regions 3 and 4) displaying declining catchability.

7.3.2 Selectivity

A very diverse set of fisheries are included in the assessment model and this is reflected in the form of the estimated selectivity functions (Figure 26). The LL fisheries were all estimated to have selectivity curves that were, or were very close to, asymptotic, even though the only fishery where asymptotic selectivity was enforced was the offshore LL fishery in region 7 (LL OS 7). The large fish handline fishery in region 7 also showed LL-like selectivity with very few fish younger than about 15 quarters of age being estimated to be caught by this fishery.

The PS, PL and miscellaneous small fish fisheries in region 7 displayed a variety of relatively complex selectivity functions (Figure 26), although in general these fisheries tend to catch very few fish over about 20 quarters of age, and some (particularly the PH and ID miscellaneous fisheries) mainly catch fish of the youngest few age classes.

7.3.3 Movement

Figure 27 portrays the origin (which region the fish were originally recruited into) of the equilibrium bigeye biomass in each region. From these it is evident that the model estimates significant movements of fish between the majority of regions. In general, there is a tendency for most regions to receive most of their fish as local recruits (particularly regions 5–9) but in several regions a significant proportion of fish ending up in a region were actually recruited to nearby regions and then moved to their final region via immigration. This latter pattern is particularly the case for regions 2–4. Because of the flexibility of MFCL with regards to assigning recruits to regions and then moving them to other regions to fit data components such as size frequencies and CPUE indices, some caution must be exercised when interpreting the movement rates.

7.3.4 Tag Reporting Rates

The estimated tag reporting rates by fishery recapture groups (see groupings in Table 3) are displayed in Figure 28. As expected, the reporting rate estimates differed among fisheries groups and across tagging programmes. In most cases, the reporting rate estimates for those groupings that received higher penalties were relatively close to the prior mean. As was the case in the 2014 assessment, reporting rate estimates for several groups were on the upper bound of 0.9. Two were miscellaneous fisheries in region 7 that did not recapture any fish for the programmes in question. The other three were the PS fisheries in regions 4 and 8 for the PTTP, and the LL AU fishery in region 9 for the CSTP. When the upper bound on the reporting rate parameter was relaxed to 0.99, only the reporting rate for the PS fishery in region 4 was estimated to be at the new bound.

7.3.5 Growth

The Von Bertalanffy function of the diagnostic case model specifies most rapid growth for the youngest age-classes, starting from a mean length of about 20–30cm for the youngest age-class, before slowing down over older ages with a mean length of 152cm for the oldest fish in the model (age-class 40). The estimated standard deviation of length-at-age increases significantly with age and estimates substantial variation in length-at-age (Figure 29). This is necessary to fit the size frequency data observed for the longline fisheries and is supported by the high apparent variation in growth observed in the otolith dataset investigated external to MFCL.

As indicated earlier, the growth function for the diagnostic case model is substantially different to the old growth model (*L2-184*) and also the growth estimate from the 2014 assessment. For the early age classes they are relatively similar, but after about eight quarters of age the old growth model estimates significantly higher mean length-at-age with substantial differences estimated at the oldest ages.

7.4 Stock assessment results

7.4.1 Recruitment

The estimated distribution of recruitment across regions must be interpreted with some caution, as MFCL has the ability to use a combination of movement and regional recruitment to distribute the population in a way that maximises the total objective function. The diagnostic case model recruitment estimates for each region and the entire assessment domain are shown in Figures 30 and 31. Note that the trend in

recruitments for the overall WCPO in Figure 30 are difficult to interpret owing to the wide confidence intervals displayed, although this trend can be more readily observed in Figure 31. The overall pattern of recruitments is similar to previous assessments: high recruitments in the first few years of the assessment period, particularly in the northern regions, although the uncertainty in the estimates in this period is very high and this is why these recruitments were not included in the estimation of the SRR (Figure 30). There were lower estimated recruitments in the 1960's and 1970's and then slightly increasing recruitments over the remaining period (Figure 31). This increase in recruitments over the assessment period has been a feature of the bigeye assessment for a number of years but is substantially reduced in this assessment.

The estimate of the SRR is presented in Figure 32 and demonstates that recruitments have been maintained at high levels, even as stock size has declined.

Several substantial recruitments are estimated to have occurred over the period around 2011–2014, and this pulse of fish moves through the model age-structure such that the increase in standardised CPUE observed in recent years for several of the LL fisheries can be fitted by the model. This is particularly the case in the equatorial regions 3, 4 and to a lesser degree region 7.

7.4.2   Biomasss

The relative pattern in spawning potential at the regional scale is consistent with the results presented for the 2014 stock assessment (Figure 31 and 33). There was a period of high abundance in the first decade of the assessment period before a period of sustained decline over most of the remaining time-series. There were several short term increases in abundance throughout the assessment period but these were never sustained more than a few years. The decline was estimated to occur in all model regions (Figure 33). Most recently, there has been a significant pulse in spawning potential in the last two years of the assessment, which occurs due to the very high recent recruitments in several regions (Section 7.4.1).

The distribution of fish for the diagnostic case model is slightly different from the 2014 assessment due to the changes in the regional structures (Figure 31). While region 4 still supports a very large biomass of fish, the proportion in this region and in region 3 have been reduced due to their smaller geographical dimensions, while the relative proportions of fish in regions 1 and 2 have increased somewhat.

7.4.3 Fishing mortality

A steady increase in fishing mortality of adult age-classes is estimated to have occurred over most of the assessment period (Figure 34), while juvenile fishing mortality was initially low, before increasing after 1970 with the increase becoming particularly rapid after the expansion of PS fishing in the 1980's. In recent years, the juvenile fishing mortality has been high, but also variable. There appears to be some stabilisation in the fishing mortality rates for both juvenile and adults in recent years, presumably due to the range of management measures implemented and the increase in recruitment and biomass in the last several years.

The fishing mortality-at-age estimates, by region (Figure 35), display intuitive results: the temperate regions where LL fishing dominates show highest fishing mortality for the vulnerable older age classes while in the equatorial regions where the PS and miscellaneous small fish fisheries occur show a high mortality for the youngest age-classes. The trends in fishing mortality-at-age by decade are presented in Figure 36 and show the general increase in fishing mortality of all age-classes over the assessment period, although the increase for the youngest age-classes is more pronounced and leads to an almost bimodal distribution with age-classes between 10 and 20 quarters having the lowest fishing mortality. The ages lower than this are most vulnerable to PS and miscellaneous fishing and the ages older are more vulnerable to LL fishing. Over time there has been a shift in the age distribution of the stock towards younger age-classes as the stock has been fished down (Figure 36).

7.5 Multimodel inference - stepwise model development, sensitivity analyses and structural uncertainty

7.5.1 One-off changes from the structural uncertainty analysis

Comparisons of the spawning potential, and depletion trajectories for the diagnostic case and one-change sensitivity runs from the structural uncertainty analysis are provided in Figures 37–39. The key reference points are compared in Appendix Tables 11–13 (one-off sensitivity models) and Tables 6–10 (structural uncertainty grid), and the likelihood components are provided in Appendix Tables 14–16.

Steepness [*h0.65, h0.95*]

The low penalties on the SRR relationship resulted in the assumed value of steepness having a negligible effect on model fit and time-series estimates of spawning potential (Figure 37). However, slightly different estimates of fisheries depletion were estimated, with the low (*h0.65*) and high (*h0.95*) steepness models suggesting a stock that is

more and less depleted, respectively (Figure 38). This is consistent with previous assessments of tuna in the WCPO - namely that low and high steepness values lead to more pessimistic and optimistic estimates of stock status, respectively. This is particularly the case for MSY and MSY-based reference points ($F_{recent}/F_{\mathrm{MSY}}$ for the diagnostic case, *h0.65* and *h0.95* of 0.67, 0.73 and 0.61, respectively) while the depletion-based reference points tend to be less sensitive to assumed steepness (Appendix Table 11). This is particularly so for the levels of depletion estimated for the diagnostic case and other model runs incorporating the new growth and maturity data – the impact of steepness on estimates of spawning potential at these reduced levels of depletion (around 0.4) is much less than for models with high depletion (<0.2).

Models with different mixing periods [*Mix1, Mix-CS*]

The model with the Coral Sea tag releases assigned an extremely long tag mixing period (*Mix-CS*; 28 quarters) to remove most of their influence from the likelihood, produced virtually indistinguishable results from the diagnostic case model, for estimates of both spawning potential and fisheries depletion (Figures 37 and 38).

The model with the mixing period reduced to one quarter (*Mix1*) from the two quarters in the diagnostic case model, scaled spawning potential downwards slightly (Figure 37) and estimated the stock to be slightly more depleted than the diagnostic case model ($\sim$ 5 percentage points more depleted; Figure 38, Appendix Table 12).

Weighting of the size-frequency data [*Size10, Size50*]

Two alternative models were run with differing weightings given to the length and weight frequency data, one up-weighting the data relative to the diagnostic case (*Size10*; divisor of 10), and the other down-weighting the data relative to the diagnostic case (*Size50*; divisor of 50). These models produced interesting dynamics. Upweighting the size data (*Size10*) scaled the population upwards moderately although the relative changes in the time-series were very similar (Figure 37). The estimates of fisheries depletion were slightly more optimistic than the diagnostic case model, although dynamic terminal depletion was very similar (Figure 38), as were the depletion reference points ($SB_{latest}/SB_{F=0}$, 0.48 vs 0.46; $SB_{recent}/SB_{F=0}$, 0.41 vs 0.38; Appendix Table 12).

The further downweighting (*Size50*) changed the time-series dynamics slightly with a higher estimated spawning potential in the first decades of the model and lower in recent decades (Figure 37), presumably due to it being able to fit the CPUE data slightly better in the very early time period. Consequently, fisheries depletion is estimated to be lower over most of the assessment period but in recent years the estimates of terminal depletion were similar to the diagnostic case model (Figure 38),

as were the depletion reference points ($SB_{latest}/SB_{F=0}$, 0.47 vs 0.46; $SB_{recent}/SB_{F=0}$, 0.39 vs 0.38).

Alternative growth functions [*AL, L2-184*]

Assumptions about the growth function of the bigeye stock had the most impact on model output of all the uncertainties investigated in the 2017 assessment. The old growth model (*L2-184*) which estimated growth in the same manner as the 2014 reference case model, produced substantially more pessimistic model output. Estimates of spawning potential were scaled downwards significantly compared to the diagnostic case model (Figure 37), although the relative fluctuations were broadly similar. An exception to the latter observation is the lack of a recent upturn in spawning potential related to the recent series of high estimated recruitments. This is due to the delayed maturity of fish in the *L2-184* model, such that, although high recruits have also been estimated for this model, they have not yet reached the definition of spawning potential. The estimates of depletion were also more pessimistic for this model, with the terminal estimates of dynamic depletion (Figure 38) and the depletion-based reference points (Appendix Table 11) being about 20 percentage points more pessimistic than the diagnostic case model, though still slightly above the limit reference point.

The alternative growth sensitivity model (*AL*) fitted the new otolith data within the MFCL model and estimated a more similar growth curve to that used in the diagnostic case model. The main differences from the diagnostic case model occurred for the youngest age-classes where *AL* estimated deviations away from the standard VB growth. This model estimated moderately more pessimistic stock status than the diagnostic case, with spawning potential scaled downwards (Figure 37), and fisheries depletion estimated to decline at a faster rate over most of the assessment period. The terminal estimates of dynamic depletion (Figure 38) and the depletion-based reference points (Appendix Table 11) were about 4 percentage points more pessimistic than the diagnostic case model.

2014 regional structure [*2014Reg*]

The assumed regional structure has a significant impact on the model estimates. Using the same MFCL model settings (though adjusted for the different fisheries structures) but input files relating to the 2014 regional structures, this model scaled stock size downwards significantly (Figure 37), although the relative time-series changes in abundance remained similar to the diagnostic case model. The estimates

of fisheries depletion for this model are similar to the diagnostic case model early in the time-series but steadily diverge such that the terminal estimates of dynamic depletion (Figure 38) and the depletion-based reference points (Appendix Table 12) are about 10 percentage points more pessimistic than the diagnostic case model. Our interpretation of this is that the new regional structure more accurately delineates the higher-exploitation tropical area from the lower exploitation sub-tropical area to the north, effectively attributing the higher exploiation to a smaller area of the overall stock, and the lower exploitation to a larger area of the overall stock. This results in lower overall stock-level depletion.

Overdispersion parameter of the negative binomial tagging likelihood [$OD2$]

When considerably less weight was given to the tagging dataset ($OD2$), the estimates of spawning potential were scaled downwards substantially (Figure 37). Similarly, the estimates of fishing depletion for this model were also significantly more pessimistic. Depletion estimates at the beginning of the time-series are below those of the diagnostic case model, and they declined at a greater rate throughout the assessment period, with terminal estimates of dynamic depletion (Figure 38) and the depletion-based reference points (Appendix Table 13) being about 10 percentage points more pessimistic than the diagnostic case model.

Quarterly SRR estimation [SRR-qtr]

The model with the SRR fitted using quarterly (consistent with the 2014 assessment), rather than annual recruitments (assumed by the diagnostic case model), produced almost indistinguishable differences with respect to stock dynamics and population parameters - the time-series of spawning potential is scaled down marginally (Figure 37) and the fisheries depletion is virtually identical to the diagnostic case (Figure 38). Yield for this model was slightly lower than for the diagnostic case model however, with MSY estimated at 156,360 versus 166,480 mt for the diagnostic case model (Appendix Table 13).

Dirichlet multinomial likelihood for the size frequency data [Dch]

The model with the Dirichlet multinomial likelihood for the size frequency data (Dch) estimated the most optimistic stock status with respect to spawning potential (Figure 37) and fishing depletion (Figure 38) of all the one-off sensitivity models. However, the temporal patterns of both these quantities were very similar to the diagnostic case model. The more optimistic outcomes result because the Dirichlet multinomial estimates of effective sample size greatly upweight the weight-frequency data (with ESS for most fishery groups approaching the specified maximum ESS of 100). While the ESS for the length-frequency data are generally downweighted relative to the

settings in the diagnostic case, the overall impact is an upweighting of the size data. As indicated by the likelihood profile (Appendix figure 54), greater weight for the size data is consistent with higher average population biomass and therefore more optimistic stock status.

Old growth with new maturity/natural mortality [*L2-184-NewMat*]

When the old growth model was run with a maturity- and natural mortality-at-age function updated with the data suggesting a younger onset of sexual maturity (*L2-184-NewMat*), the estimates of spawning potential were scaled upwards significantly from the standard old growth model (*L2-184*), and estimates were roughly half way between that model and the diagnostic case model (Figure 37). With this model there also becomes evidence for the very recent upturn in spawning potential related to the recent high recruitment events. These were previously not included in the definition of spawning potential with the later maturity assumed by model *L2-184*. However, the trajectory of fisheries depletion for this sensitivity remained very similar to model *L2-184*, with only very minor differences in temporal variation over the time-series (Figure 38). Apparently the inclusion of both more young fish and more older fish in the spawning potential when using the new maturity schedule (Figure 11) has an offsetting effect in terms of fisheries depletion.

Estimate natural mortality *M*-at-age [*Lorenzen*]

The *M*-at-age function estimated using the Lorenzen parameterisation was very consistent with the fixed values used in the other models (Figure 10). Because of this, most outputs for this model were broadly similar with the diagnostic case model, with spawning potential estimated to be slightly higher earlier in the time-series but very similar in recent years (Figure 37). In contrast, estimates of fisheries depletion were more pessimistic for this model, with terminal estimates of dynamic depletion (Figure 38) and the depletion-based reference points (Appendix Table 11) being about 5–10 percentage points more pessimistic than the diagnostic case model.

Alternative standardised CPUE indices [*CPUE-Proxy, CPUE-Geostat*]

The model with the alternative GLM-based CPUE indices (*CPUE-Proxy*) produced estimates of spawning potential that showed some minor differences in variation across the time-series compared to the diagnostic case model, but the scale of the estimates was very similar (Figure 37). The estimates of fishing depletion were very similar for most of the assessment period but became more pessimistic than the diagnostic case model in about the last 10 years (Figure 38).

The model with the geostatistical CPUE indices (*CPUE-Geostat*) estimated a moderately lower spawning potential over the entire assessment period (Figure 37), and

the trajectory of fishing depletion (Figure 38) was also more pessimistic than the diagnostic case model, with depletion occurring at a higher rate over most of the time-series. Terminal estimates of dynamic depletion (Figure 38) and the depletion-based reference points (Appendix Table 13) were about 5–10 percentage points more pessimistic than the diagnostic case model.

Inclusion of tags from the Japanese Tagging Programme [*JP*]

When JPTP tags were included in the model the resulting estimates of spawning potential showed very similar temporal variation but the time-series was scaled downwards slightly (Figure 37). The estimates of fishing depletion were also very similar for most of the assessment period but gradually became more pessimistic than the diagnostic case model, with terminal depletion estimates about 4 percentage points lower than the diagnostic case model in the terminal years (Figure 38).

Relaxation of reporting rate parameter bounds [*RR0.99*]

This sensitivity model scaled spawning potential upwards marginally in the first decade of the assessment period, but the trajectories over the remainder of the period were virtually indistinguishable from the the diagnostic case model (Figure 37). The estimates of fisheries depletion over the entire time-series are indistinguishable between *RR0.99* and the diagnostic case model (Figure 38).

Uncertainty in longline catch reporting [*C-Uncert*]

The sensitivity model that assumed an alternative catch history with higher catches for the longline fisheries than have been reported, scaled the spawning potential upwards slightly, mainly in the early part of the assessment period (Figure 37). By the end of the time-series, where the level of catch deviates most from the diagnostic case model, spawning potential was estimated to be slightly lower than the diagnostic case model. This resulted in estimates of fisheries depletion that were very similar over most of the time-series but become more pessimistic following the assumed increase in under reported catch in the mid-late 2000's, with terminal depletion estimates about 3 percentage points lower than the diagnostic case model in the terminal years (Figure 38).

7.5.2   Structural uncertainty analysis

The results of the structural uncertainty analysis are summarised in several forms – time-series plots of fisheries depletion for all models in the grid (Figure 39), boxplots of F_{recent}/F_{MSY} and $SB_{latest}/SB_{F=0}$ for the different levels of each of the five axes of uncertainty (Figure 40), Majuro plots showing the estimates of F_{recent}/F_{MSY} and $SB_{latest}/SB_{F=0}$ (and $SB_{recent}/SB_{F=0}$ for comparison) across all models in the grid

(Figure 41), and averages and quantiles across the full grid of 48 models (and specific subsets of models) for all the reference points and other quantities of interest (Tables 6–10) that have also been presented for the diagnostic case model and one-off sensitivity models.

Many of the results of the structural uncertainty analysis are consistent with the results of previous assessments of tuna stocks in the WCPO that used the same uncertainty axes. However, additional axes have been included in the 2017 assessment and these have significant consequences for summarising stock status and deriving management recommendations.

The general features of the structural uncertainty analysis are as follows:

– The grid contains a wide range of models with significantly different estimates of stock status, trends in abundance and reference points. The uncertainty identified is higher than previous assessments of bigeye in the WCPO.

– The most influential axis with respect to model output is growth, with the grid being able to be separated into two discrete groups of models depending on whether the new growth of the diagnostic case model (optimistic) or the old growth (L2 = 184cm) used in 2014 (pessimistic) is assumed (Figure 39).

– The subset of new growth models estimate fisheries depletion to be well above $20\%SB_{F=0}$, with individual model runs estimating depletion at the end of the assessment period being largely between about 0.25 and 0.50, depending on whether dynamic depletion (Figure 39), or the reference points $SB_{latest}/SB_{F=0}$ or $SB_{recent}/SB_{F=0}$ (Tables 7–8) are used.

– The subset of old growth models are substantially more pessimistic with an estimated depletion at the end of the assessment period being largely between about 0.08 and 0.26, depending on whether dynamic depletion (Figure 39), or the reference points $SB_{latest}/SB_{F=0}$ or $SB_{recent}/SB_{F=0}$ (Tables 9–10) are used. This set therefore falls either side of the limit reference point, $20\%SB_{F=0}$.

– The next most influential axis was the assumed regional structure and it effectively further separates the growth subsets such that there are four relatively mutually exclusive groups of models (Figure 39) which, in order from optimistic to pessimistic, are; new growth/2017 regions, new growth/2014 regions, old growth/2017 regions and old growth/2014 regions.

– The steepness axis displayed largely predictable results, with steepness of 0.65 and 0.95 producing more pessimistic and optimistic estimates, respectively, than the 0.8 assumed in the diagnostic case model. The lower the steepness the more depleted the stock, and the higher the fishing mortality with respect to fishing

mortality at MSY.

– The remaining two axes, weighting of the size frequency data and weighting of the tagging data, produced further uncertainty in model output, but their impacts were much reduced in comparison to growth and the regional structure (Figure 39; Tables 6–10). Downweighting the size frequencies led to generally more pessimistic results than upweighting (relative to the settings in the diagnostic case model). Similarly, downweighting the tagging data produced generally more pessimistic results but the differences were not substantial (Figure 39).

– For all models with new growth there is a recent upturn in spawning potential and consequently fisheries depletion (Figure 39). This results in the estimates of $SB_{latest}/SB_{F=0}$ for these models being higher than the fully dynamic depletion ratio, because we are dividing a single, more optimistic spawning potential (2015), by a series of less optimistic spawning potentials for an unfished stock (2005–2014). Therefore, if the reference point $SB_{recent}/SB_{F=0}$ is plotted as well (blue points, panel 2; Figure 41) the general, curved relationship between depletion and the fishing mortality reference point is more consistent between the growth assumptions.

7.5.3 Further analyses of stock status

There are several ancillary analyses related to stock status that are typically undertaken on the reference case model (dynamic Majuro analyses, fisheries impacts analyses etc.). The shift towards relying more on multimodel inference for the 2017 assessment makes it more difficult to present these results over a large number of model runs. In this section we rely heavily on the tabular results of the structural uncertainty grid (Tables 6–10), and where space allows, diagnostic plots are presented for example models from the grid with different estimates of stock status.

Fishery impacts for example models

We measure fishery impact at each time step as the ratio of the estimated spawning potential relative to the spawning potential that is estimated to have occurred in the historical absence of fishing. This is a useful quantity to monitor, as it can be computed both at the region level, and for the WCPO as a whole. This information is plotted in two ways, firstly as the fished and unfished spawning potential trajectories (Figures 42–43), and secondly as the depletion ratios themselves (Figure 44–45). The latter is relevant for the agreed reference points and are discussed in more detail in the following sections. Example plots of these values are displayed for the diagnostic case model and the old growth sensitivity model (*L2-184*).

The diagnostic case model estimated that steady declines in spawning potential have occurred in all regions (Figure 44), although the rate and extent of decline differs among regions. The equatorial regions (3, 4 and 7) tend to show more rapid decline than the more temperate regions, with most decline attributed to fishing impacts. By contrast, most of the declines in spawning potential in the large southern regions (5 and 6) are estimated to have occurred due to declining recruitment, rather than the impacts of fishing in these regions. The results for *L2-184* show the same general patterns except that the difference between the fished and unfished trajectories is higher, which may partly be attributed to the increased trend in recruitments over the time-series for this model. These patterns are also reflected in the depletion ratio plots (Figures 44–45), with similar sustained declines in all regions for both models, although the magnitude of the declines is more pronounced for *L2-184*.

It is also possible to attribute the fishery impact with respect to depletion levels to specific fishery components (grouped by gear-type), in order to estimate which types of fishing activity have the most impact on spawning potential (Figure 46–47). While the overall depletion levels of the 2017 diagnostic case model are significantly lower than the 2014 reference case model, the relative contributions of different fishing gears are very similar (Figure 46). The early impacts on the population were primarily driven by LL fishing, but as the PS fishery expanded from the 1980's the impact of associated PS fishing has steadily increased, such that it has led to that gear having more impact than LL fishing in all the equatorial regions. The substantial increase in fishing effort by miscellaneous gears in region 7 has also had a significant impact on spawning potential in that region, and also in neighbouring regions that are linked by movement. Unassociated PS fishing has had little impact due to far fewer bigeye being caught in these sets. Similar results were estimated using *L2-184* (Figure 47), with the main difference being the significantly higher overall fisheries impacts for this model. There was also a tendency for the impacts of PS fishing on biomass in the northern regions to be lower for this model, which can be related to different movement dynamics among regions between the models.

Yield analysis and equilibrium estimates across the grid

The yield analyses conducted in this assessment incorporate the spawner recruitment relationship (Figure 32) into the equilibrium biomass and yield computations. Importantly, in the diagnostic case model, the steepness of the SRR was fixed at 0.8 so only the scaling parameter was estimated. Other models in the one-off sensitivity analyses and structural uncertainty analyses assumed steepness values of 0.65 and 0.95.

Across the structural uncertainty grid the equilibrium virgin spawning potential in the absence of fishing ($SB_{F=0}$) was estimated to be between 1,317,335 and 2,460,410 mt (Table 6), and the spawning potential that would support the MSY (SB_{MSY}) was

estimated to be between 219,500 and 710,000 mt. The ratio of SB_{MSY} to SB_0 was estimated to be between 0.22 and 0.29 (mean = 0.25). The ratio of SB_{MSY} to $SB_{F=0}$ was estimated to be between 0.17 and 0.29 (mean = 0.22).

A plot of the yield distribution under different values of fishing effort relative to the current effort are shown in Figure 48, for example models from the four main subsets of structural uncertainty grid models. These support the general observations for other model output. That is, models with new growth/2017 regions estimate that MSY would be achieved with moderately higher fishing effort, new growth/2014 regions and old growth/2017 regions indicate that current effort is close to that giving MSY, and old growth/2014 regions suggest that fishing effort should be reduced to acheive MSY.

The yield analysis also enables an assessment of the MSY level that would be theoretically achievable under the different patterns of age-specific fishing mortality observed through the history of the fishery (Figures 49–50). Again we present two example plots for the diagnostic case model and the old growth sensitivity model, *L2-184*. Prior to 1970, the WCPO bigeye fishery was almost exclusively conducted using LL gear, with a low exploitation of small bigeye. The associated age-specific selectivity pattern resulted in a much higher MSY in the early period compared to the recent estimates. This pronounced decline occurred after the expansion of the small-fish fisheries in region 7 and, soon after, the rapid expansion of the PS fishery which shifted the age composition of the catch dramatically towards much younger fish. These patterns were common to both example models, although the absolute estimates of MSY were slightly lower for model *L2-184*.

Dynamic Majuro plots and comparisons with Limit Reference Points

The section summarising the structural uncertainty grid (Section 7.5.2) presents terminal estimates of stock status in the form of Majuro plots. Further analyses can estimate the time-series of stock status in the form of Majuro plots, the methods of which are presented in Section 5.7.4. The large number of model runs in the structural uncertainty grid precludes undertaking and presenting this process for all runs, however an example from each of the four subsets of model runs (combinations of new and old growth, and 2017 and 2014 regional structure) in the grid is presented in Figure 51. These examples produce intuitive results with respect to the terminal results already presented in Section 7.5.2. All models at the start of the assessment period were close to an $SB/SB_{F=0}$ of one and an F/F_{msy} approaching zero, but each progressively shifted towards the overfishing and overfished definitions over the remaining period. The diagnostic case model (representative of new growth/2017 regions) never reaches $20\%SB_{F=0}$ or an F/F_{msy} of one. Both the sensitivity models (*L2-184, 2014Reg*) representing old growth/2017 regions and new growth/2014 regions,

respectively, move into the overfishing region of the plot but do not reach an overfished state. The grid model (*A0B1C0D0E1*; old growth/2014 regions) progressively shifted into the overfishing region in the early 2000's and then progressed into the overfished state for the last decade of the assessment period. The equivalent dynamic Kobe plots are displayed in Figure 52.

7.5.4 Key differences between new and old growth models

One of the major influences of stock status and other model output in the current assessment is the assumed growth function. The new and old growth models result in optimistic and pessimistic estimates, respectively. Both models are able to fit the various data components relatively well, including the size frequency data, but the manner in which they do so is quite different. The new growth model estimates that the average asymptotic size of the oldest fish is relatively low (Figure 29), but there is substantial variation around the growth at old ages, and so even the largest fish caught by the LL fisheries can be predicted by the model (Figures 18–19), although it should be noted that the proportion of fish caught at these large sizes is very low. This model suggests exactly this, there are very few fish in the population at these sizes and the LL fisheries fully select these fish. This is reflected in the asymptotic selectivity functions estimated for most LL fisheries (Figure 26).

In contrast, the old growth function estimates a much higher asymptotic size for the oldest fish and the estimated variation around the length-at-age is less than for the diagnostic case model. This model effectively suggests that there would be a high number of very large fish in the population in the absence of fishing, but because proportionally few at this size are caught, even by the LL fisheries, then most fish are not reaching this size due to fishing mortality. This is why the depletion for this model is considerably higher, even in the early years when fishing mortality was relatively low. This model also estimates that most LL fisheries do not fully select these very old/large fish, which is due to few fish of this size evident in the observed size frequencies and results in estimates of selectivity functions for the LL fisheries usually having a significant declining right-hand limb, which was also observed in the 2014 assessment (Figure 53). This result conflicts with the usual assumptions about LL fishing which are generally thought to fully select the old/large fish in the population, although in some cases it could be interpreted as the fisheries operating in areas of a region where the oldest fish are either absent or unavailable (e.g. if the oldest fish show a preference for deeper habitat where they are less vulnerable to LL gear). However, the fact that the selectivities have this feature across nearly all LL fisheries and regions, even those regions where the thermocline depths are known to be relatively shallow (e.g. region 4), does question the reliability of this result.

Another important factor influencing the estimates of spawning potential and fisheries depletion between the old and new growth models is the definition of spawning potential. For the new growth model, the spawning potential function is updated for new maturity estimates which shifts the spawning potential function several age-classes younger (Figure 11). This definition therefore includes a significant number of extra fish in the spawning population which is evident in the higher estimates of spawning potential for the new growth models. When the sensitivity model *L2-184-NewMat* was constructed with the new maturity estimates, these younger fish are then also included in the spawning potential and the estimates are scaled up considerably compared to model *L2-184* (Figure 37). The fisheries depletion for these two models remained similar however, probably due to more fish in the older age-classes also being included in the spawning definition of *L2-184-NewMat* (Figure 11). These older age-classes become more depleted as the stock is fished down and so appear to offset the less depleted younger age-classes that were also included in the spawning potential for this model.

8 Discussion and conclusions

8.1 Changes to the previous assessment

The 2017 stock assessment introduces a number of changes from the 2014 assessment that have had a large influence on the resulting estimates of stock status. Three additional years of data (tagging, catch, effort, size frequencies) were included in the assessment, covering a period of strong El Nino conditions, and a period of catch stabilisation. Within this period there has been a significant increase in several of the standardised CPUE indices in the terminal years of the model. The model attributes this to a period of high recruitments several years before the upturn in the CPUE (which indexes the abundance of older fish vulnerable to longline gear). This has also resulted in an increase in stock status indicators in the last several years. A similar pattern has recently been observed in the EPO, with IATTC estimating a very high recruitment event in 2012 which resulted in an increase in their spawner biomass ratio from the historically low value of 0.16 in 2013 to 0.26 at the beginning of 2017 (Aires-da Silva et al., 2017). This also reflects anecdotal evidence from longline fisheries in parts of the WCPO, for which increased catch rates of bigeye have been reported. Of particular interest over the coming years will be whether this increase in CPUE is maintained, whether further individuals enter the age-classes vulnerable to longline fishing after the current assessment period (and thus influence future measures of stock status) and whether these high recruitments can be attributed to recent climatic conditions and/or recent patterns in fishing mortality and conservation

and management measures.

Other changes made to the model included implementing minor developments to MFCL that have become available since the 2014 assessment. These included developments in the modelling of recruitment (annual SRR, arithmetic rather than geometric mean of other recruitments), a trial of the Dirichlet multinomial likelihood for modelling the size frequency data and estimation of a Lorenzen-type relationship between natural mortality and the size of fish. These implementations did not have significant consequences for the estimates of stock status for the current assessment, especially compared to some of the other uncertainties investigated in alternative model runs.

Undoubtedly the most important changes from the 2014 assessment were the investigation of alternative growth functions (and associated age-specific natural mortality and maturity) and the use of alternative spatial structures in the 2017 assessment. These topics will now be discussed in detail in Sections 8.2 and 8.3.

8.2 Investigation of alternative growth and maturity functions

The most important factor driving changes in the 2017 stock assessment concerned growth. Very large differences in estimated stock status resulted from growth being assumed to be similar to that used in the 2014 assessment, or whether it is accurately represented by the new otolith dataset presented by Farley et al. (2017). This dataset suggests a significantly lower asymptotic size for old fish in the stock. The growth axis of the structural uncertainty grid (and to a lesser degree the regional structure axis) effectively separate model runs into discrete categories with respect to model output and stock status. All models with the new growth function suggest significantly more optimistic stock status than the 2014 assessment with all estimates of $SB_{latest}/SB_{F=0}$, $SB_{recent}/SB_{F=0}$ and dynamic depletion being well above those estimated in the 2014 assessment and also well above the agreed limit reference point $20\%SB_{F=0}$.

The Scientific Committee must assess the evidence to support each of the two divergent growth options as this will have a strong impact on the management recommendations that can be formulated from this assessment. If recommendations are to be made based solely on models with the new growth function then the bigeye stock in the WCPO appears not to be in an overfished state although several model runs are at the boundary of the designation of overfishing. However, even for most these models, the stock is estimated to have declined substantially over most of the assessment period and remains in a state somewhat more depleted than most other tuna stocks in the WCPO.

If models with the old growth are considered for formulating recommendations, then estimates of stock status are less clear cut. The old growth models with the 2017 regional structure are more optimistic than the 2014 assessment results, with most the model runs in this category terminating around, or above the limit reference point, $20\%SB_{F=0}$. In contrast the same models but assuming the 2014 regional structure are more pessimistic and are generally more consistent with the 2014 assessment results. All models in this set terminate below the limit reference point.

Given the critical nature of the growth assumptions on the assessment, it may be worthwhile to consider further investment in research to improve the existing growth data set and its analysis, including:

– **The collection and analysis of additional otoliths of larger bigeye tuna, > 140cm**. This would improve the estimates of the age distribution of these large fish, which seems to be the key determinant of the impacts on stock status. It would also allow better estimation of the variation of length-at-age for the oldest fish in the population.

– **The continued development of a high-confidence tagging data set for growth analysis, with particular focus on larger bigeye tuna**. Such data would assist with the validation of the age estimates of large bigeye in the WCPO, and could possibly be incorporated directly into the assessment model as a supporting growth data set.

∗ **Collaboration with the IATTC to analyse bigeye growth from otolith and tagging data collected across the entire Pacific, to better characterise the apparent regional differences in growth between the WCPO and EPO and possible environmental determinants of such differences.**

– **The analysis of the same otoliths by different laboratories, to build confidence in the aging estimates and estimate aging errors.**

8.3 Investigation of a modified regional structure

A recommendation of the PAW in April 2017 (Pilling and Brouwer, 2017) was to investigate a new regional structure to better reflect the limited movement between the equatorial zone (10° S–10° N) and more temperate areas that has been observed for fish tagged throughout the Pacific (Schaefer et al., 2015). This structure was also proposed as a means to better reflect the spatial division of the equatorial purse seine fishing zone and areas where fishing is more dominated by longline fishing. The assumed regional structure has a substantial impact on estimated stock status which was observed in the stepwise model development, the one-off sensitivity model runs

and when the two regional structures were included in the structural uncertainty grid (the second most important axis after growth). The 2017 regional structure led to significantly more optimistic results, which can be attributed to the model redistributing the stock biomass from the more heavily exploited equatorial regions to the northern temperate regions, which are less exploited. Consequently, the depletion trajectories for individual regions for the 2014 and 2017 regional structures are relatively similar among regions (higher in equatorial regions, lower in northern temperate regions), though because more of the population is assumed to be present in the northern regions under the 2017 regional structure, the net depletion estimated for the WCPO is lower.

It is important to note that the opposite phenomenon occurred when the regional structure was modified between the 2004 and 2005 bigeye stock assessments (Hampton et al., 2005). That is, the 2005 assessment shifted biomass towards the more exploited equatorial regions and the stock status was consequently estimated to be significantly more pessimistic.

8.4 MFCL and other modelling considerations

8.4.1 Maturity and natural mortality-at-age

The investigation of alternative growth functions forced consideration of the associated fixed vectors of natural mortality- and spawning potential-at-age (McKechnie et al., 2017b). The latter is often considered to be a relatively unimportant biological parameter, with the main influence being through the use of spawning biomass (in the case of bigeye - spawning potential) in the estimation of the SRR. However, this assessment demonstrated the importance of this function when many of the estimates of stock status rely on definitions of spawning potential and its comparison with spawning potential in the absence of fishing ($SB_{F=0}$). Shifts in spawning potential-at-age to include younger or older age-classes can significantly affect the scale of estimates of spawning potential and depletion, because depletion is always higher for older age-classes as a stock is fished down.

Given the importance of these parameters, **further biological studies on the components (e.g. egg production-at-age) would be beneficial in reducing uncertainty in the values used in assessments.** Additional analyses investigating the determinants of the observed shifts in sex ratio of the catch with size would also be useful. For example, the spatial, biological and fisheries factors influencing biased sex ratios in the catch of bigeye by longline vessels would help improve estimates of natural mortality-at-age, which is also used in the calculation of maturity-at-age

functions (McKechnie et al., 2017b).

A related issue is that the current parameterisation of MFCL is based on natural mortality- and maturity-at-age which requires analyses to be undertaken whenever the growth function is modified to make sure all the biological parameters are consistent. This is because many of the component factors are calculated at-length and must be backtransformed to at-age using the desired growth function. This adds considerable work and bookkeeping to the assessment to make sure consistency is maintained. **Re-parameterising MFCL to allow maturity-at-length for example would ameliorate this problem, with the backtransformation from at-length to at-age occurring naturally within the model itself.**

8.4.2   Caution in the use of $SB_{latest}/SB_{F=0}$

The 2017 stock assessment of bigeye raised similar issues about the use of the reference point $SB_{latest}/SB_{F=0}$ that were previously noted in the 2016 assessment of skipjack in the WCPO (McKechnie et al., 2016a). The problem arises when the estimated spawning potential changes significantly in the terminal year/s of the assessment, usually due to a significant recruitment event. The reference point $SB_{latest}/SB_{F=0}$ is therefore based on a numerator that is very high (the last year/s of the assessment) and a denominator that includes many values that were part of a period of lower unexploited biomass (2005–2014). This leads to a high reference point, with significant deviation from the fully dynamic estimate of depletion. This is particularly the case when comparing models assuming the old and new growth functions, with only the latter models estimating the high recent recruitments being part of the spawning potential in the terminal years of the assessment. Consequently, comparison of $SB_{latest}/SB_{F=0}$ between the model sets tends to overestimate the difference in these reference points, given that in the coming years the high recruitments will also progress to the spawning potential in the old growth models as well. We therefore recommend considering additional reference points ($SB_{recent}/SB_{F=0}$ and fully dynamic depletion estimates) when assessing the recent stock status from these models.

8.4.3   Issues with the weighting of data components

This assessment is the first for bigeye in the WCPO that quantifies the influence of the different data components on the scaling of the biomass estimates. The new method of conducting likelihood profiles with respect to the derived parameter of mean total biomass has proven to be successful, and without some of the problems associated with more traditional profiles on the population scaling parameter (total population scaling parameter in MFCL). However, the results indicate conflict between the various data

components. Now that these conflicts have been assessed, a major task in the intersessional period will be investigating means to reduce conflict and/or develop further methods of objectively weighting the various data components. This is an area of ongoing research for integrated fisheries models particularly for tagging data, for which methods are still in their infancy. It should be noted however that we have investigated a range of weightings of the components as part of the one-off sensitivity analyses and the structural uncertainty grid, and the estimates of stock status were proven to be robust to the different assumptions tested. Of particular promise is the Dirichlet-multinomial, and similar self-scaling likelihoods for compositional data, which have the potential to estimate relatively objective effective sample sizes for length- and weght-frequency data, and potentially conditional age-at-length data. The Dirichlet-multinomial and another self-scaling multinomial method that models autocorrelation in observations through random effects, have been recently coded in MFCL. Following further testing and evaluation, it is likely that one of these approaches will be used as a future standard for modelling compositional data in MFCL.

8.5 Main assessment conclusions

The main conclusions of the current assessment are more difficult to construct than in previous bigeye assessments. The Scientific Committee will have to assess the plausibility of the different models in the structural uncertainty grid, particularly four groups of models resulting from different combinations of new and old growth/maturity, and the 2017 and 2014 regional structure. To this end, we summarise the general conclusions of this assessment as follows:

– All models that assume the new growth function estimate significantly more optimistic stock status than the 2014 assessment, with the stock above $20\%SB_{F=0}$ in all cases.

– All models with the new growth estimate a significant recent recruitment event that has increased spawning potential in the last several years, and it is expected that for the old growth models these recruits will soon progress into the spawning potential and increase stock status, at least in the short-term.

– Of the four sets of models in the structural uncertainty grid (the combinations of old/new growth and 2017/2014 regions), only the old growth/2014 regions models estimate spawning potential to be below $20\%SB_{F=0}$ for all models in the set. These models estimate $SB_{latest}/SB_{F=0}$ to be between 0.08 and 0.17 which is slightly more pessimistic than the structural uncertainty grid of the 2014 assessment (between 0.1 and 0.2).

– A substantial decline in bigeye abundance was estimated by all models in the assessment and recent estimates of depletion with respect to estimates earlier in the assessment period, and with respect to estimates in the absence of fishing are significant and appear to be ongoing, at least on a multi-year scale.

– The significance of the recent high recruitment events and the progression of these fish to the spawning potential component of the stock are encouraging, although whether this is a result of management measures for the fishery or beneficial environmental conditions is currently unclear. It is noteworthy, however, that recent favourable recruitment events have also been estimated for skipjack (McKechnie et al., 2016a) and yellowfin (Tremblay-Boyer et al., 2017) in the WCPO, and bigeye in the EPO (Aires-da Silva et al., 2017), which may give weight to the favourable environmental conditions hypothesis. Whether these trends are maintained in coming years will help tease these factors apart and will likely provide more certainty about the future trajectories of the stock.

Acknowledgements

References

Abascal, F., Lawson, T., and Williams, P. (2014). Analysis of purse seine size data for skipjack, bigeye and yellowfin tunas. WCPFC-SC10-2014/SA-IP-05, Majuro, Republic of the Marshall Islands, 6–14 August 2014.

Aires-da Silva, A., Minte-Vera, C., and Maunder, M. N. (2017). Status of bigeye tuna in the eastern Pacific Ocean in 2016 and outlook for the future. IATTC, Scientific Advisory Committee, Eighth Meeting, SAC-08-04a.

Aires-da-Silva, A. M., Maunder, M. N., Schaefer, K. M., and Fuller, D. W. (2015). Improved growth estimates from integrated analysis of direct aging and tag-recapture data: An illustration with bigeye tuna (*Thunnus obesus*) of the eastern pacific ocean with implications for management. *Fisheries Research*, 163:119–126.

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.

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.

Cordue, P. L. (2013). Review of species and size composition estimation for the western and central pacific purse seine fishery. WCPFC-SC9-2013/ST-IP-02, Pohnpei, Federated States of Micronesia, 6–14 August 2013.

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.

Davies, N., Fournier, D. A., Hampton, J., Hoyle, S., Bouye, F., and Harley, S. (2017). Recent developments in the MULTIFAN-CL stock assessment software. WCPFC-SC13-2017/SA-IP-05, Rarotonga, Cook Islands, 9–17 August 2017.

Davies, N., Hoyle, S., Harley, S., Langley, A., Kleiber, P., and Hampton, J. (2011). Stock assessment of bigeye tuna in the western and central Pacific Ocean. WCPFC-SC7-2011/SA-WP-02, Pohnpei, Federated States of Micronesia, 9–17 August 2011.

Evans, K., Langley, A., Clear, N., Williams, P., Patterson, T., Sibert, J., Hampton, J., and Gunn, J. (2008). Behaviour and habitat preferences of bigeye tuna (*Thunnus*

obesus) and their influence on longline fishery catches in the western coral sea. *Canadian Journal of Fisheries and Aquatic Sciences*, 65:2427–2443.

Farley, J., Eveson, P., Krusic-Golub, K., Sanchez, C., Roupsard, F., McKechnie, S., Nichol, S., Leroy, B., Smith, N., and Chang, S.-K. (2017). Age, growth and maturity of bigeye tuna in the western and central Pacific Ocean. WCPFC-SC13-2017/SA-WP-01, Rarotonga, Cook Islands, 9–17 August 2017.

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. and Archibald, C. P. (1982). A general theory for analyzing catch at age data. *Canadian Journal of Fisheries and Aquatic Science*, 39.

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. (1992). Use of risk analysis to assess fishery management strategies: A case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. *Canadian Journal of Fisheries and Aquatic Science*, 49:922–930.

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.).

Grewe, P. and Hampton, J. (1998). An assessment of bigeye (thunnus obesus) population structure in the pacific ocean based on mitochondrial dna and dna microsatellite analysis. Technical report, JIMAR Contribution 98-330.

Hampton, J. (2000). Natural mortality rates in tropical tunas: size really does matter. *Canadian Journal of Fisheries and Aquatic Sciences*, 57:1002–1010.

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.

Hampton, J., Kleiber, P., Langley, A., Takeuchi, Y., and Ichinokawa, M. (2005). Stock assessment of bigeye tuna in the western and central Pacific Ocean. WCPFC-SC1-2005/SA-WP-02, Noumea, New Caledonia, 8–19 August 2005.

Hampton, J. and Williams, P. (2005). A description of tag-recapture data for bigeye tuna in the western and central pacific ocean. *Col. Vol. Sci. Pap. ICCAT*, 57(2):85–93.

Hampton, J. and Williams, P. (2017). Annual estimates of purse seine catches by species based on alternative data sources. WCPFC-SC13-2017/SA-IP-03, Rarotonga, Cook Islands, 9–17 August 2017.

Harley, S. J. (2011). A preliminary investigation of steepness in tunas based on stock assessment results. WCPFC-SC7-2011/SA-IP-08, Pohnpei, Federated States of Micronesia, 9–17 August 2011.

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.

Harley, S. J., Davies, N., Tremblay-Boyer, L., Hampton, J., and McKechnie, S. (2015). Stock assessment of south Pacific albacore tuna. WCPFC-SC11-2015/SA-WP-06, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Harley, S. J., Hoyle, S., Hampton, J., and Kleiber, P. (2010). Stock assessment of bigeye tuna in the western and central pacific ocean. WCPFC-SC6-2010/SA-WP-01, Nuku'alofa, Tonga, 10–19 August 2010.

Harley, S. J., Hoyle, S., Langley, A., Hampton, J., and Kleiber, P. (2009). Stock assessment of bigeye tuna in the Western and Central Pacific Ocean. WCPFC-SC5-2009/SA-WP-04, Port Vila, Vanuatu, 10–21 August 2009.

Harley, S. J. and Maunder, M. N. (2003). A simple model for age-structured natural mortality based on changes in sex ratios. Technical Report SAR-4-01, Inter-American Tropical Tuna Commission, La Jolla, California, USA, 19–21 May 2003.

Hoyle, S., Langley, A. D., and Hampton, J. (2008). 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. and Nicol, S. (2008). Sensitivity of bigeye stock assessment to alternative biological and reproductive assumptions. WCPFC-SC4-2008/ME-WP-01, Port Moresby, Papua New Guinea, 11–22 August 2008.

Hoyle, S. D. (2008). Adjusted biological parameters and spawning biomass calculations for south Pacific albacore tuna, and their implications for stock assessments. WCPFC-SC4-2008/ME-WP-02, Port Moresby, Papua New Guinea, 11–22 August 2008.

Hoyle, S. D. and Okamoto, H. (2011). Analyses of Japanese longline operational catch and effort for Bigeye and Yellowfin Tuna in the WCPO. WCPFC-SC7-2011/SA-IP-01, Pohnpei, Federated States of Micronesia, 9–17 August 2011.

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.

ISSF (2011). Report of the 2011 ISSF stock assessment workshop. Technical Report ISSF Technical Report 2011-02, Rome, Italy, March 14–17.

Itano, D. (2000). The reproductive biology of yellowfin tuna (*Thunnus albacares*) in Hawaiian waters and the western tropical Pacific Ocean: Project summary. JIMAR Contribution 00-328 SOEST 00-01.

Kleiber, P., Fournier, D., Hampton, J., Davies, N., Bouye, F., and Hoyle, S. (2017). *MULTIFAN-CL User's Guide*. http://www.multifan-cl.org/.

Langley, A., Hampton, J., Kleiber, P., and Hoyle, S. (2008). Stock assessment of bigeye tuna in the western and central pacific ocean, including an analysis of management options. WCPFC-SC4/SA-WP-1.

Langley, A., Okamoto, H., Williams, P., Miyabe, N., and Bigelow, K. (2006). A summary of the data available for the estimation of conversion factors (processed to whole fish weights) for yellowfin and bigeye tuna. WCPFC-SC2-2006/ME-IP-03, Manila, Philippines, 7–18 August 2006.

Lawson, T. (2013). Update on the estimation of the species composition of the catch by purse seiners in the Western and Central Pacific Ocean, with responses to recent independent reviews. WCPFC-SC9-2013/ST-WP-03, Pohnpei, Federated States of Micronesia, 6–14 August 2013.

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.

Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystem and aquaculture. *Journal of Fish Biology*, 42:627–647.

McArdle, B. (2013). To improve the estimation of species and size composition estimation of the western and central Pacific purse seine fishery from observer-based sampling of the catch. WCPFC-SC9-2013/ST-IP-04, Pohnpei, Federated States of Micronesia, 6–14 August 2013.

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. (2015a). 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.

McKechnie, S., Hampton, J., Pilling, G. M., and Davies, N. (2016a). Stock assessment of skipjack tuna in the western and central Pacific Ocean. WCPFC-SC12-2016/SA-WP-04, Bali, Indonesia, 3–11 August 2016.

McKechnie, S., Harley, S. J., Davies, N., Rice, J., Hampton, J., and Berger, A. (2014a). Basis for regional structures used in the 2014 tropical tuna assessments, including regional weights. WCPFC-SC10-2014/SA-IP-02, Majuro, Republic of the Marshall Islands, 6–14 August 2014.

McKechnie, S., Harley, S. J. Chang, S.-K., Liu, H.-I., and Yuan, T.-L. (2014b). Analysis of longline catch per unit effort data for bigeye and yellowfin tunas. WCPFC-SC10-2014/SA-IP-03, Majuro, Republic of the Marshall Islands, 6–14 August 2014.

McKechnie, S., Ochi, D., Kiyofuji, H., Peatman, T., and Caillot, S. (2016b). Construction of tagging data input files for the 2016 skipjack tuna stock assessment in the western and central Pacific Ocean. WCPFC-SC12-2016/SA-IP-05, Bali, Indonesia, 3–11 August 2016.

McKechnie, S., Pilling, G., and Hampton, J. (2017a). Stock assessment of bigeye tuna in the western and central Pacific Ocean. WCPFC-SC13-2017/SA-WP-05, Rarotonga, Cook Islands, 9–17 August 2017.

McKechnie, S., Tremblay-Boyer, L., and Harley, S. J. (2015b). Analysis of Pacific-wide operational longline CPUE data for bigeye tuna. WCPFC-SC11-2015/SA-WP-03, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

McKechnie, S., Tremblay-Boyer, L., and Pilling, G. (2017b). Background analyses for the 2017 stock assessments of bigeye and yellowfin tuna in the western and central Pacific Ocean. WCPFC-SC13-2017/SA-IP-06, Rarotonga, Cook Islands, 9–17 August 2017.

MRAG Asia Pacific (2016). *Towards the quantification of illegal, unreported and unregulated (IUU) fishing in the Pacific Islands Region*. MRAG Asia Pacific, 101p.

Peatman, T., Caillot, S., Leroy, B., McKechnie, S., Roupsard, F., Sanchez, C., Nicol, S., and Smith, N. (2016). Analysis of tag seeding data and reporting rates. WCPFC-SC12-2016/SA-IP-13, Bali, Indonesia, 3–11 August 2016.

Pilling, G. and Brouwer, S. (2017). Report from the spc pre-assessment workshop, noumea, april 2017. Technical Report WCPFC-SC13-2017/SA-IP-02, Rarotonga, Cook Islands, 9–17 August 2017.

Powers, J. E. (2013). Review of SPC estimation of species and size composition of the western and central Pacific purse seine fishery from observer-based sampling of the catch. WCPFC-SC9-2013/ST-IP-03, Pohnpei, Federated States of Micronesia, 6–14 August 2013.

Schaefer, K. and Fuller, D. (2002). Movements, behavior, and habitat selection of bigeye tuna (thunnus obesus) in the eastern equatorial pacific, ascertained through archival tags. *Fisheries Bulletin*, 100:765–788.

Schaefer, K., Fuller, D., Hampton, J., Caillot, S., Leroy, B., and Itano, D. (2015). Movements, dispersion, and mixing of bigeye tuna (*Thunnus obesus*) tagged and released in the equatorial Central Pacific Ocean, with conventional and archival tags. *Fisheries Research*, 161:336–355.

Schaefer, K. M., Fuller, D. W., and Miyabe, N. (2005). Reproductive biology of bigeye tuna (*Thunnus obesus*) in the eastern and central Pacific Ocean. Technical Report 23(1). 35p., IATTC Bulletin.

SPC-OFP (2013). Purse seine effort: a recent issue in logbook reporting. WCPFC-TCC9-2013-18, Pohnpei, Federated States of Micronesia.

Thorson, J. T., Johnson, K. F., Methot, R. D., and Taylor, I. G. (2017). Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. *Fisheries Research*, 192:84–93.

Tremblay-Boyer, L., McKechnie, S., Pilling, G., and Hampton, J. (2017). Stock assessment of yellowfin tuna in the Western and Central Pacific Ocean. WCPFC-SC13-2017/SA-WP-06, Rarotonga, Cook Islands, 9–17 August 2017.

Tremblay-Boyer, L. and Pilling, G. (2017a). Geo-statistical analyses of operational longline CPUE data. WCPFC-SC13-2017/SA-WP-03, Rarotonga, Cook Islands, 9–17 August 2017.

Tremblay-Boyer, L. and Pilling, G. (2017b). Use of operational vessel proxies to account for vessels with missing identifiers in the development of standardised CPUE time series. WCPFC-SC13-2017/SA-WP-04, Rarotonga, Cook Islands, 9–17 August 2017.

Williams, P. and Terawasi, P. (2016). Overview of tuna fisheries in the Western and Central Pacific Ocean, including economic conditions–2015. WCPFC-SC12-2016/GN-WP-01, Bali, Indonesia, 3–11 August 2016.

Williams, P. and Terawasi, P. (2017). Overview of tuna fisheries in the Western and Central Pacific Ocean, including economic conditions–2016. WCPFC-SC13-2017/GN-WP-01, Rarotonga, Cook Islands, 9–17 August 2017.

9 Tables

Table 1: Definition of fisheries for the MULTIFAN-CL bigeye analysis. Gears: PL = pole and line; PS = purse seine unspecified set type; LL = longline; DOM = the range of artisanal gear types operating in the domestic fisheries of Philippines and Indonesia. Flag/fleets: JPN = Japan; PH = Philippines; ID = Indonesia; AU = Australia; US = United States; ALL = all nationalities.

Fishery	Nationality	Gear	Region
F1 L-ALL-1	ALL	LL	1
F2 L-ALL-2	ALL	LL	2
F3 L-US-2	US	LL	2
F4 L-ALL-3	ALL	LL	3
F5 L-OS-3	OS	LL	3
F6 L-OS-7	OS	LL	7
F7 L-ALL-7	ALL	LL	7
F8 L-ALL-8	ALL	LL	8
F9 L-ALL-4	ALL	LL	4
F10 L-AU-5	AU	LL	5
F11 L-ALL-5	ALL	LL	5
F12 L-ALL-6	ALL	LL	6
F13 S-ASS-ALL-3	ALL	PS	3
F14 S-UNA-ALL-3	ALL	PS	3
F15 S-ASS-ALL-4	ALL	PS	4
F16 S-UNA-ALL-4	ALL	PS	4
F17 Z-PH-7	PH	Dom	7
F18 Z-ID.PH-7	ID.PH	Dom	7
F19 S-JP-1	JP	PS	1
F20 P-JP-1	JP	PL	1
F21 P-ALL-3	ALL	PL	3
F22 P-ALL-8	ALL	PL	8
F23 Z-ID-7	ID	Dom	7
F24 S-ID.PH-7	ID.PH	PS	7
F25 S-ASS-ALL-8	ALL	PS	8
F26 S-UNA-ALL-8	ALL	PS	8
F27 L-AU-9	AU	LL	9
F28 P-ALL-7	ALL	PL	7
F29 L-ALL-9	ALL	LL	9
F30 S-ASS-ALL-7	ALL	PS	7
F31 S-UNA-ALL-7	ALL	PS	7
F32 Z-VN-7	VN	Dom	7

Table 2: Summary of the number of release events, tag releases and recoveries by region and program

Prog	CSTP			PTTP			RTTP		
Years	1991–2001			2006–2014			1989–1992		
Category	Grps	Rel	Rec	Grps	Rel	Rec	Grps	Rel	Rec
1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0
3	0	0	0	11	1,233	376	5	293	65
4	0	0	0	10	7,911	4,082	3	877	107
5	0	0	0	1	28	7	1	66	18
6	0	0	0	0	0	0	0	0	0
7	0	0	0	3	379	117	5	1,217	370
8	0	0	0	13	1,826	800	5	485	62
9	5	3,571	340	0	0	0	0	0	0
Total	5	3,571	340	38	11,377	5,382	19	2,938	622

Table 3: Summary of the groupings of fisheries within the assessment for estimation of selectivity, catchability (used for the implementation of regional weights), tag recaptures, and tag reporting rates. Note that effort is missing for many Z fisheries and so effort deviation penalties only apply to the last four quarters (see Section 5.3.3). See Table 1 for further details on each fishery.

Fishery	Region	Selectivity	SeasCat	TimVarCat	TimVarCatCV	EffPen	EffPenCV	Recaptures	Reporting
F1 L-ALL-1	1	1	Y	N	NA	time-variant	0.20	1	1
F2 L-ALL-2	2	1	Y	N	NA	time-variant	0.20	2	1
F3 L-US-2	2	2	Y	Y	0.1	scaled	0.41	3	20
F4 L-ALL-3	3	3	Y	N	NA	time-variant	0.20	4	1
F5 L-OS-3	3	4	Y	Y	0.1	scaled	0.41	5	1
F6 L-OS-7	7	5	Y	Y	0.1	scaled	0.41	6	1
F7 L-ALL-7	7	6	Y	N	NA	time-variant	0.20	7	1
F8 L-ALL-8	8	7	Y	N	NA	time-variant	0.20	8	1
F9 L-ALL-4	4	3	Y	N	NA	time-variant	0.20	9	1
F10 L-AU-5	5	8	Y	Y	0.1	scaled	0.41	10	21
F11 L-ALL-5	5	3	Y	N	NA	time-variant	0.20	11	22
F12 L-ALL-6	6	3	Y	N	NA	time-variant	0.20	12	1
F13 S-ASS-ALL-3	3	9	Y	Y	0.7	scaled	0.41	13	23
F14 S-UNA-ALL-3	3	11	Y	Y	0.7	scaled	0.41	13	23
F15 S-ASS-ALL-4	4	10	Y	Y	0.7	scaled	0.41	14	24
F16 S-UNA-ALL-4	4	15	Y	Y	0.7	scaled	0.41	14	24
F17 Z-PH-7	7	12	N	N	NA	constant	0.41	15	25
F18 Z-ID.PH-7	7	13	N	N	NA	constant	0.41	16	26
F19 S-JP-1	1	14	Y	Y	0.1	scaled	0.41	17	27
F20 P-JP-1	1	14	Y	Y	0.1	scaled	0.41	18	28
F21 P-ALL-3	3	14	Y	Y	0.1	scaled	0.41	19	29
F22 P-ALL-8	8	14	Y	Y	0.1	scaled	0.41	19	30
F23 Z-ID-7	7	12	N	N	NA	constant	0.41	20	31
F24 S-ID.PH-7	7	9	N	N	NA	constant	0.41	21	32
F25 S-ASS-ALL-8	8	9	Y	Y	0.7	scaled	0.41	22	33
F26 S-UNA-ALL-8	8	11	Y	Y	0.7	scaled	0.41	22	33
F27 L-AU-9	9	8	Y	Y	0.1	scaled	0.41	23	34
F28 P-ALL-7	7	12	N	Y	0.1	scaled	0.41	24	35
F29 L-ALL-9	9	3	Y	N	NA	scaled	0.41	25	36
F30 S-ASS-ALL-7	7	9	Y	Y	0.7	scaled	0.41	26	23
F31 S-UNA-ALL-7	7	11	Y	Y	0.7	scaled	0.41	26	23
F32 Z-VN-7	7	12	N	N	NA	constant	0.41	27	37

Table 4: Description of symbols used in the yield and stock status analyses. For the purpose of this assessment, "recent" is the average over the period 2011–2014 and 'latest' is 2015

Symbol	Description
C_{latest}	Catch in the last year of the assessment (2015)
F_{recent}	Average fishing mortality-at-age for a recent period (2011–2014)
F_{MSY}	Fishing mortality-at-age producing the maximum sustainable yield (MSY)
MSY	Equilibrium yield at F_{MSY}
$F_{recent}/F_{\mathrm{MSY}}$	Average fishing mortality-at-age for a recent period (2011–2014) relative to F_{MSY}
SB_0	Equilibrium unexploited spawning potential
SB_{latest}	Spawning potential in the latest time period (2015)
SB_{recent}	Spawning potential for a recent period (2011–2014)
$SB_{F=0}$	Average spawning potential predicted to occur in the absence of fishing for the period 2005–2014
SB_{MSY}	Spawning potential that will produce the maximum sustainable yield (MSY)
$SB_{latest}/SB_{F=0}$	Spawning potential in the latest time period (2015) relative to the average spawning potential predicted to occur in the absence of fishing for the period 2005–2014
$SB_{latest}/SB_{\mathrm{MSY}}$	Spawning potential in the latest time period (2015) relative to that which will produce the maximum sustainable yield (MSY)
$SB_{recent}/SB_{F=0}$	Spawning potential in for a recent period (2011–2014) relative to the average spawning potential predicted to occur in the absence of fishing for the period 2005–2014
$20\%SB_{F=0}$	WCPFC adopted limit reference point – 20% of spawning potential in the absence of fishing average over years $t-10$ to $t-1$ (2005–2014)

Table 5: Description of the structural sensitivity grid used to characterise uncertainty in the assessment.

Axis	Levels	Option
Steepness	3	0.65, 0.80, or 0.95
Growth	2	Low growth L2 = 152, High growth L2 = 184
Tagging overdispersion	2	Default level, Fixed (moderate) level
Size frequency weighting	2	sample sizes divided by 10, or 50
Regional structure	2	2017 regions, 2014 regions

Table 6: Summary of reference points over all 48 individual models in the structural uncertainty grid

	Mean	Median	Min	25%	75%	Max
C_{latest}	149,361	153,621	130,903	143,812	155,554	157,725
MSY	153,444	151,380	124,120	139,140	164,760	204,040
$Y_{F_{recent}}$	145,970	145,660	118,000	134,660	155,640	187,240
f_{mult}	1.09	1.01	0.57	0.83	1.35	1.85
F_{MSY}	0.05	0.05	0.04	0.04	0.05	0.06
F_{recent}/F_{MSY}	1.01	0.99	0.54	0.74	1.21	1.76
SB_{MSY}	421,281	425,350	219,500	319,050	509,775	710,000
SB_0	1,632,875	1,633,500	1,009,000	1,320,750	1,932,000	2,509,000
SB_{MSY}/SB_0	0.25	0.26	0.22	0.24	0.27	0.29
$SB_{F=0}$	1,853,576	1,930,603	1,317,335	1,624,408	2,046,314	2,460,410
$SB_{MSY}/SB_{F=0}$	0.22	0.23	0.17	0.19	0.25	0.29
SB_{latest}/SB_0	0.31	0.30	0.11	0.22	0.41	0.51
$SB_{latest}/SB_{F=0}$	0.28	0.28	0.08	0.18	0.40	0.49
SB_{latest}/SB_{MSY}	1.23	1.21	0.42	0.92	1.58	2.09
$SB_{recent}/SB_{F=0}$	0.25	0.25	0.08	0.18	0.34	0.42
SB_{recent}/SB_{MSY}	1.05	1.02	0.32	0.74	1.39	1.86

Table 7: Summary of reference points over the 12 models in the structural uncertainty grid within the subset of new growth/ 2017 regions

	Mean	Median	Min	25%	75%	Max
C_{latest}	143,358	143,157	130,903	131,495	155,057	155,972
MSY	173,043	172,920	144,600	160,500	185,720	204,040
$Y_{F_{recent}}$	159,990	160,820	135,200	147,270	169,000	187,240
f_{mult}	1.59	1.59	1.36	1.48	1.67	1.85
F_{MSY}	0.05	0.05	0.04	0.04	0.05	0.05
F_{recent}/F_{MSY}	0.63	0.63	0.54	0.60	0.67	0.74
SB_{MSY}	553,425	537,900	428,000	503,875	612,150	710,000
SB_0	2,090,583	2,072,000	1,729,000	1,963,500	2,189,500	2,509,000
SB_{MSY}/SB_0	0.26	0.26	0.24	0.25	0.28	0.28
$SB_{F=0}$	2,125,165	2,135,463	1,812,463	2,030,472	2,193,277	2,460,411
$SB_{MSY}/SB_{F=0}$	0.26	0.26	0.23	0.24	0.28	0.29
SB_{latest}/SB_0	0.46	0.46	0.40	0.43	0.48	0.51
$SB_{latest}/SB_{F=0}$	0.45	0.45	0.40	0.43	0.46	0.49
SB_{latest}/SB_{MSY}	1.74	1.77	1.40	1.55	1.92	2.09
$SB_{recent}/SB_{F=0}$	0.38	0.39	0.35	0.37	0.39	0.42
SB_{recent}/SB_{MSY}	1.62	1.62	1.39	1.53	1.69	1.86

Table 8: Summary of reference points over the 12 models in the structural uncertainty grid within the subset of new growth/ 2014 regions

	Mean	Median	Min	25%	75%	Max
C_{latest}	147,375	147,434	141,388	141,679	153,142	153,153
MSY	152,203	152,040	134,800	144,330	158,400	170,760
$Y_{F_{recent}}$	150,530	150,220	133,400	143,450	155,640	170,000
f_{mult}	1.17	1.17	0.98	1.10	1.23	1.35
F_{MSY}	0.05	0.05	0.04	0.04	0.05	0.05
F_{recent}/F_{MSY}	0.86	0.86	0.74	0.81	0.91	1.02
SB_{MSY}	470,208	466,050	372,400	421,850	515,875	579,800
SB_0	1,719,500	1,712,500	1,466,000	1,617,000	1,791,000	1,999,000
SB_{MSY}/SB_0	0.27	0.27	0.25	0.26	0.29	0.29
$SB_{F=0}$	1,886,275	1,880,531	1,637,143	1,801,576	1,954,738	2,144,857
$SB_{MSY}/SB_{F=0}$	0.25	0.25	0.23	0.23	0.26	0.27
SB_{latest}/SB_0	0.39	0.39	0.32	0.36	0.41	0.45
$SB_{latest}/SB_{F=0}$	0.35	0.36	0.29	0.34	0.37	0.40
SB_{latest}/SB_{MSY}	1.43	1.44	1.12	1.26	1.60	1.75
$SB_{recent}/SB_{F=0}$	0.29	0.29	0.24	0.28	0.31	0.33
SB_{recent}/SB_{MSY}	1.19	1.20	0.97	1.12	1.27	1.39

Table 9: Summary of reference points over the 12 models in the structural uncertainty grid within the subset of old growth/ 2017 regions

	Mean	Median	Min	25%	75%	Max
C_{latest}	150,982	151,641	144,501	147,021	155,639	155,965
MSY	143,143	141,480	124,120	136,600	149,020	164,640
$Y_{F_{recent}}$	142,323	141,080	123,880	135,780	148,180	163,760
f_{mult}	0.93	0.91	0.80	0.90	0.97	1.03
F_{MSY}	0.04	0.04	0.04	0.04	0.05	0.05
$F_{recent}/F_{\mathrm{MSY}}$	1.08	1.10	0.97	1.03	1.11	1.24
SB_{MSY}	360,092	334,400	278,100	288,800	430,800	470,400
SB_0	1,469,167	1,395,000	1,256,000	1,284,250	1,651,250	1,767,000
SB_{MSY}/SB_0	0.24	0.24	0.22	0.23	0.26	0.27
$SB_{F=0}$	1,720,062	1,707,605	1,474,692	1,564,316	1,930,175	2,001,488
$SB_{MSY}/SB_{F=0}$	0.21	0.20	0.18	0.19	0.22	0.24
SB_{latest}/SB_0	0.26	0.25	0.22	0.24	0.27	0.28
$SB_{latest}/SB_{F=0}$	0.22	0.22	0.19	0.20	0.23	0.26
$SB_{latest}/SB_{\mathrm{MSY}}$	1.06	1.06	0.86	0.93	1.17	1.29
$SB_{recent}/SB_{F=0}$	0.22	0.23	0.19	0.20	0.23	0.26
$SB_{recent}/SB_{\mathrm{MSY}}$	0.89	0.87	0.70	0.85	0.95	1.05

Table 10: Summary of reference points over the 12 models in the structural uncertainty grid within the subset of old growth/ 2014 regions

	Mean	Median	Min	25%	75%	Max
C_{latest}	155,730	155,354	154,089	154,782	156,708	157,725
MSY	145,387	140,280	126,560	131,410	164,760	170,080
$Y_{F_{recent}}$	131,037	131,000	118,000	121,590	135,300	149,400
f_{mult}	0.69	0.67	0.57	0.62	0.73	0.83
F_{MSY}	0.05	0.05	0.04	0.05	0.05	0.06
$F_{recent}/F_{\mathrm{MSY}}$	1.48	1.48	1.20	1.37	1.62	1.76
SB_{MSY}	301,400	282,300	219,500	245,625	378,875	393,200
SB_0	1,252,250	1,187,500	1,009,000	1,110,500	1,471,500	1,519,000
SB_{MSY}/SB_0	0.24	0.24	0.22	0.22	0.26	0.26
$SB_{F=0}$	1,682,806	1,603,006	1,317,336	1,442,138	2,008,122	2,137,922
$SB_{MSY}/SB_{F=0}$	0.18	0.18	0.17	0.17	0.18	0.19
SB_{latest}/SB_0	0.16	0.15	0.11	0.14	0.18	0.22
$SB_{latest}/SB_{F=0}$	0.12	0.11	0.08	0.10	0.14	0.17
$SB_{latest}/SB_{\mathrm{MSY}}$	0.67	0.64	0.42	0.55	0.79	0.98
$SB_{recent}/SB_{F=0}$	0.12	0.11	0.08	0.10	0.14	0.17
$SB_{recent}/SB_{\mathrm{MSY}}$	0.51	0.49	0.32	0.40	0.58	0.74

10 Figures
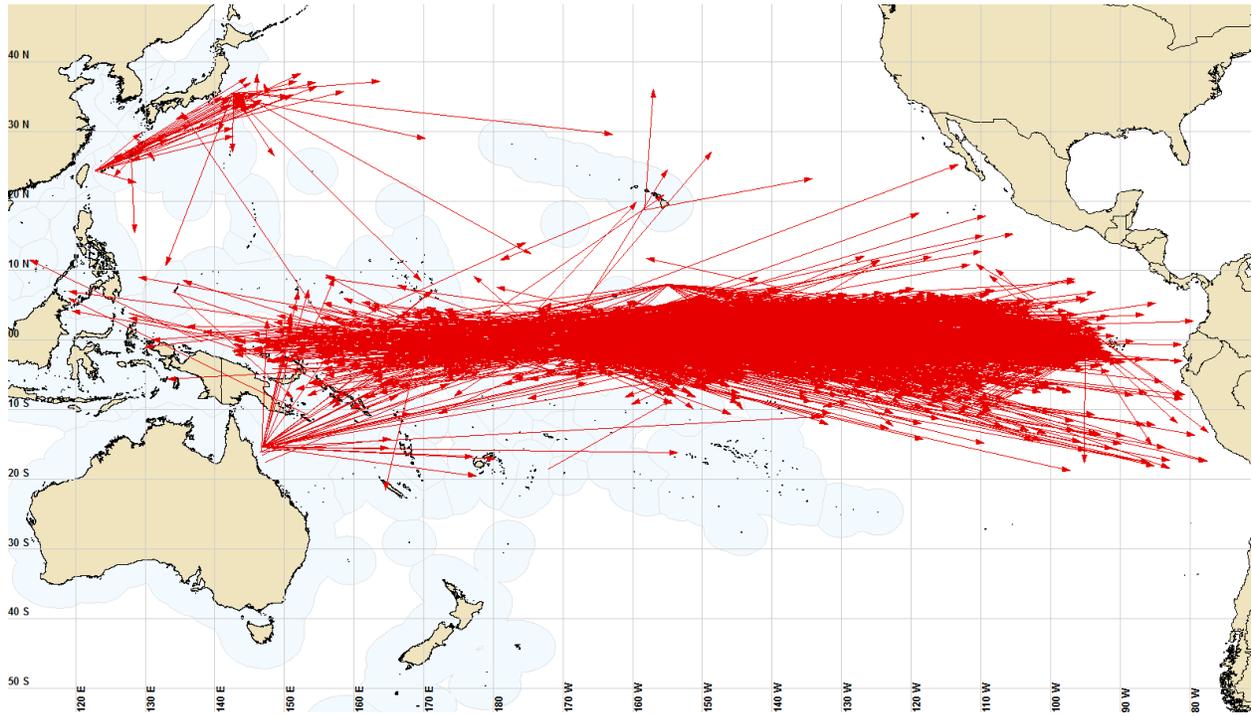
Figure 1: Map of the movements of tagged bigeye released in the Pacific Ocean and subsequently recaptured more than 1,000 nautical miles from their release site.
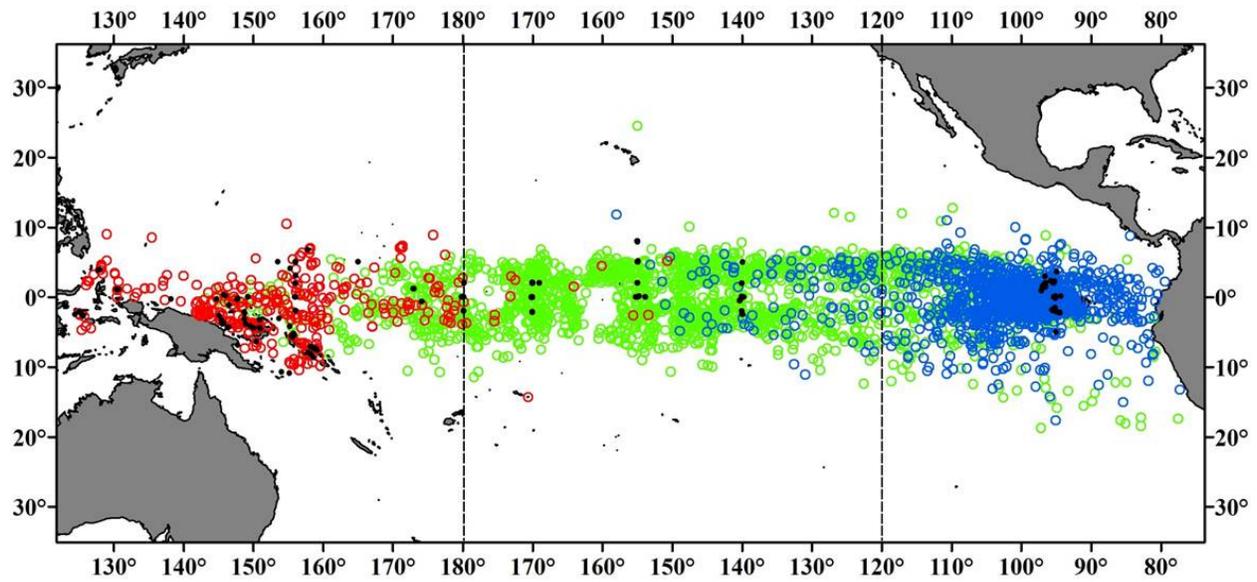
Figure 2: Map of the movements of tagged bigeye released in the Pacific Ocean and subsequently recaptured. The figure is sourced from (Schaefer et al., 2015), and shows the three regions they split their data by. The small black points are the release locations, the red points are the recapture locations of fish released in the western region, the green points are the recapture locations of fish released in the central region and the blue points are the recapture locations of fish released in the eastern region.
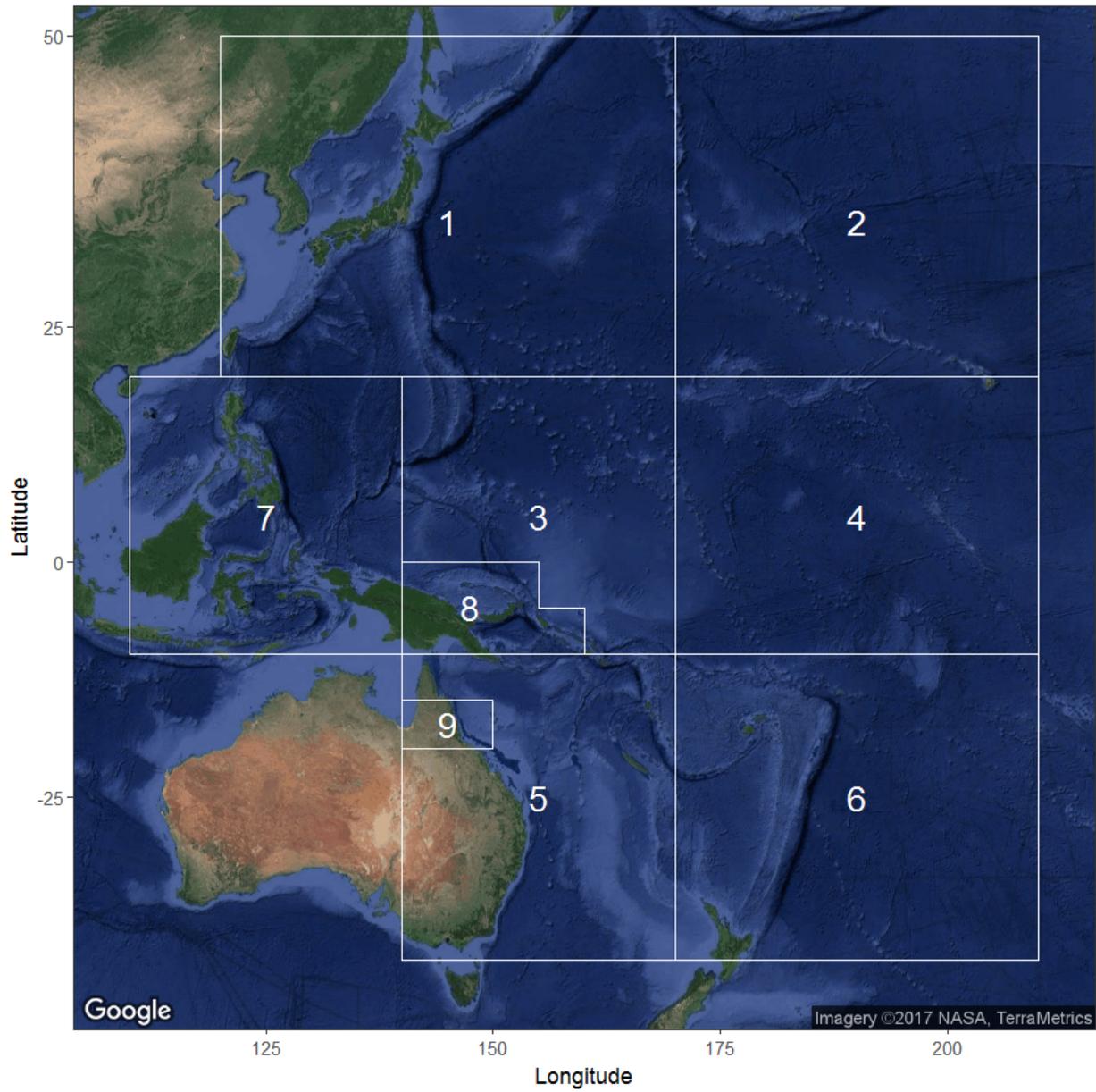
Figure 3: The geographical area covered by the stock assessment and the boundaries for the 9 regions when using the "2014 regional structure".
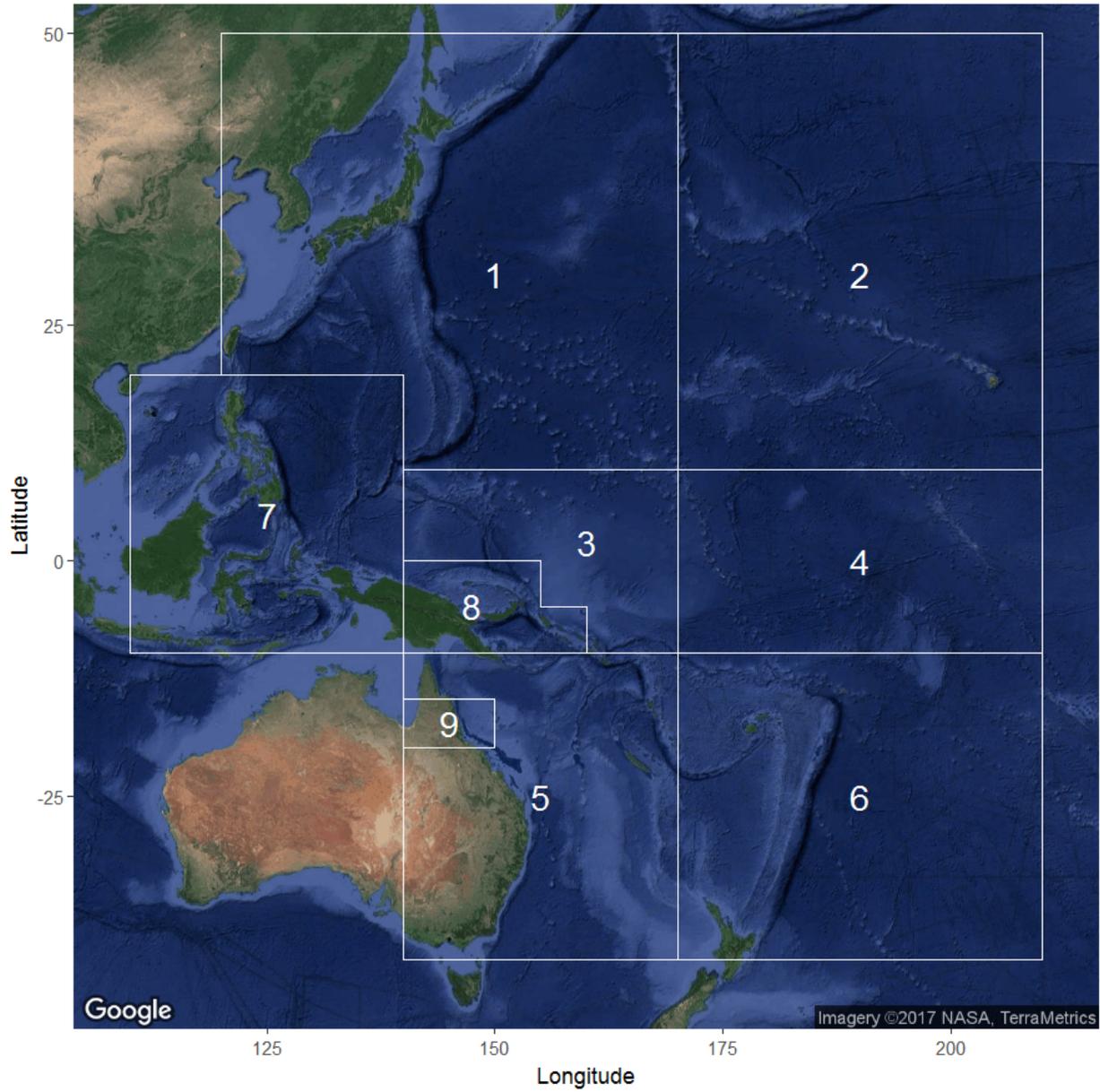
Figure 4: The geographical area covered by the stock assessment and the boundaries for the 9 regions when using the "2017 regional structure".
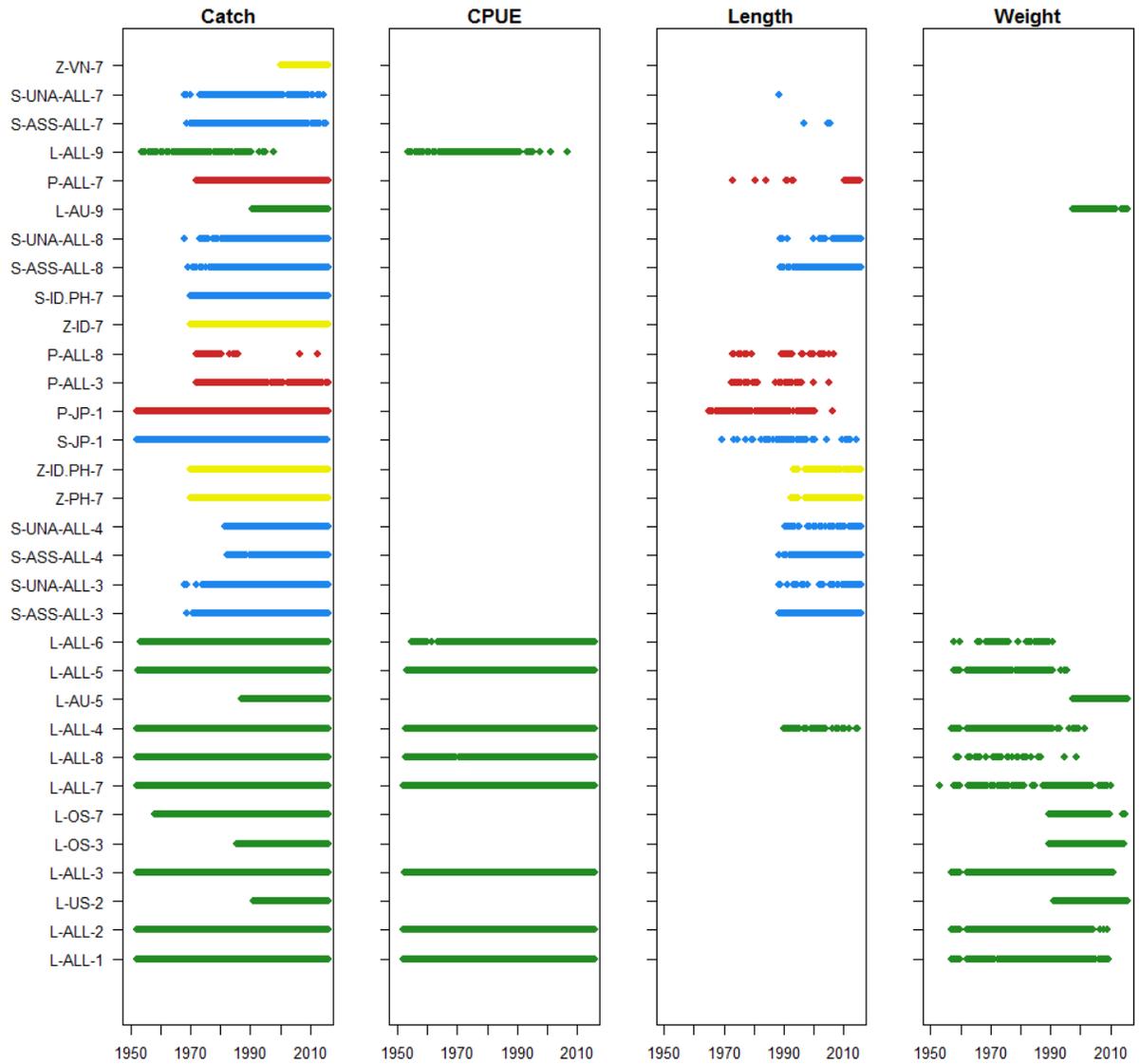
Figure 5: Presence of catch, standardised CPUE, length frequency and weight frequency data by year and fishery for the diagnostic case model (2017 regional structure, hence 32 fisheries). The different colours denote gear-type of the fishery: longline (green); pole-and-line (red); purse seine (blue); and miscellaneous (yellow).
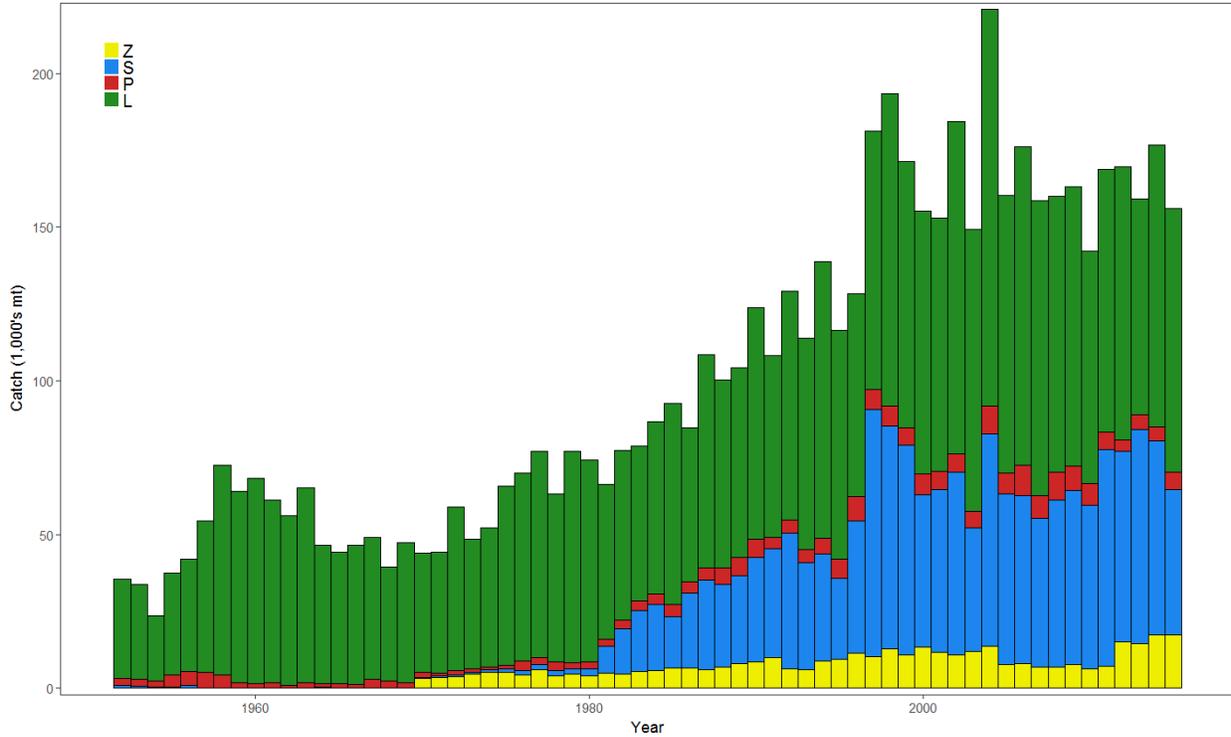
Figure 6: Time series of total annual catch (1000's mt) by fishing gear for the diagnostic case model over the full assessment period. The different colours refer to longline (green), pole-and-line (red), purse seine (blue) and miscellaneous (yellow). Note that the catch by longline gear has been converted into catch-in-weight from catch-in-numbers and so estimates differs from the annual catch estimates presented in (Williams and Terawasi, 2017), however these catches enter the model as catch-in-numbers.
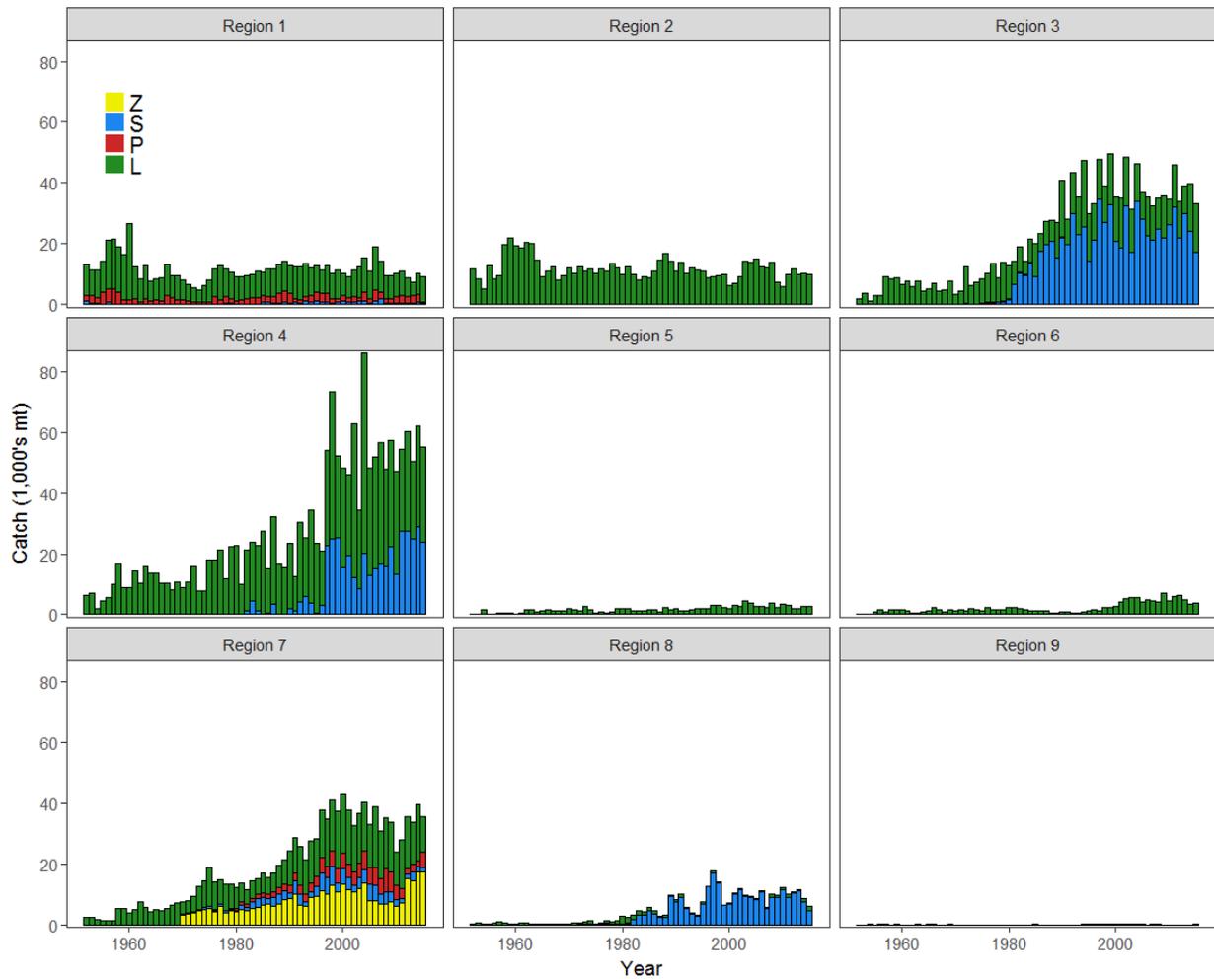
Figure 7: Time series of total annual catch (1000's mt) by fishing gear and assessment region from the diagnostic case model over the full assessment period. The different colours denote longline (green), pole-and-line (red), purse seine (blue) and miscellaneous (yellow).
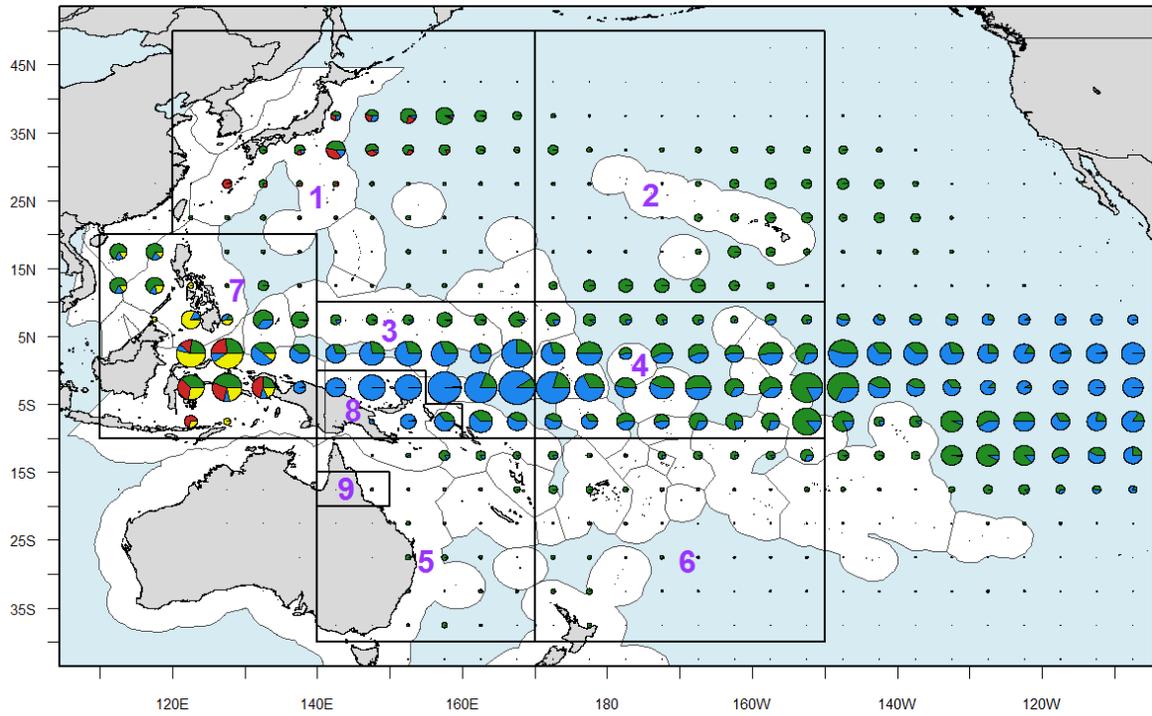
Figure 8: Distribution and magnitude of bigeye tuna catches for the most recent decade of the stock assessment (2006-2015) by 5° square and fishing gear: longline (green), pole-and-line (red), purse seine (blue) and miscellaneous (yellow), for the WCPO and part of the EPO. Overlayed are the regional boundaries for the stock assessment (2017 regional structure).
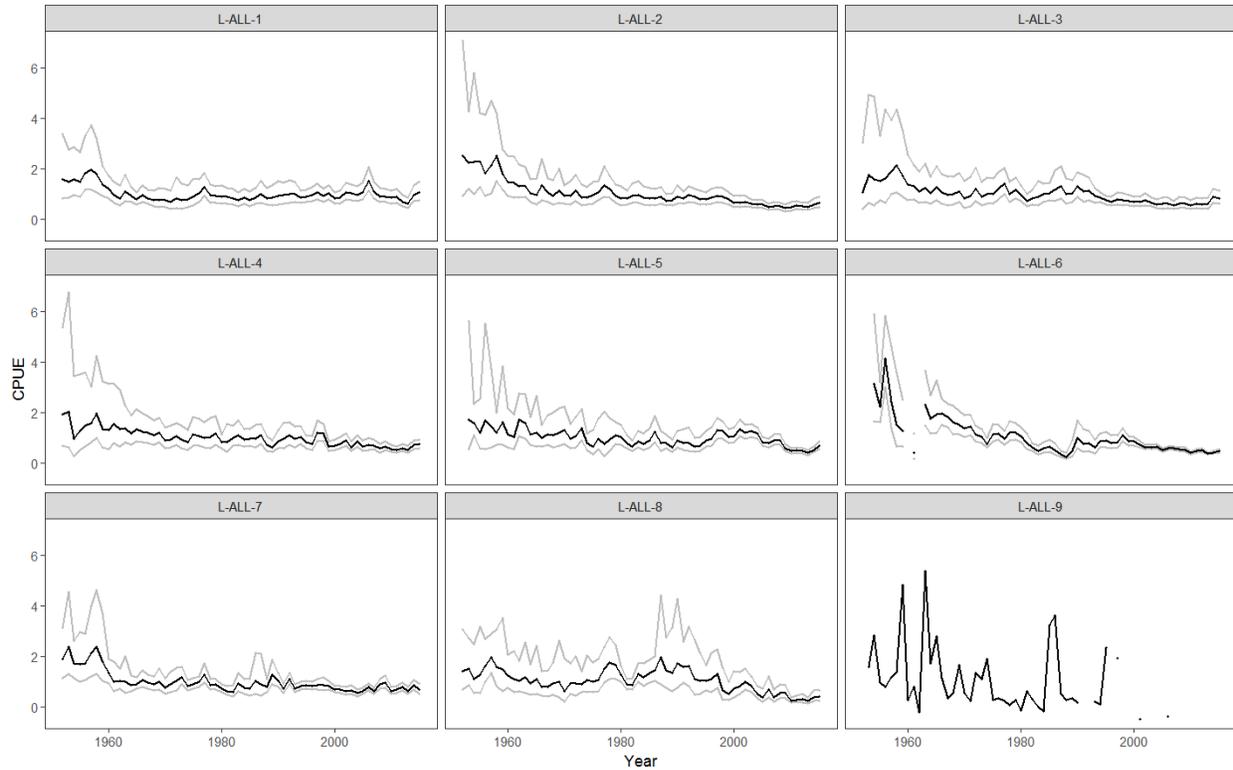
Figure 9: Standardised catch-per-unit-effort (CPUE) indices for the longline fisheries in regions 1–9 used in the diagnostic case model. See McKechnie et al. (2017a) for further details of the estimation of these CPUE indices. The light grey lines represent the 95% confidence intervals derived from the effort deviation penalties used in the diagnostic case model.
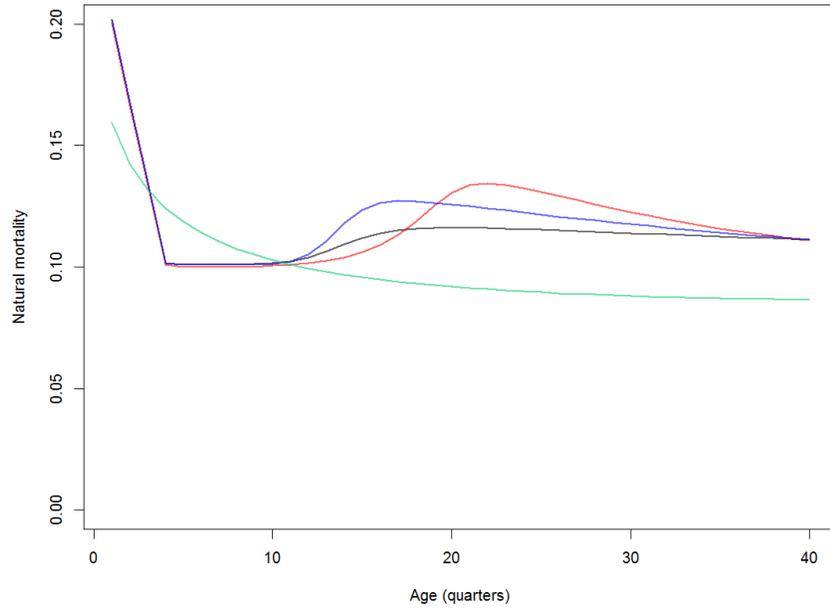
Figure 10: Quarterly natural mortality-at-age as used the diagnostic case model (new growth/new maturity; black line) and the sensitivity models *L2-184* (old growth/old maturity; red line) and *L2-184-NewMat* (old growth/new maturity; blue line). The green line shows the estimated m-at-age function for the sensitivity model assuming a Lorenzen-type function (*Lorenzen*) between mortality and fish size (presented in Section 6.2).
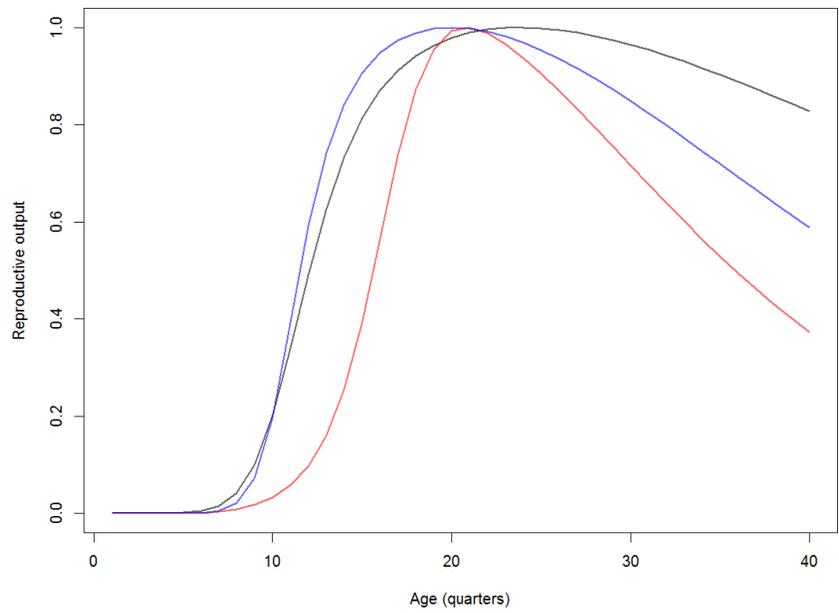
Figure 11: Maturity-at-age as used in the diagnostic case model (new growth/new maturity; black line) and the sensitivity models *L2-184* (old growth/old maturity; red line) and *L2-184-NewMat* (old growth/new maturity; blue line).
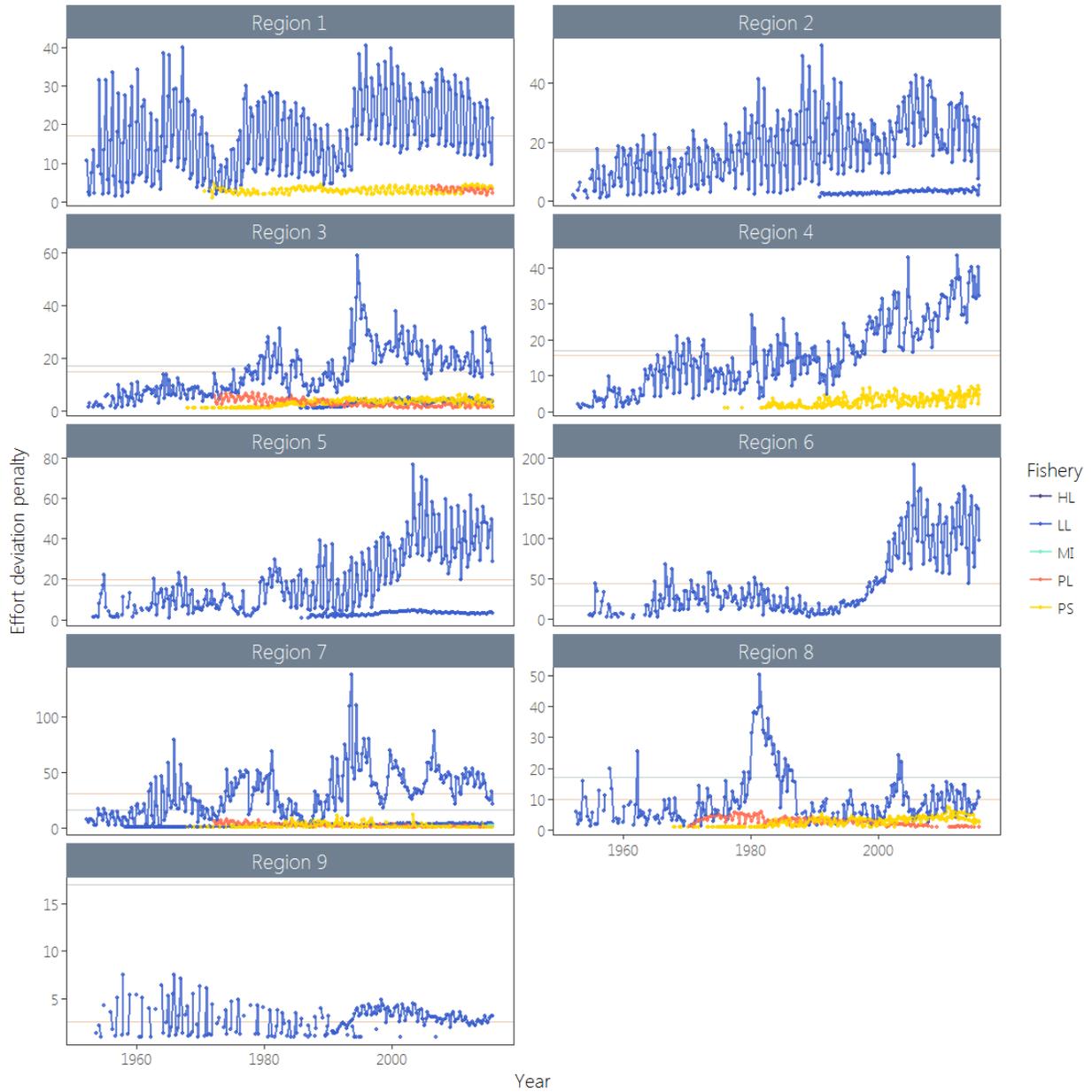
Figure 12: Plot of the effort deviation penalties applied to each fishery, by region, with the colours of the lines representing the gear of the fishery. In several cases there is more than one fishery for a given gear-type in a region (e.g. regions 3 and 7 both have two longline fisheries, though only a single fishery receives a standardised CPUE index in both regions). A higher penalty gives more weight to the CPUE of that fishery and so the high weightings applied to the standardised CPUE indices are evident.

94
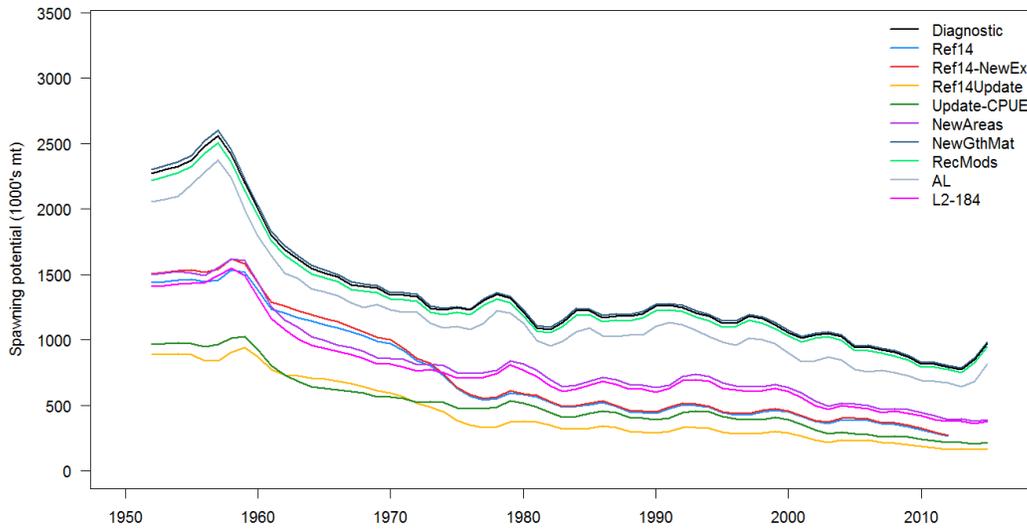
(a) Progression from 2014 to 2017 diagnostic model - spawning potential
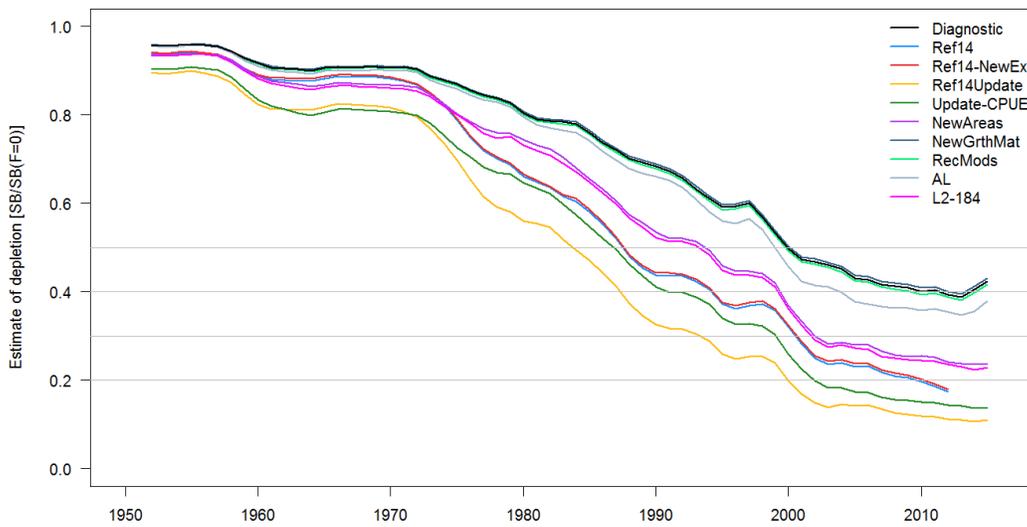
(b) Progression from 2014 to 2017 diagnostic model - depletion

Figure 13: Stepwise changes in spawning potential (a), and fishing depletion (b) from the 2014 reference case model through to the 2017 diagnostic case model.
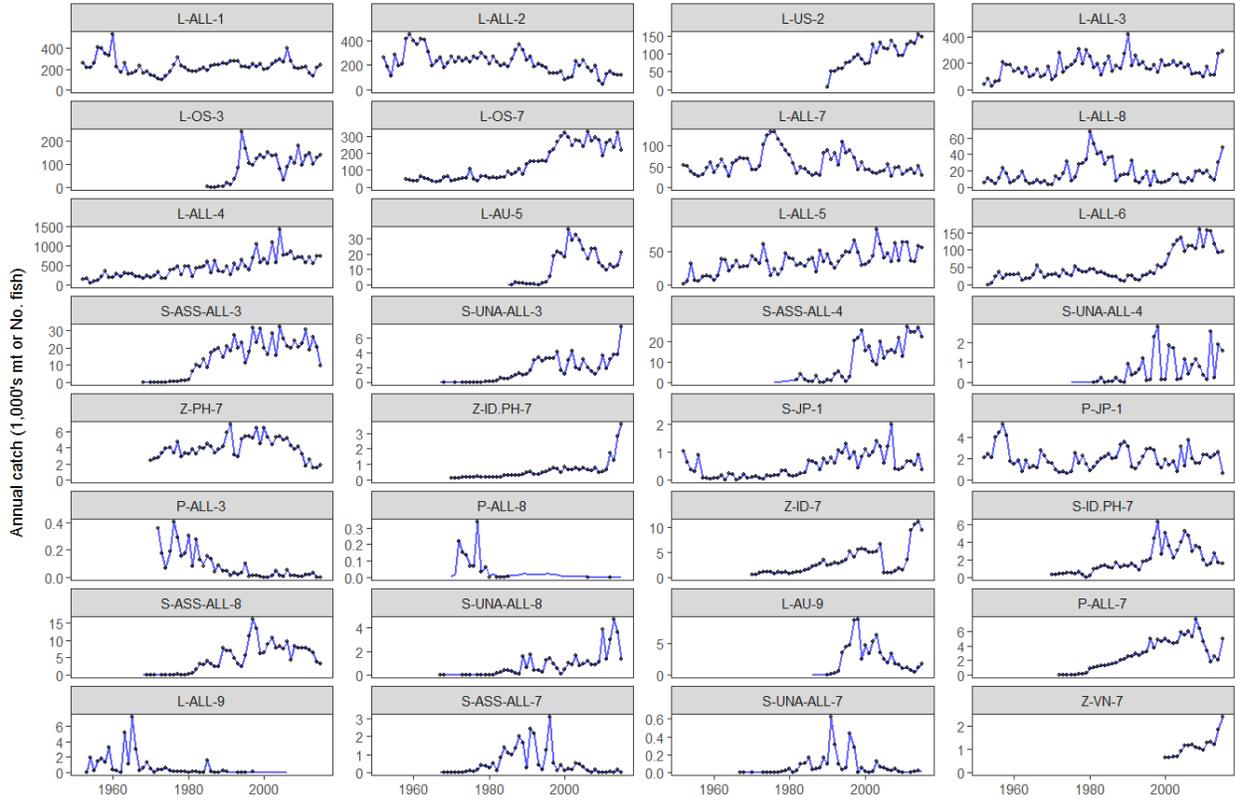
Figure 14: Observed (black points) and model-predicted (blue lines) catch for the 32 fisheries in the diagnostic case model. The y-axis is in catch-in-numbers for the longline fisheries and catch-in-weight for the other fisheries, both divided by 1,000.
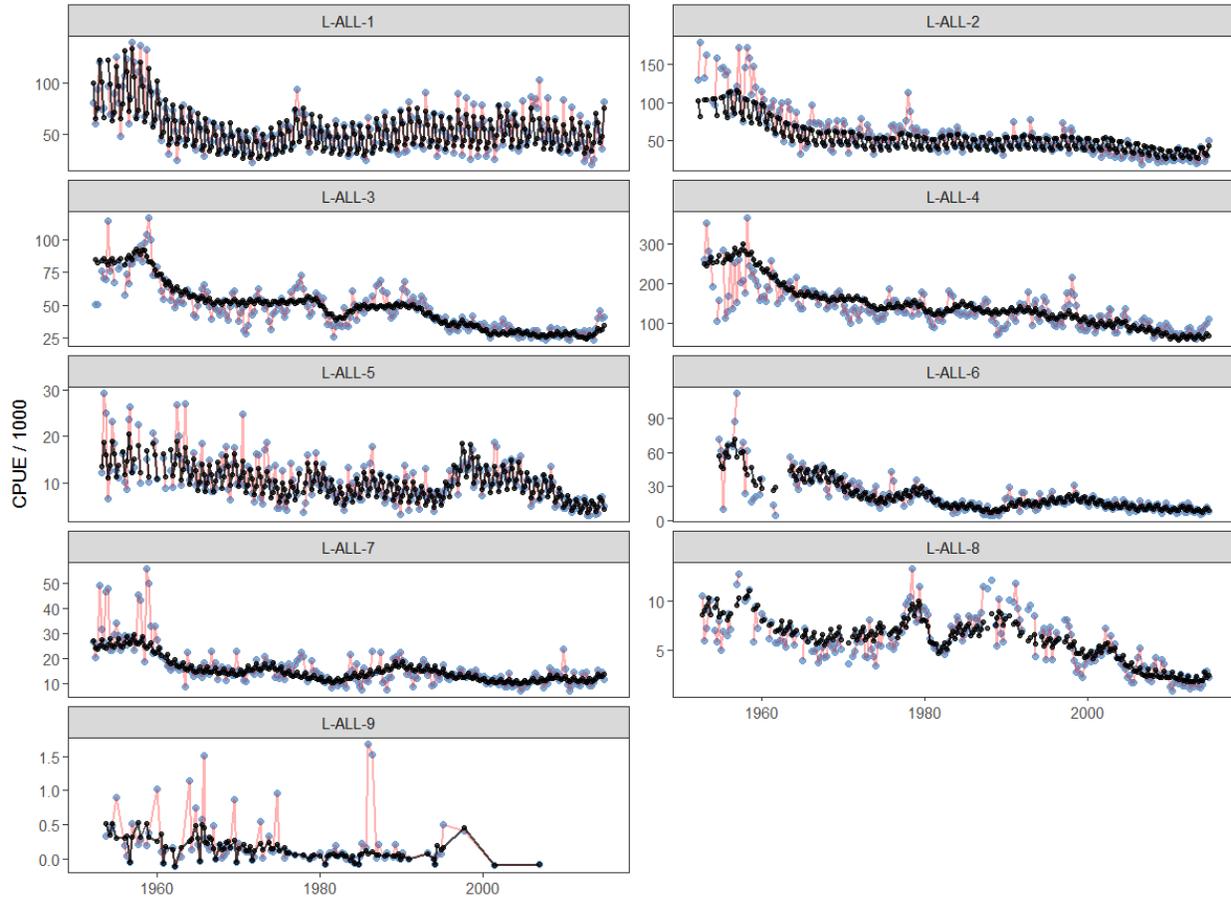
Figure 15: Observed (blue points and red lines) and model-predicted (black points and lines) CPUE for the eight fisheries which received standardised CPUE indices in the diagnostic case model, and the nominal CPUE index used for region 9.
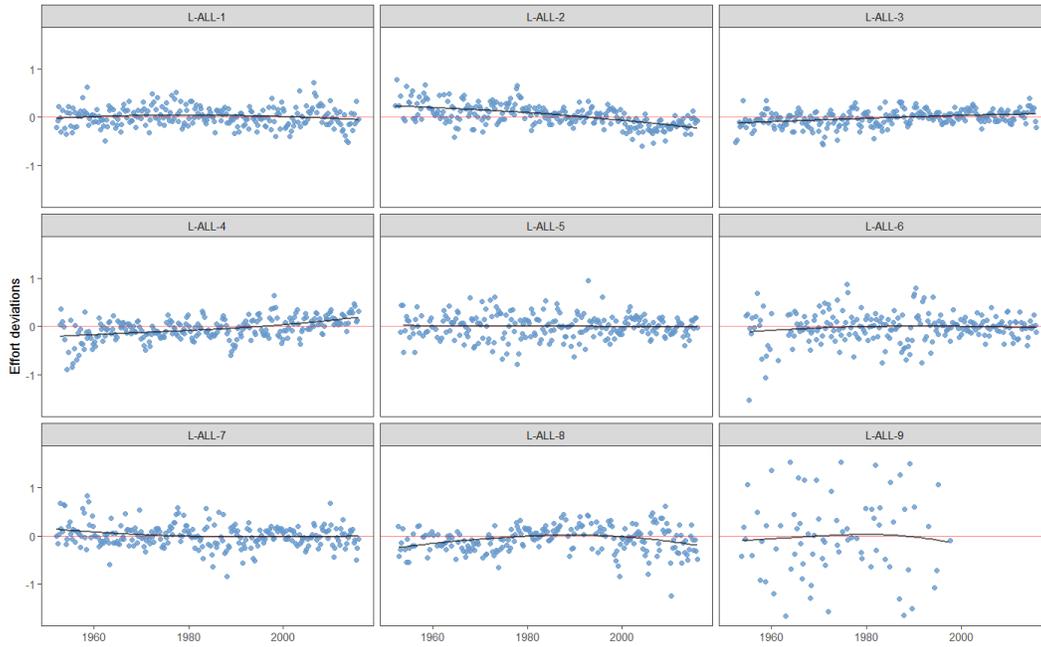
Figure 16: Effort deviations by time period for each of the fisheries receiving standardised CPUE indices in the diagnostic case model. The dark line represents a lowess smoothed fit to the effort deviations.
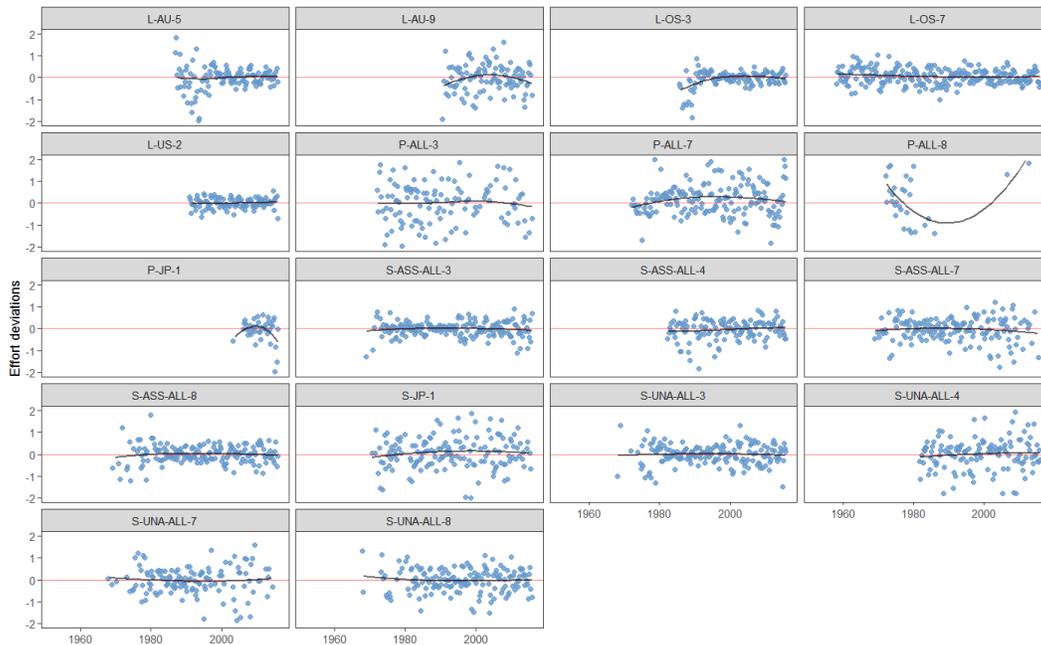
Figure 17: Effort deviations by time period for each of the fisheries that did not receive standardised CPUE indices in the diagnostic case model. The dark line represents a lowess smoothed fit to the effort deviations.
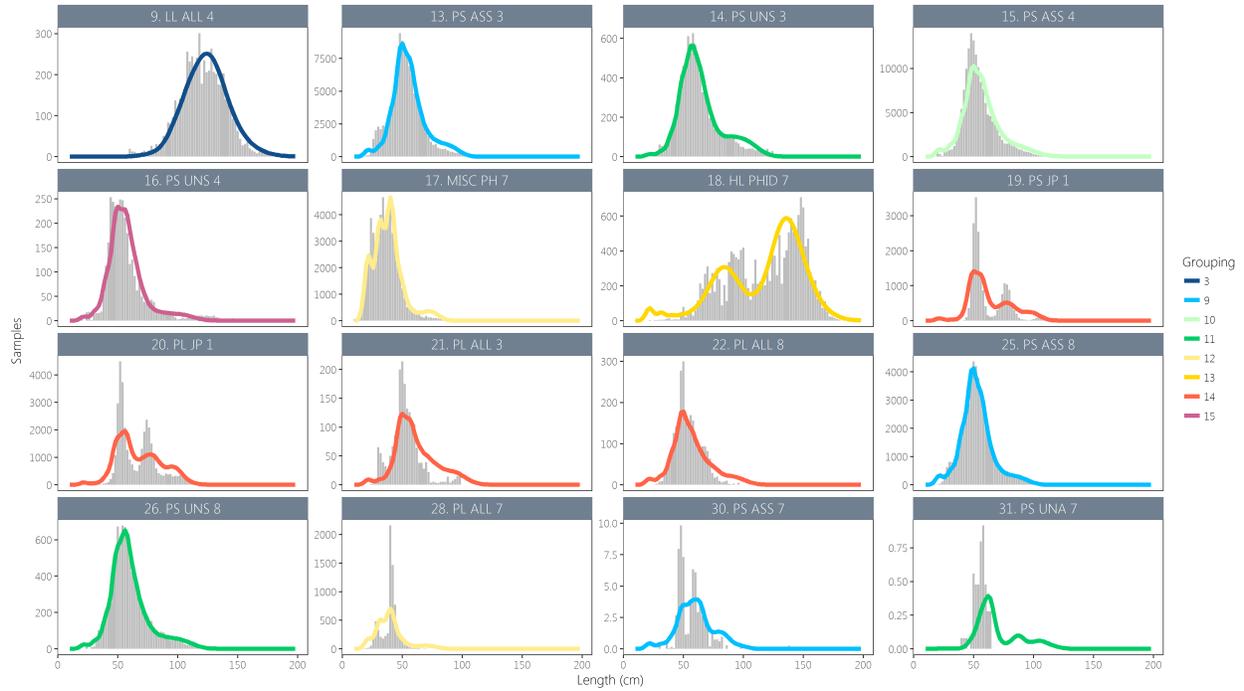
Figure 18: Composite (all time periods combined) observed (grey histograms) and predicted (coloured lines) catch-at-length for all fisheries with samples for the diagnostic case model. The colours indicate the groupings of the fisheries with respect to selectivity, such that fisheries sharing the same colour in the plot also share selectivity functions in the model.
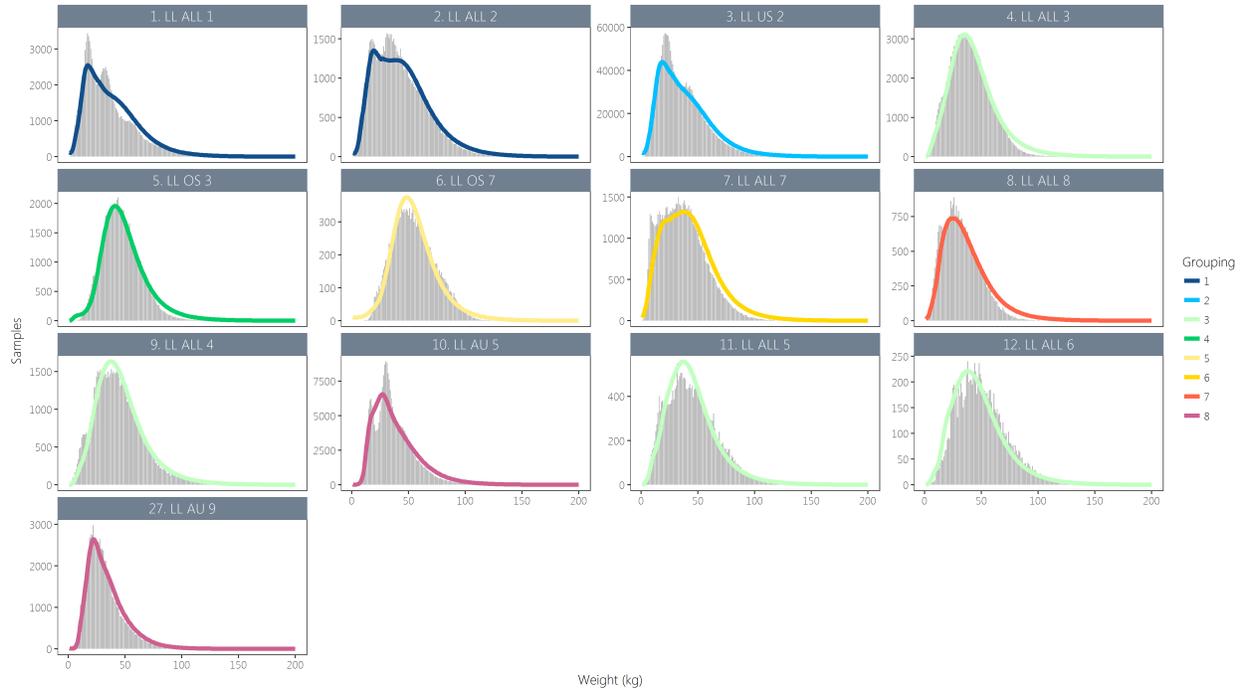
Figure 19: Composite (all time periods combined) observed (grey histograms) and predicted (coloured lines) catch-at-weight for all fisheries with samples for the diagnostic case model. The colours indicate the groupings of the fisheries with respect to selectivity, such that fisheries sharing the same colour in the plot also share selectivity functions in the model.
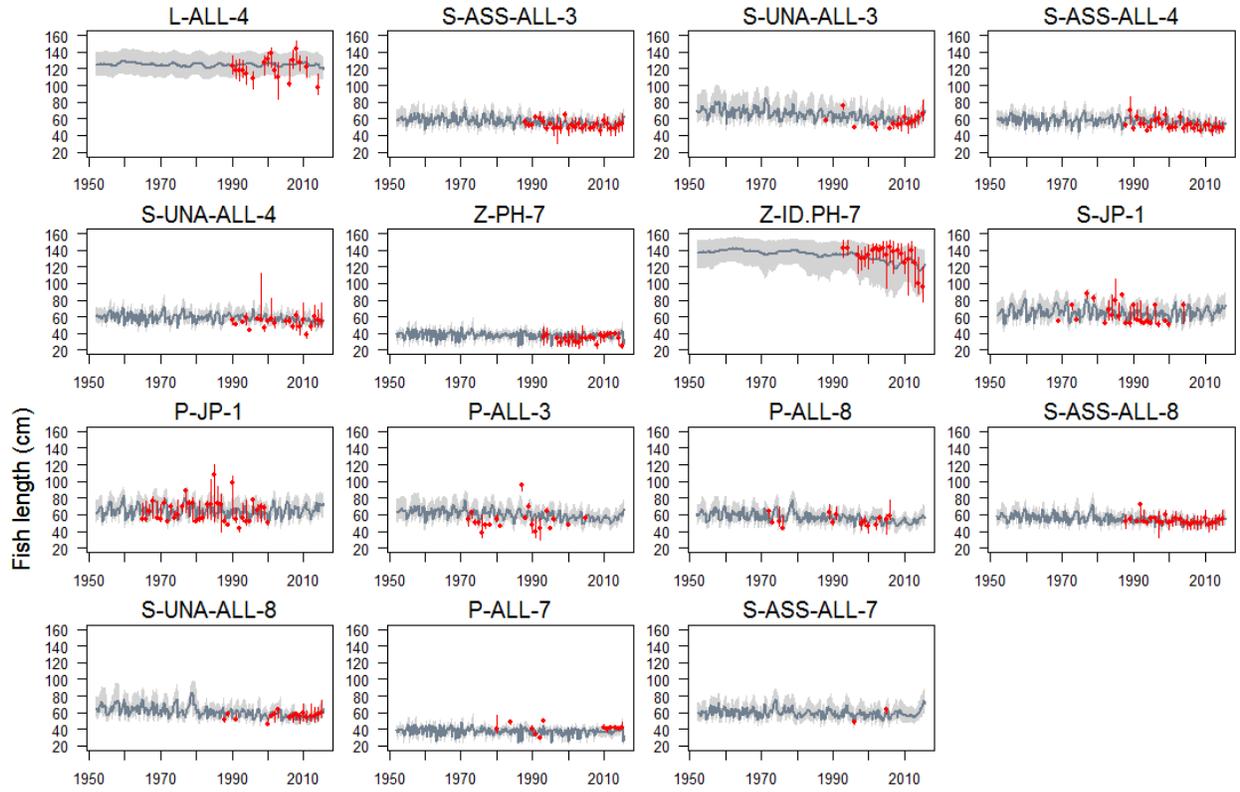
Figure 20: A comparison of the observed (red points) and predicted (grey line) median fish length (FL, cm) for all fisheries with samples for the diagnostic case model. The uncertainty intervals (grey shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are aggregated by year and only length samples with a minimum of 30 fish per year are plotted.
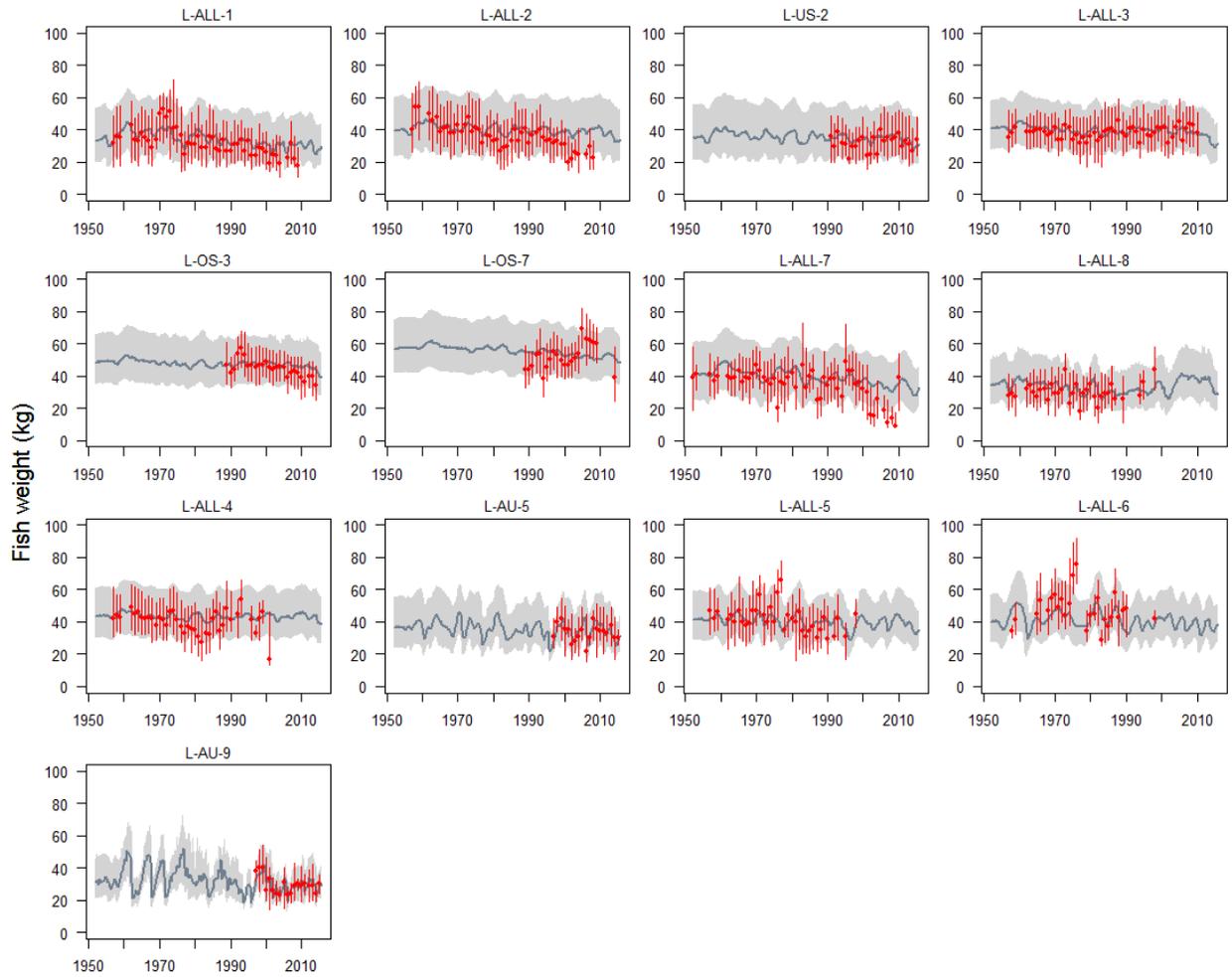
Figure 21: A comparison of the observed (red points) and predicted (grey line) median fish weight (kg) for all fisheries with samples for the diagnostic case model. The uncertainty intervals (grey shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are aggregated by year and only length samples with a minimum of 30 fish per year are plotted.
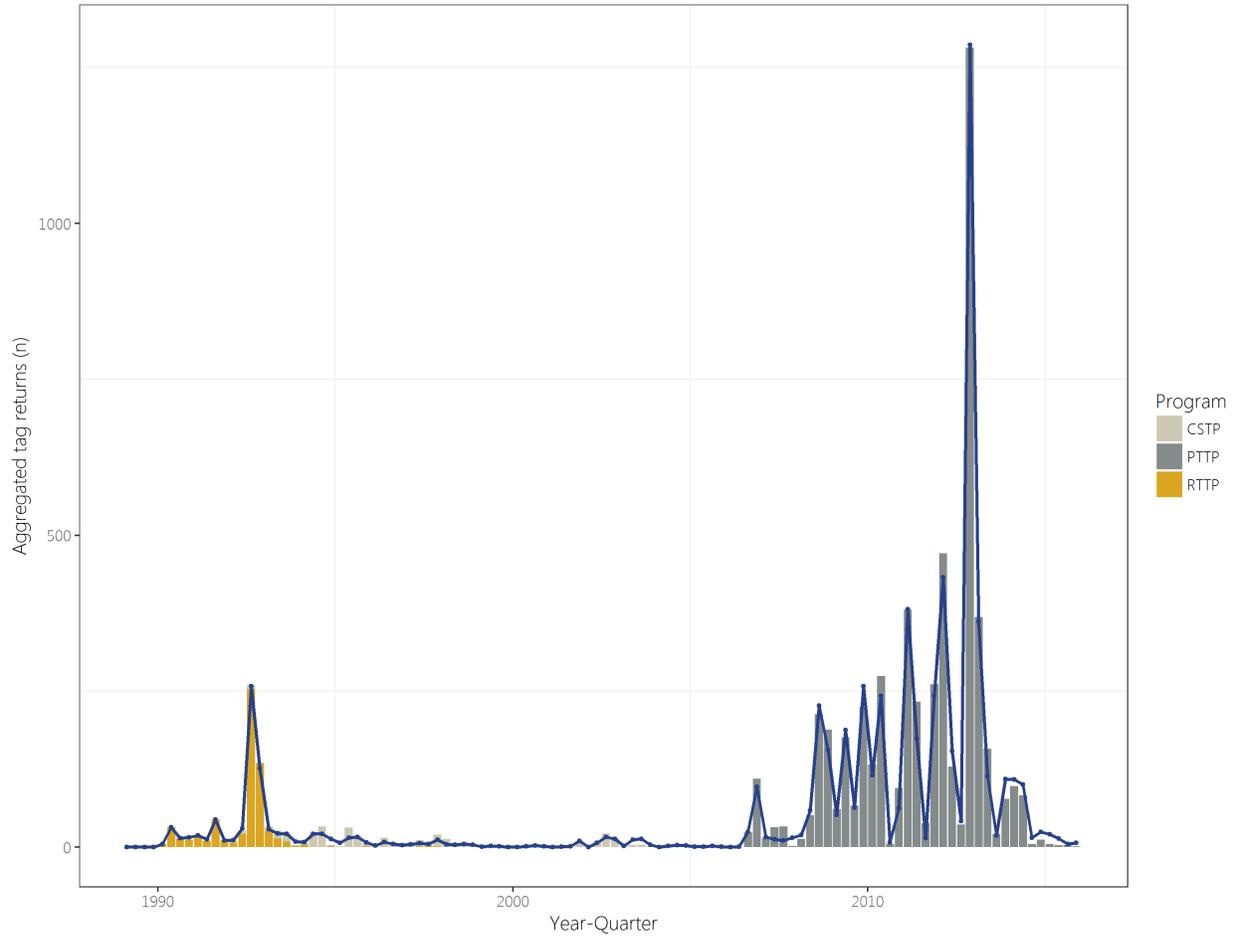
Figure 22: Observed (coloured bars) and model-predicted (blue line) tag returns over time for the diagnostic case model across all tag release events with all tag recapture groupings aggregated. The colour of the bars denotes the tagging programme from which the recaptured fish were released.
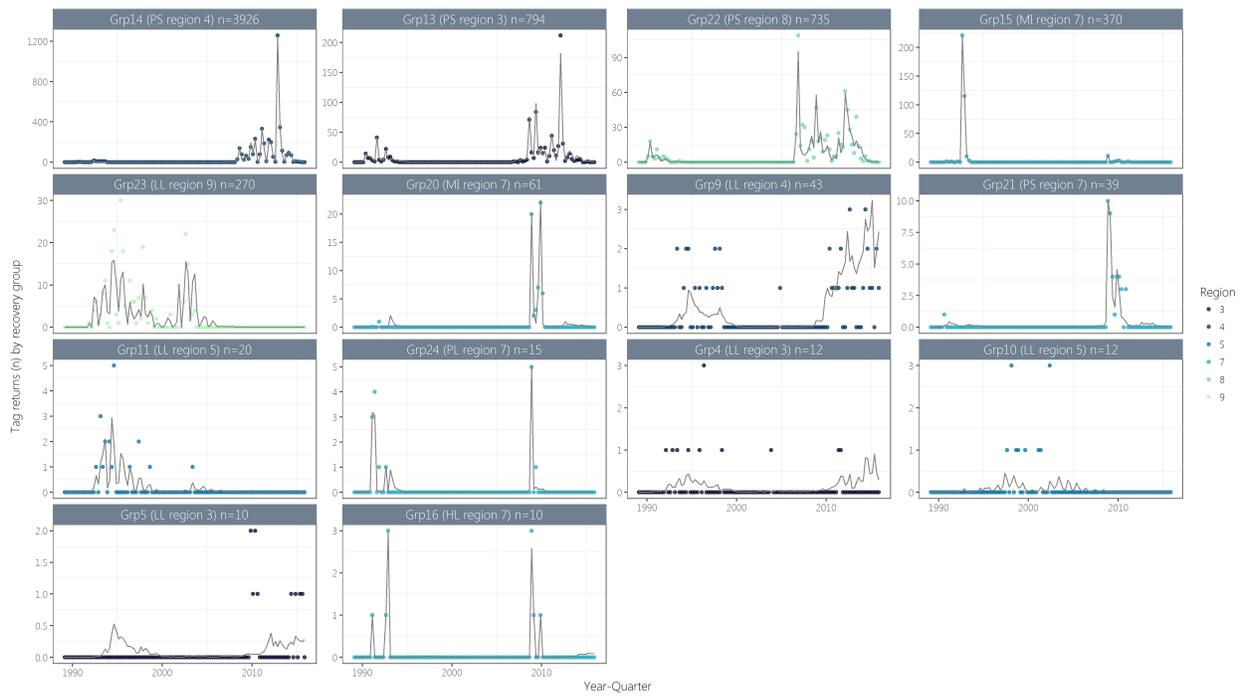
Figure 23: Observed (points) and model-predicted (black line) tag returns over time for the diagnostic case model by recapture group for groups with at least 10 recaptures. Groups with extremely low numbers of recaptures are uninformative and not shown. The colours of the points denote the region where the recapture groups are located.

104
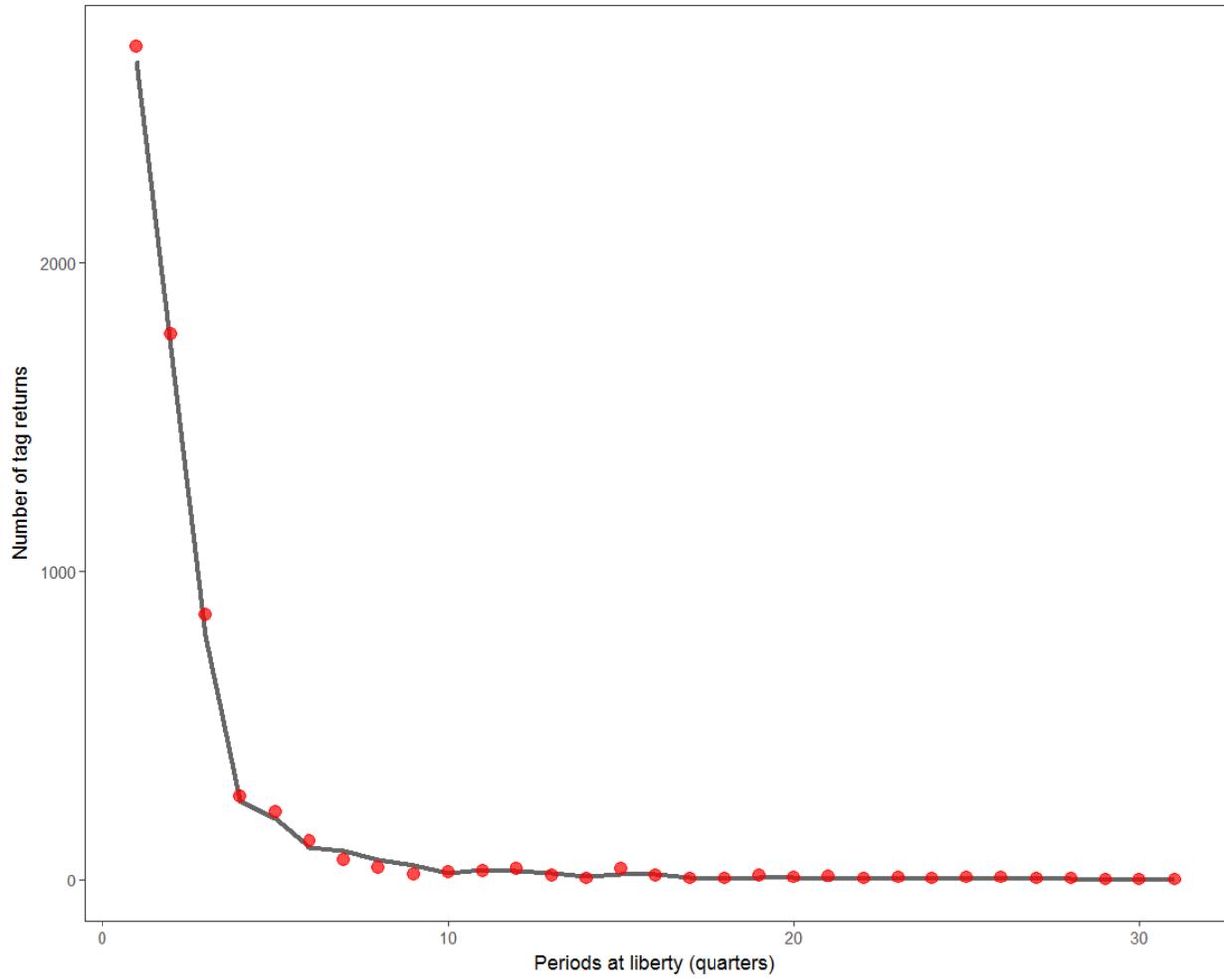
Figure 24: Observed and model-predicted tag attrition across all tag release events for the diagnostic case model.
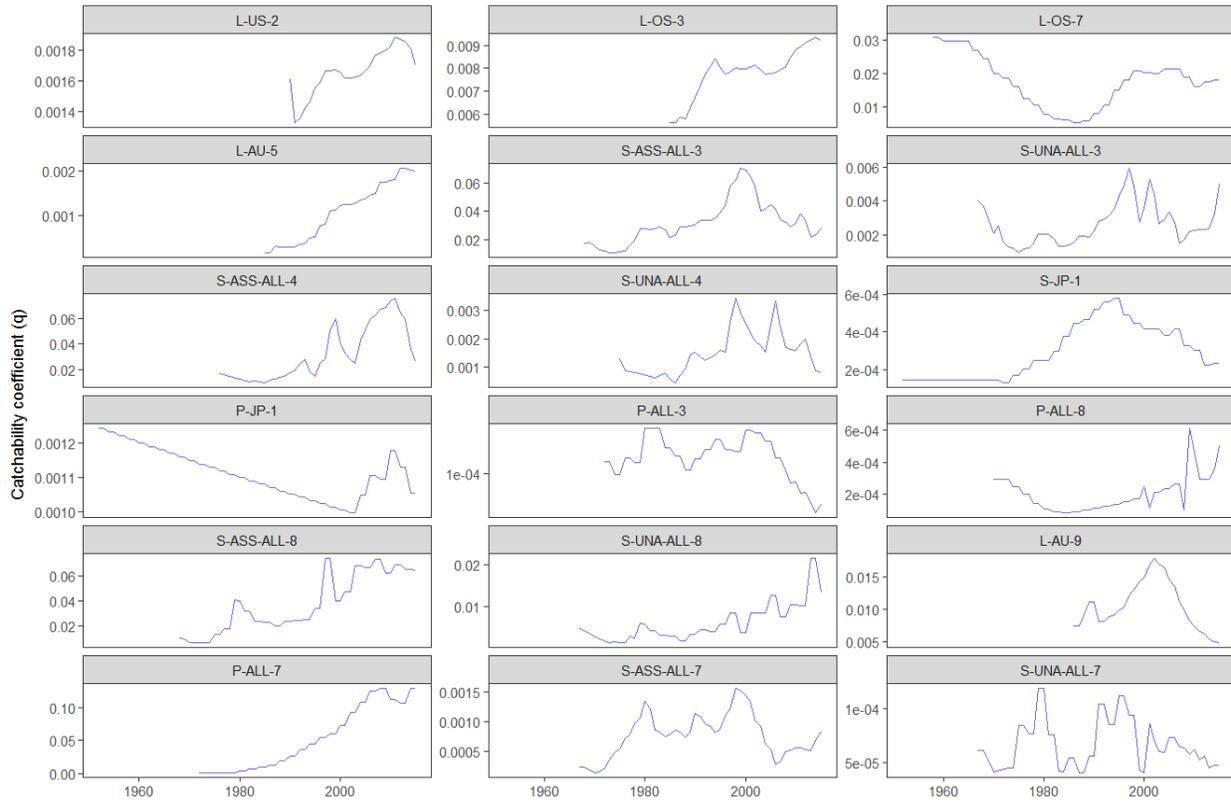
Figure 25: Estimated time series of catchability for those fisheries assumed to have random walk in these parameters. Values shown are the annual means which removes seasonal variability.
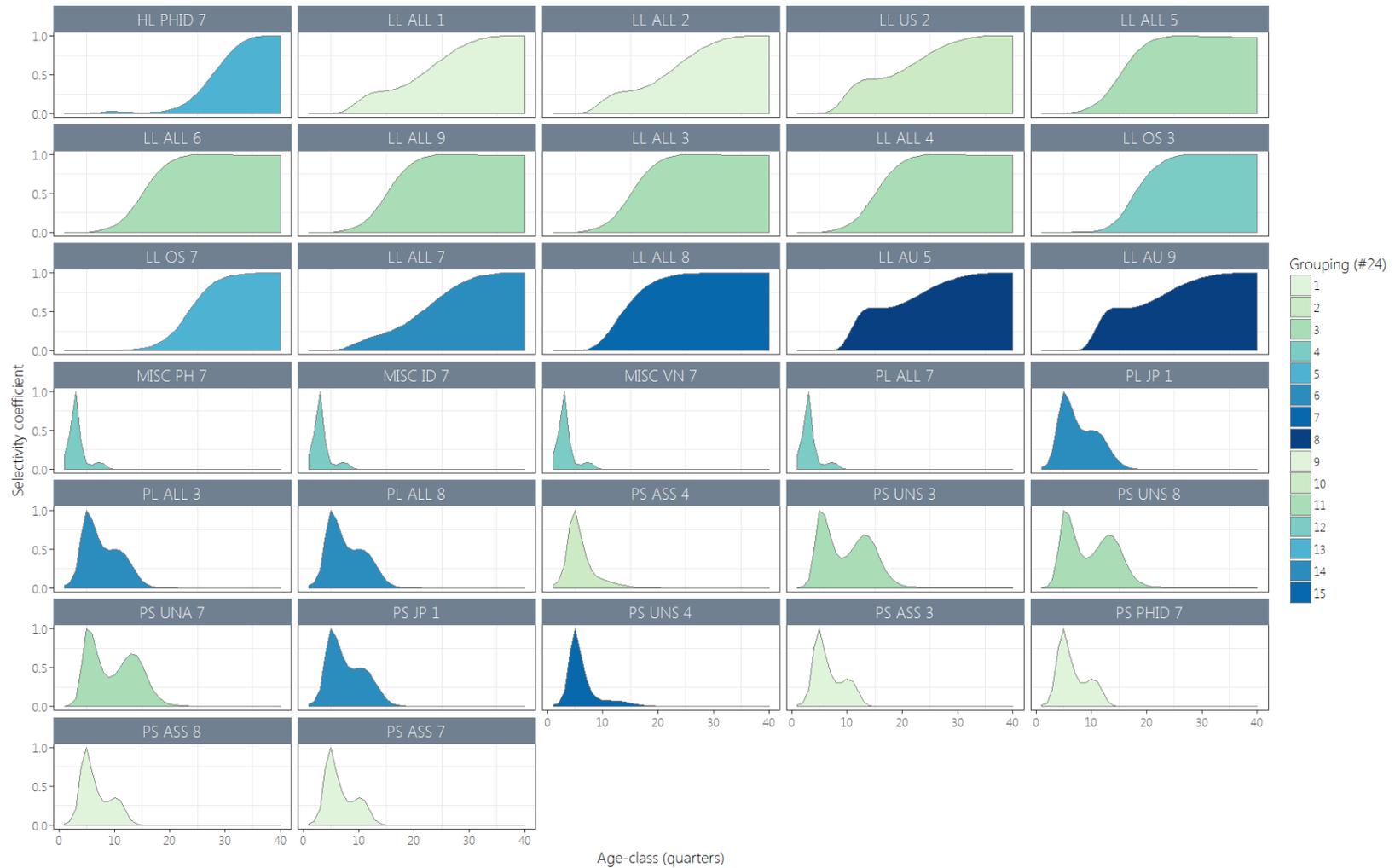
Figure 26: Estimated age-specific selectivity coefficients by fishery for the diagnostic case model. The colours indicate the groupings of the fisheries with respect to selectivity, such that fisheries sharing the same colour in the plot also share selectivity functions in the model.
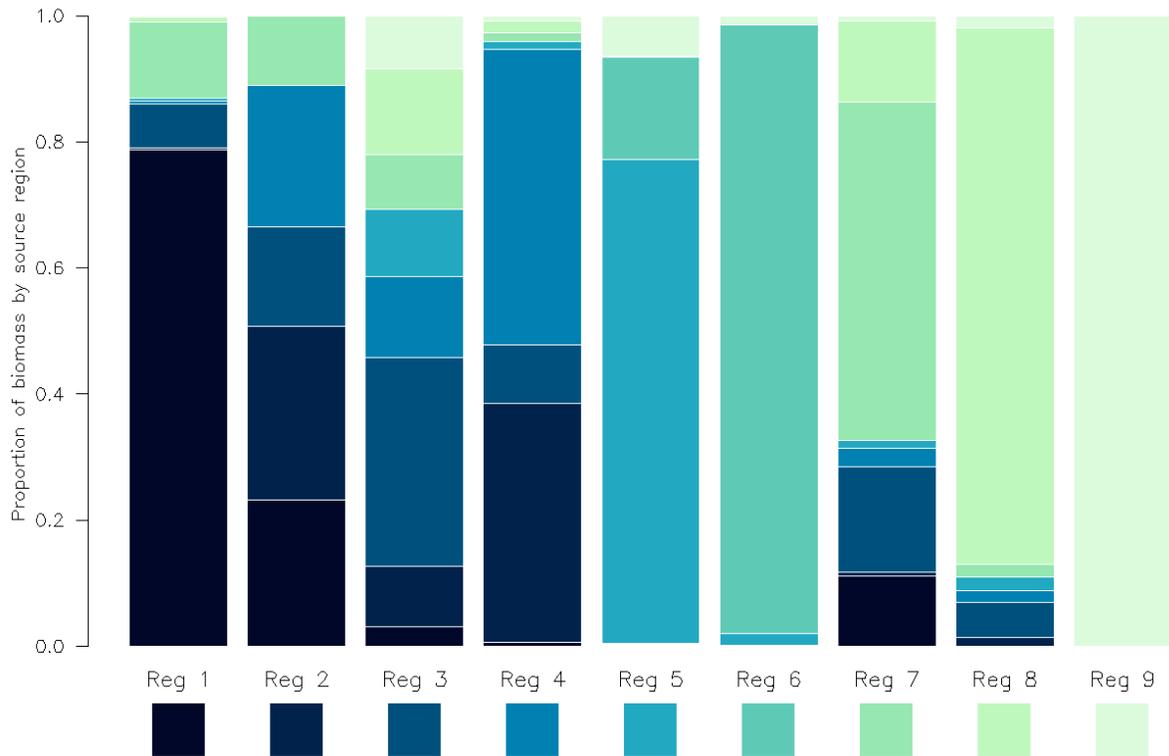
Figure 27: Proportional distribution of total biomass (by weight) in each region apportioned by the source region of the fish, for the diagnostic case model. The colour of the home region is presented below the corresponding label on the x-axis. The biomass distributions are calculated based on the long-term average distribution of recruitment between regions, estimated movement parameters, and natural mortality.
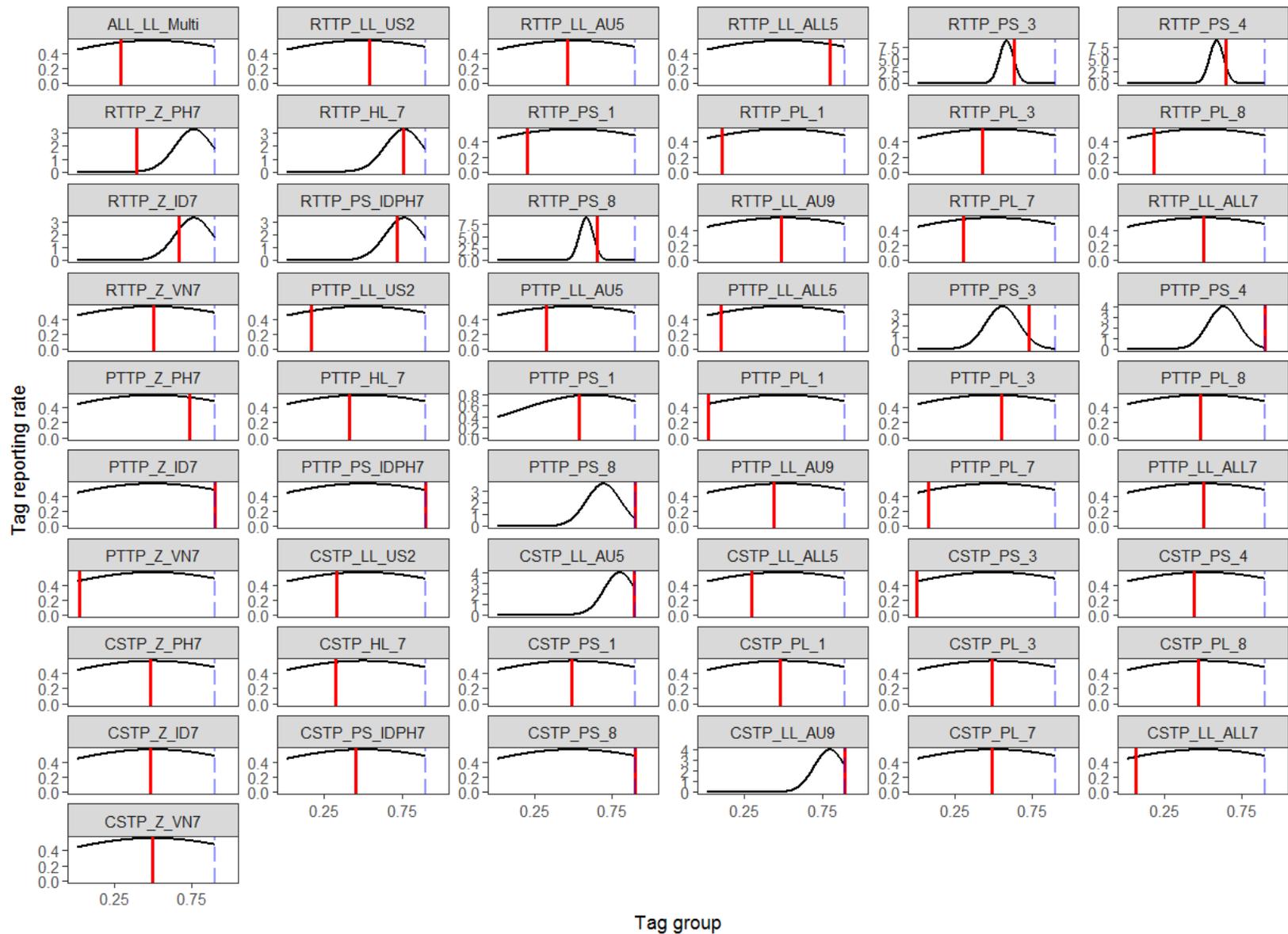
Figure 28: Estimated reporting rates for the diagnostic case model (red lines) and the prior distribution (black lines) for each reporting rate group. The imposed upper bound (0.9) on the reporting rate parameters are show as a blue dashed line. Reporting rates can be estimated separately for each release program and recapture fishery group but in practice are aggregated over some recapture groups to reduce dimensionality.
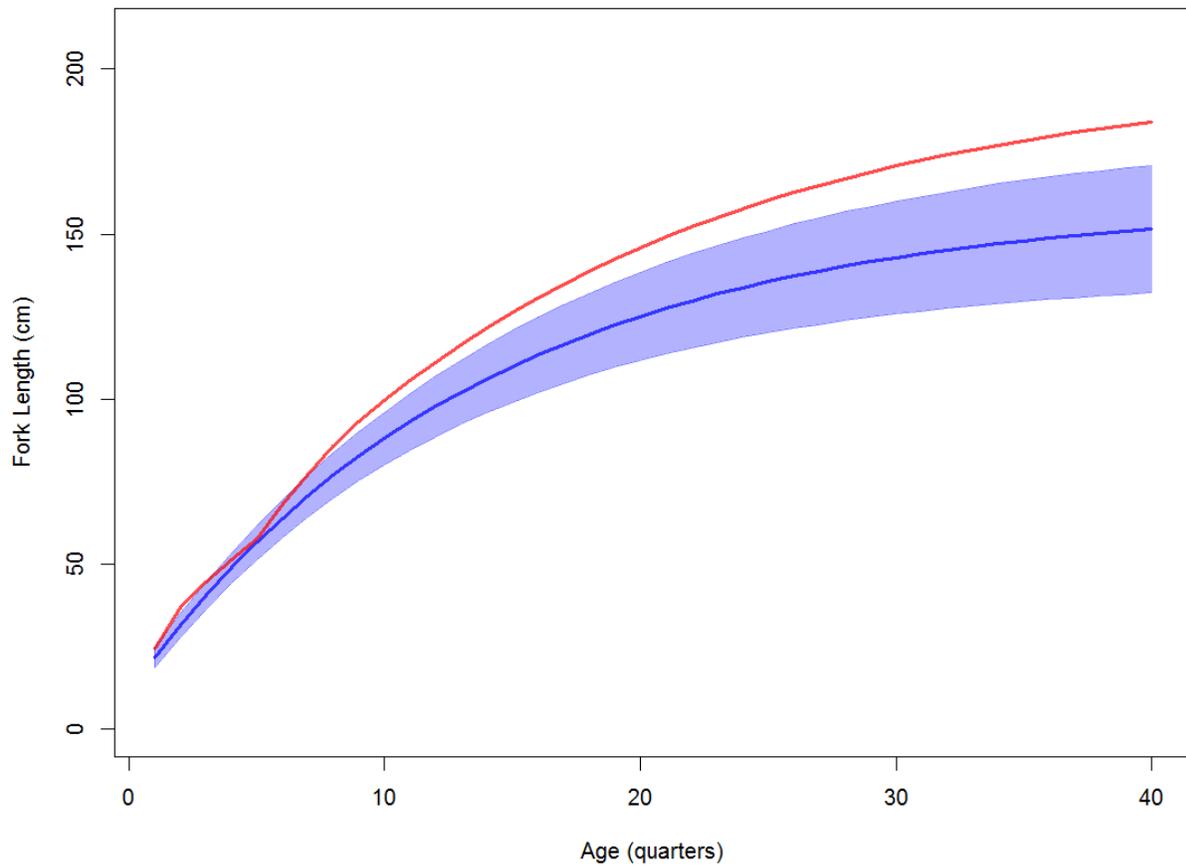
Figure 29: Estimated growth for the diagnostic case model and the sensitivity model *L2-184*. The blue line represents the estimated mean fork length (cm) at-age and the blue region represents the length-at-age within one standard deviation of the mean, for the diagnostic case model. The red line represents the estimated mean length-at-age for model *L2-184*.
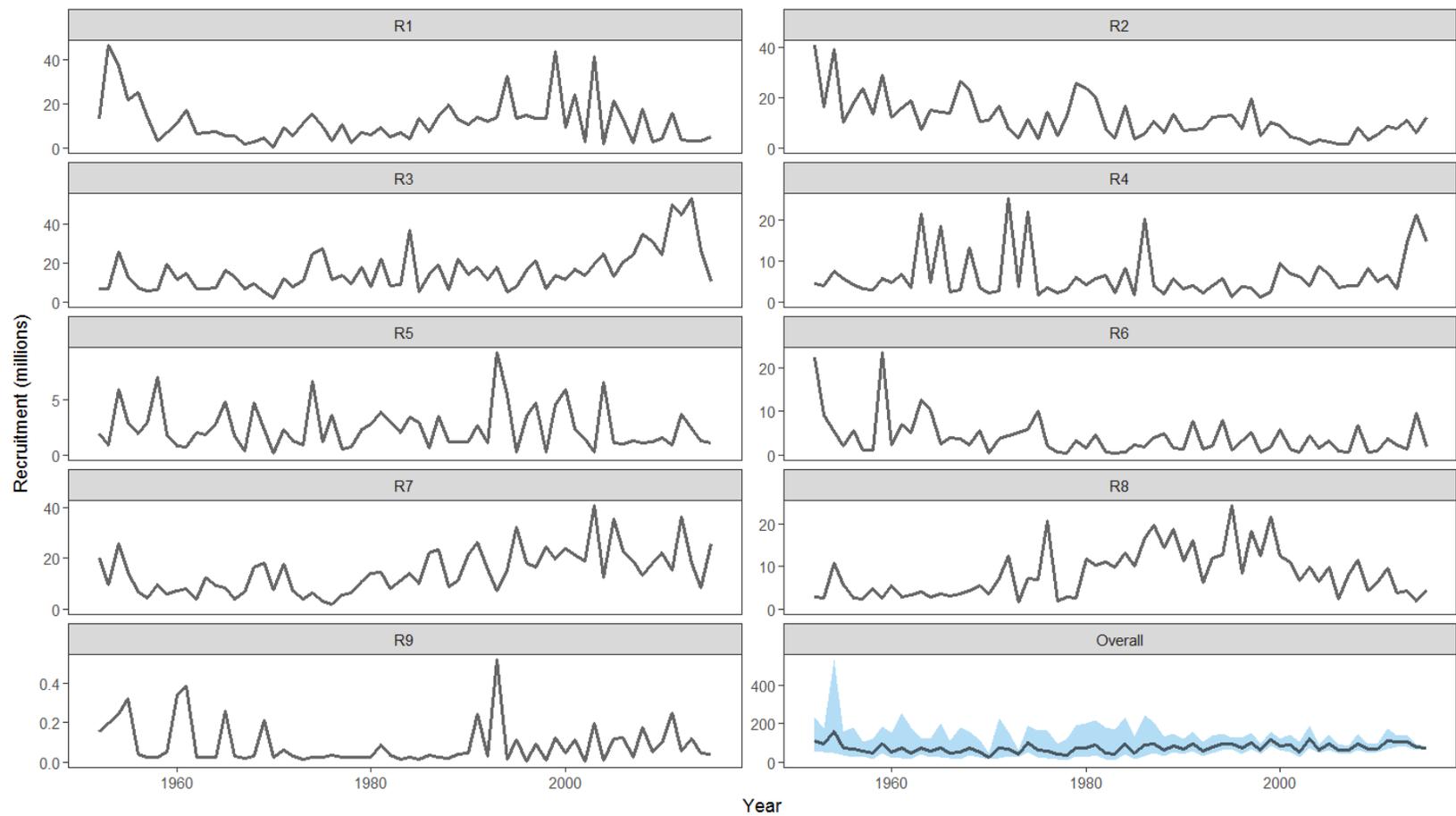
Figure 30: Estimated annual, temporal recruitment by model region for the diagnostic case model. The figure for the overall recruitments shows the estimated 95% confidence intervals as the blue shaded regions. Because of the wide uncertainty in the early years, the y-axis scale makes it difficult to interpret trends, but these can be seen more clearly in Figure 31. Note that the scale of the y-axis is not constant across regions.
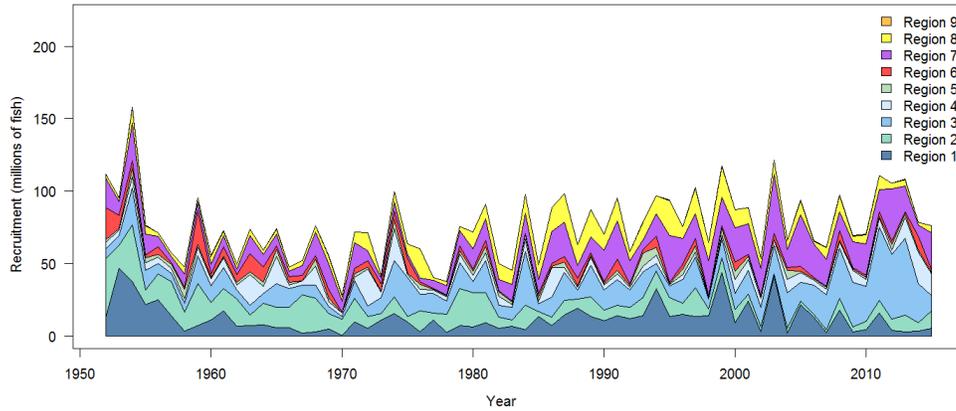
(a) Regional recruitment
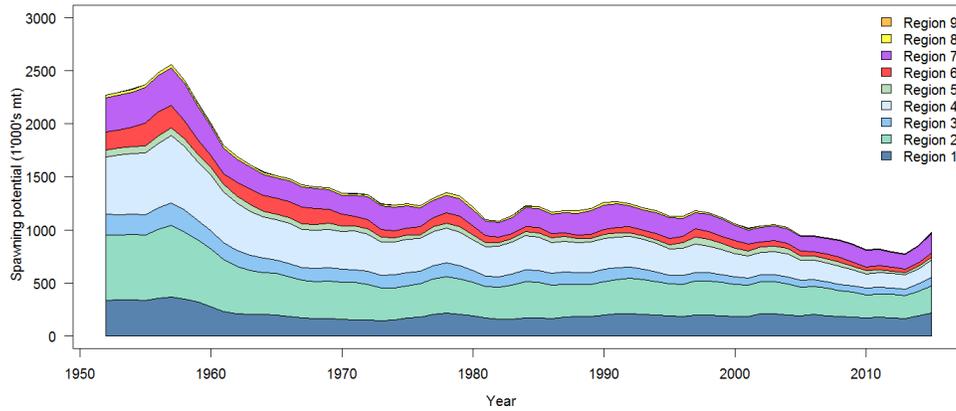
(b) Spawning potential
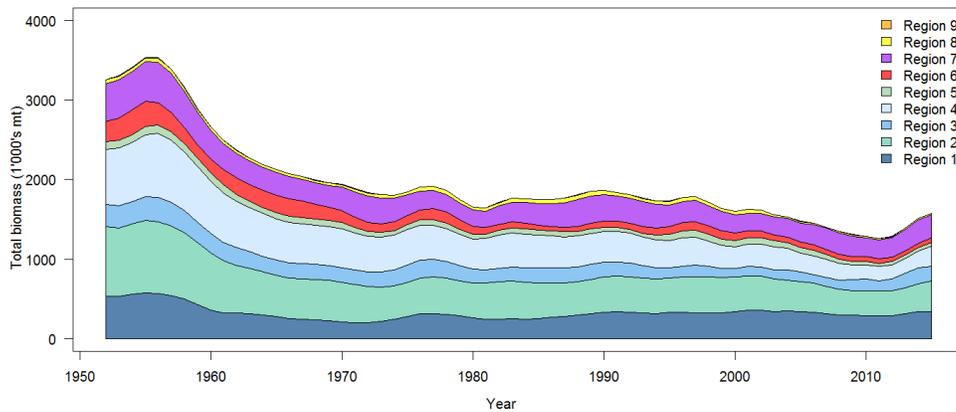
(c) Total biomass

Figure 31: Estimated annual average recruitment, spawning potential and total biomass by model region for the diagnostic case model, showing the relative sizes among regions.
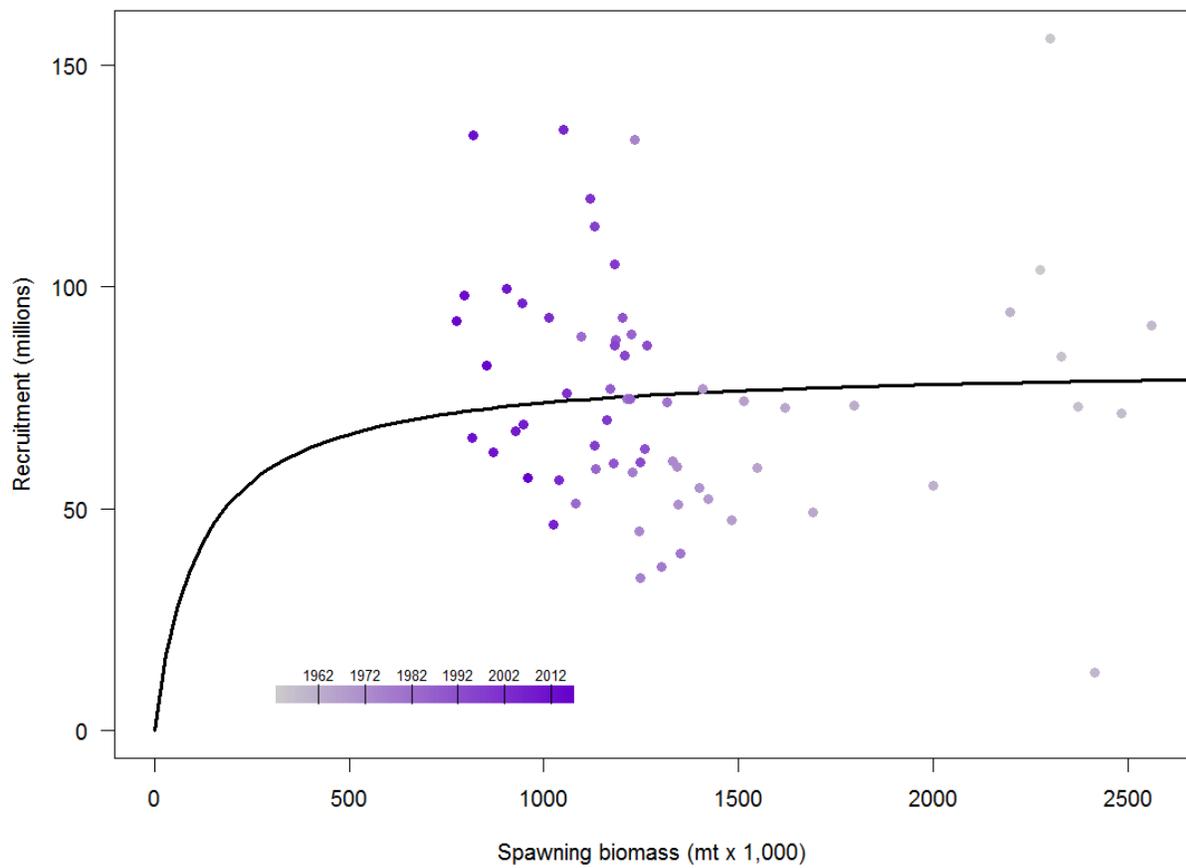
Figure 32: Estimated relationship between recruitment and spawning potential based on annual values for the diagnostic case model. The darkness of the circles changes from light to dark through time.
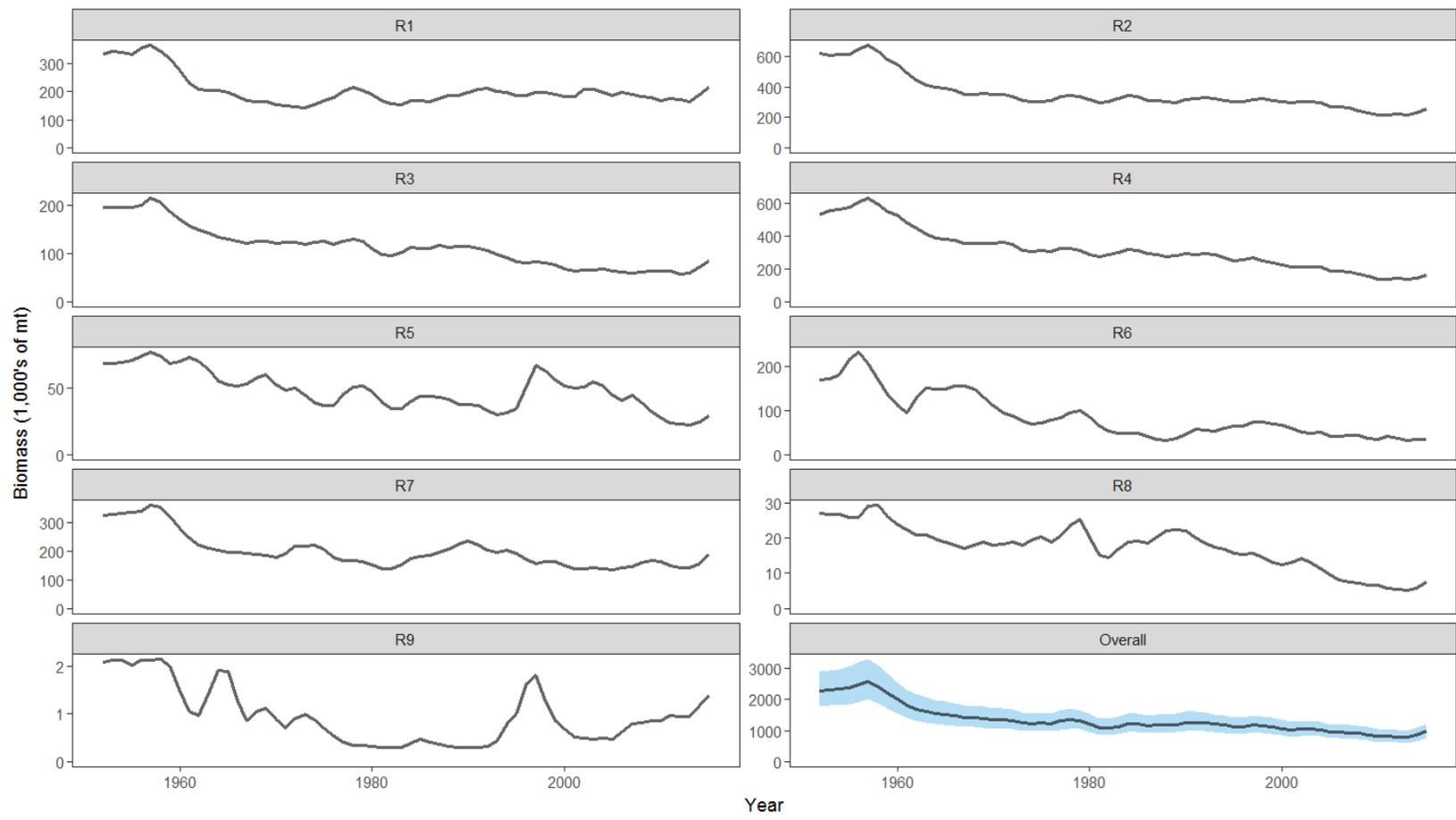
Figure 33: Estimated temporal spawning potential by model region for the diagnostic case model. The figure for the overall recruitments shows the estimated 95% confidence intervals as the blue shaded regions. Note that the scale of the y-axis is not constant across regions.
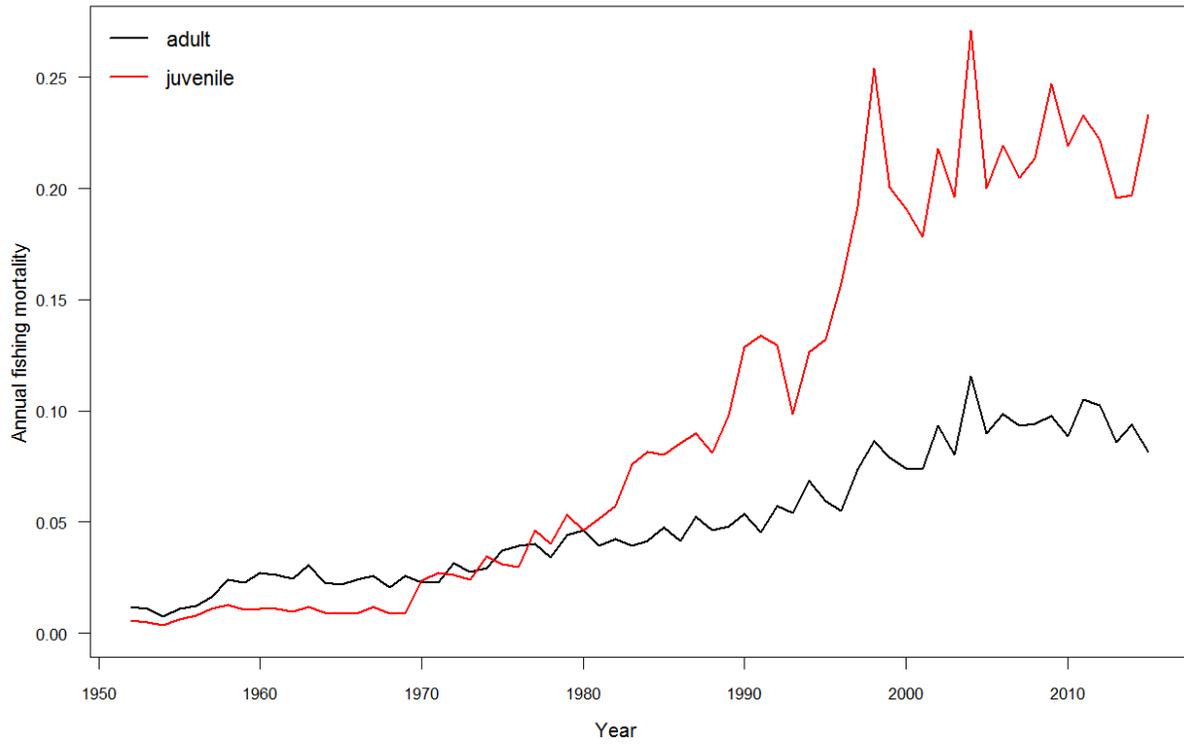
Figure 34: Estimated annual average juvenile and adult fishing mortality for the diagnostic case model.
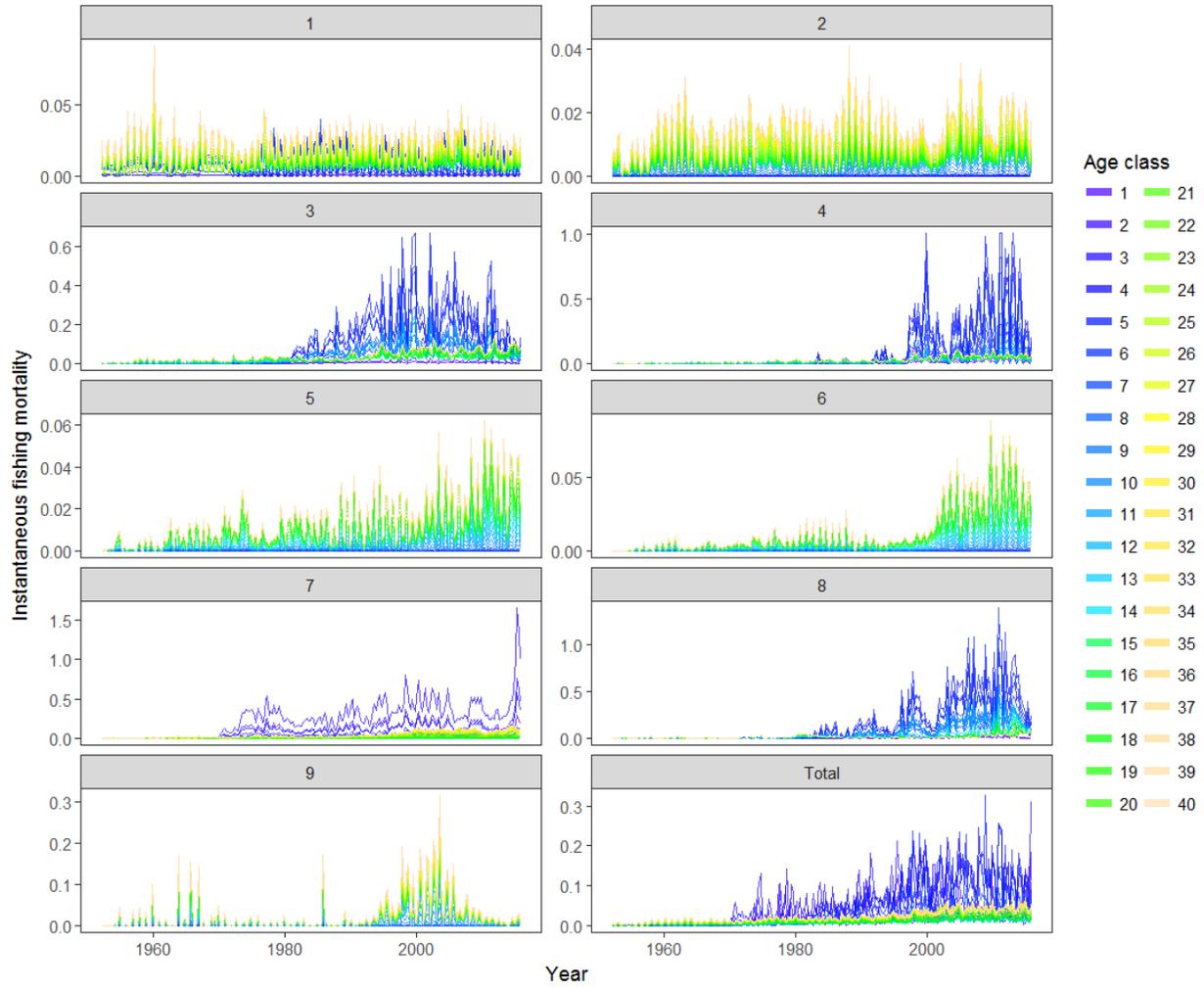
Figure 35: Estimated age-specific fishing mortality for the diagnostic case model, by region and overall.
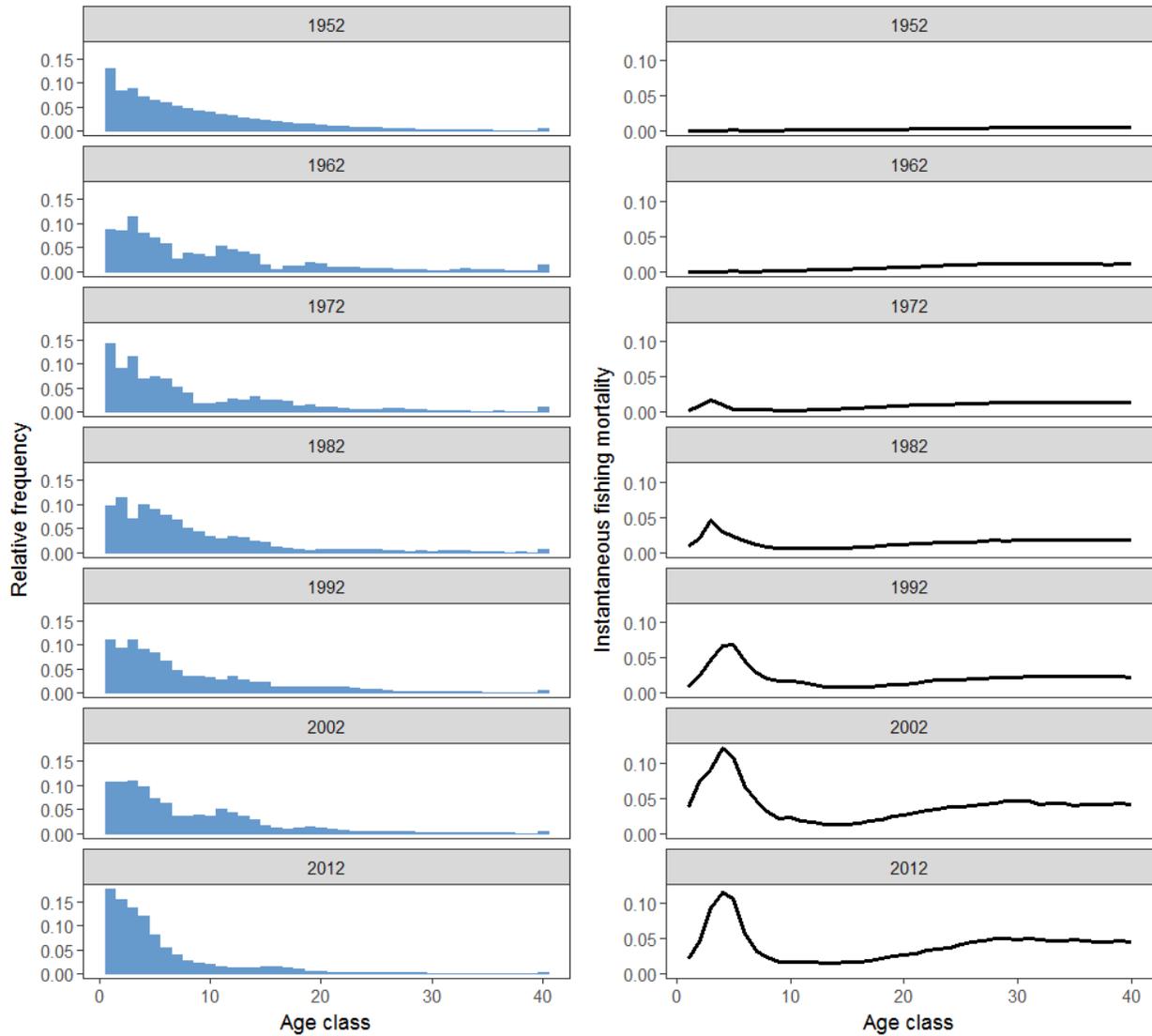
Figure 36: Estimated proportion of the population at-age (quarters; left panels) and fishing mortality-at-age (right panels), at decadal intervals, for the diagnostic case model.
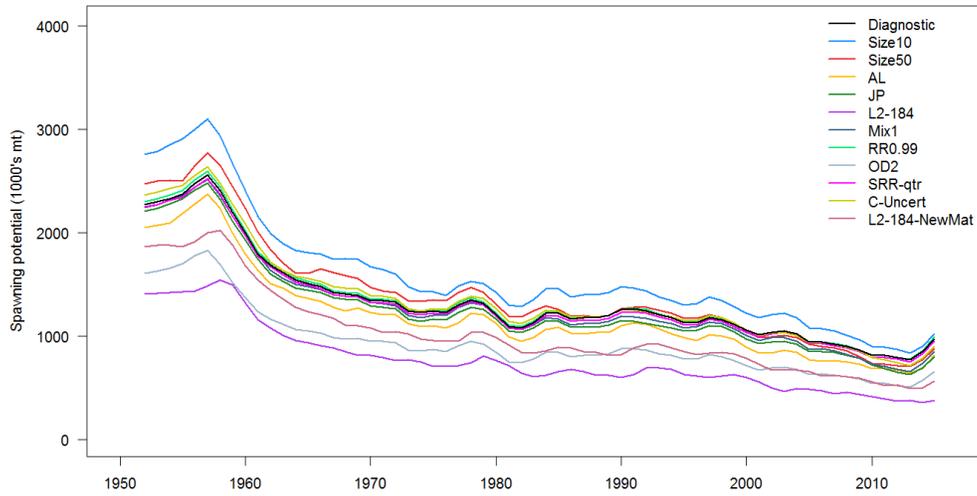
(a) First set of sensitivities
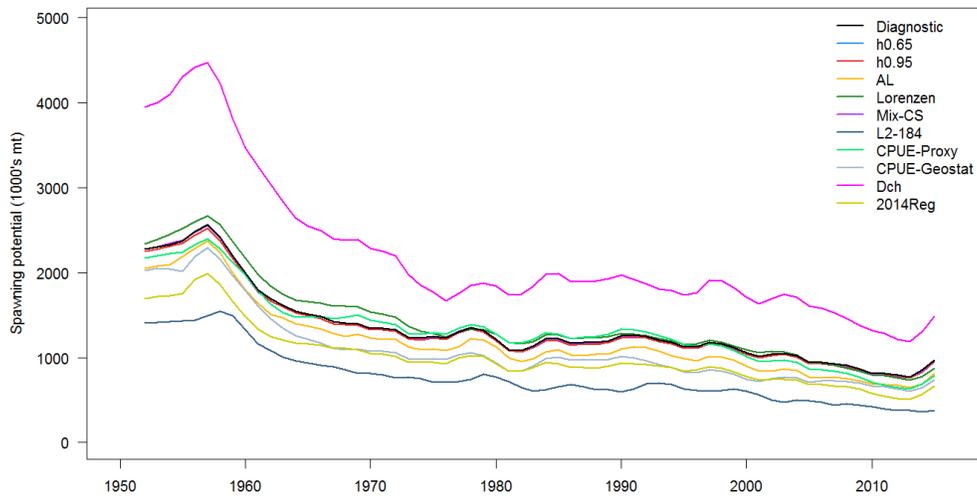
(b) Second set of sensitivities

Figure 37: Estimated spawning potential for each of the one-off sensitivity models investigated in the assessment. The models are separated into two groups, (a) and (b), to prevent obstruction of lines. The old growth sensitivity model (*L2-184*) is included in both sets for comparison. Details of the models can be found in Section 6.2.
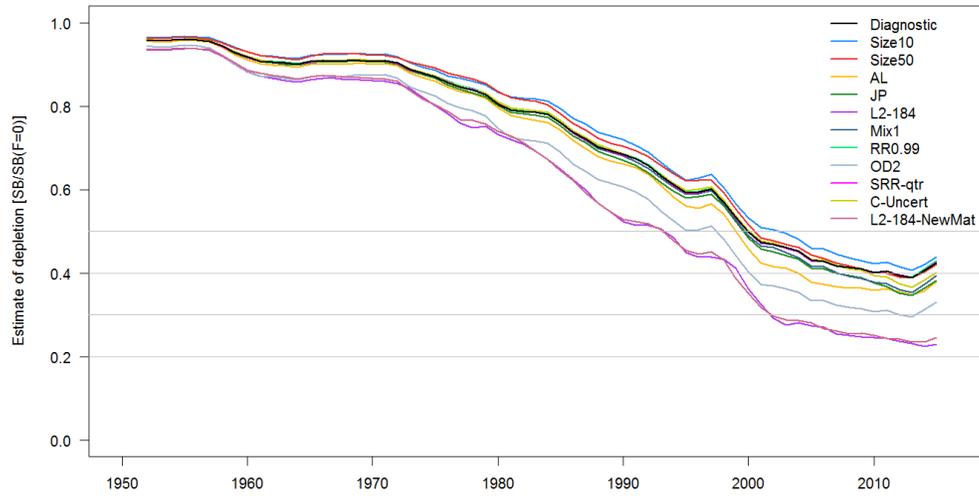
(a) First set of sensitivities
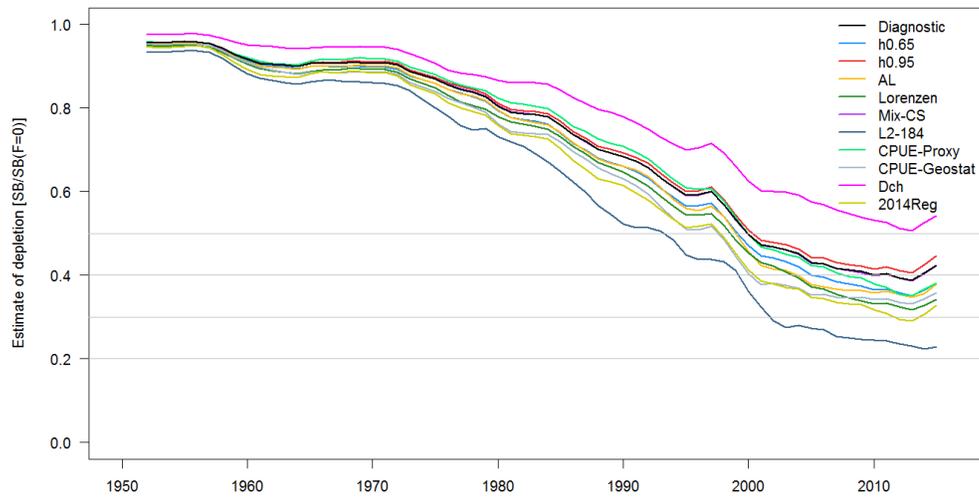
(b) Second set of sensitivities

Figure 38: Estimated fishing depletion (of spawning potential) for each of the one-off sensitivity models investigated in the assessment. The models are separated into two groups, (a) and (b), to prevent obstruction of lines. The old growth sensitivity model (*L2-184*) is included in both sets for comparison. Details of the models can be found in Section 6.2.
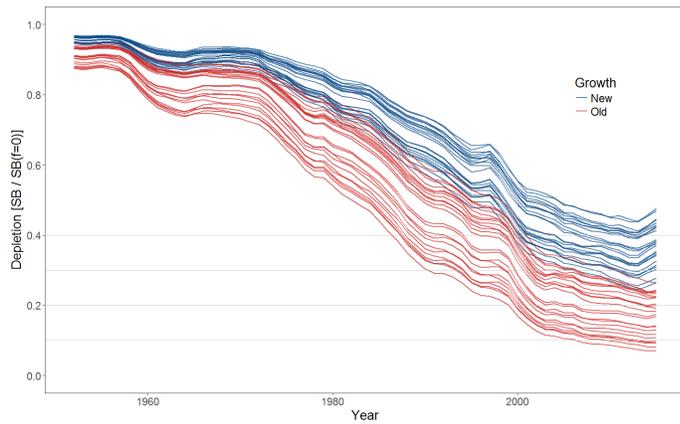
(a) Old and new growth
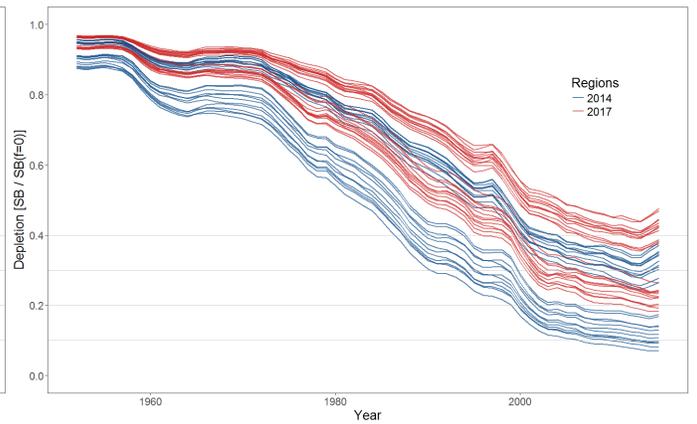
(b) 2014 and 2017 regions
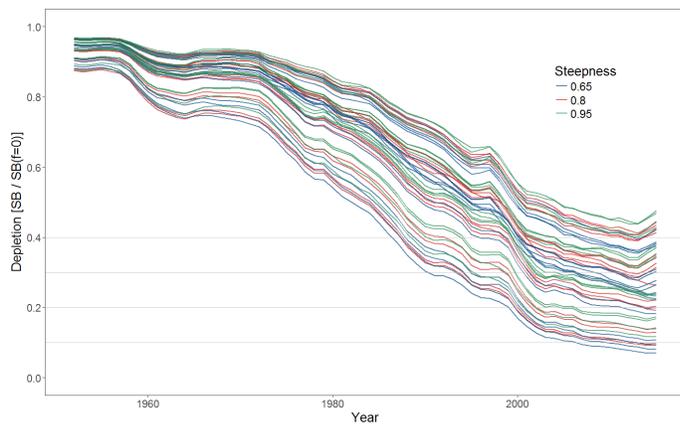
(c) Steepness
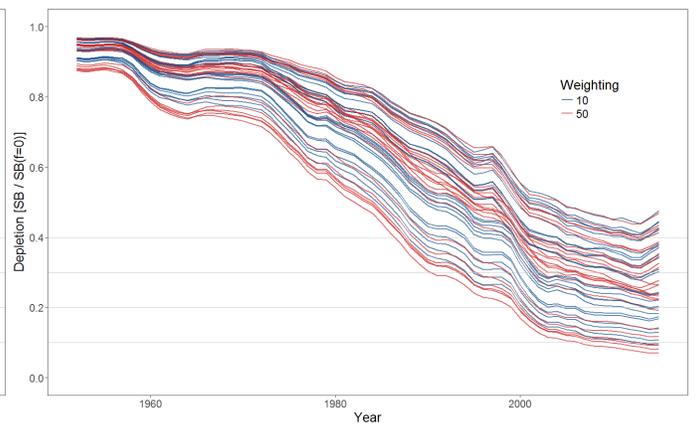
(d) Weighting of size data
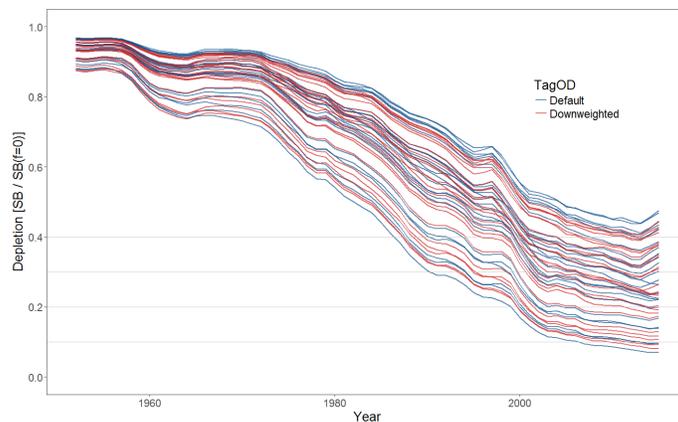
(e) Weighting of tagging data

Figure 39: Plots showing the trajectories of fishing depletion (of spawning potential) for the model runs included in the structural uncertainty grid (see Section 6.2 for details of the structure of the grid models). The five panels show the models separated on the basis of the five axes used in the grid, with the colour denoting the level within the axes for each model.
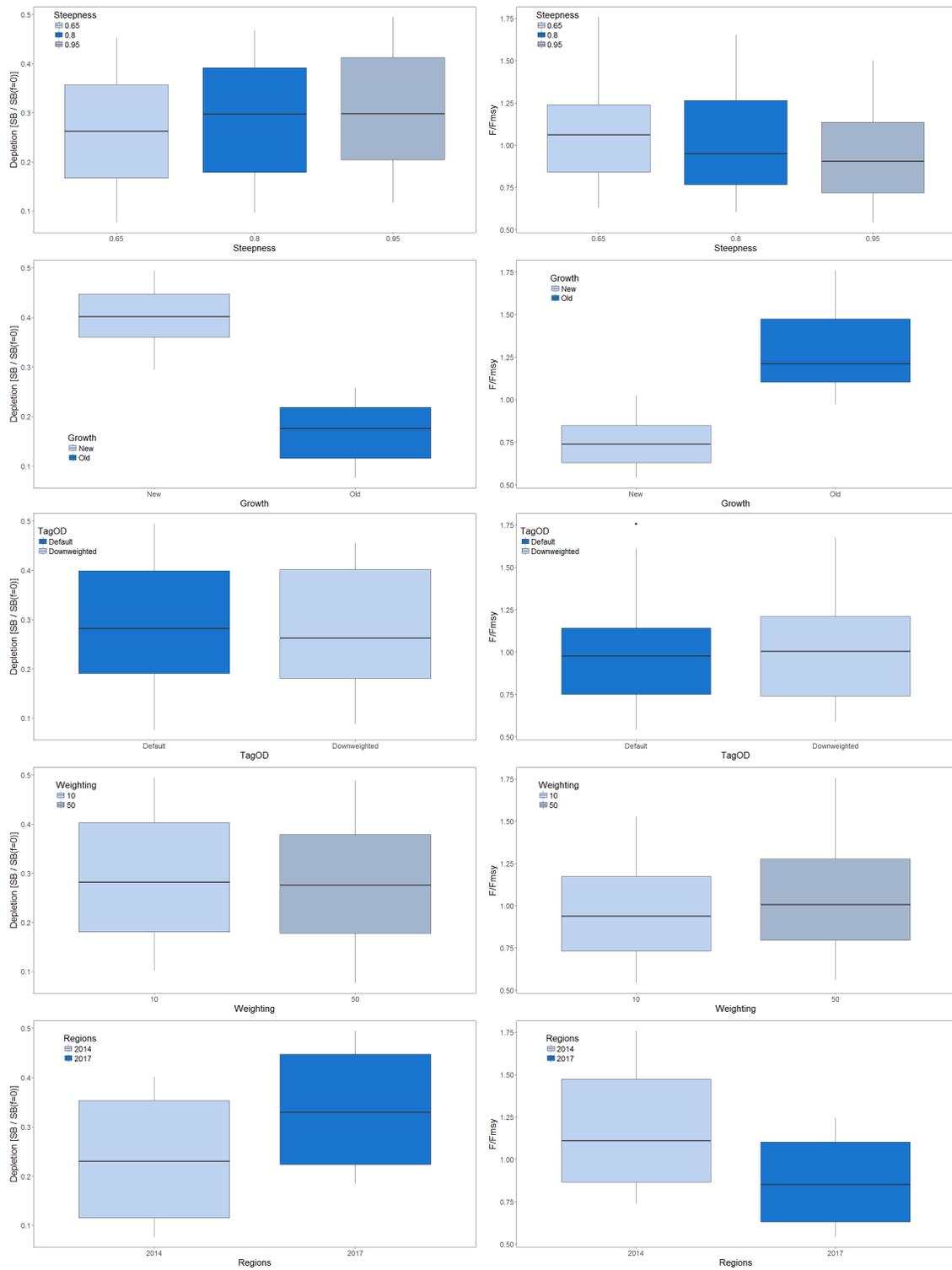
Figure 40: Boxplots summarising the results of the structural uncertainty grid with respect to the spawning potential reference point (left panels), and the fishing mortality reference point $F_{recent}/F_{\mathrm{MSY}}$ (right panels). The colours indicate the level of the model with respect to each uncertainty axis.
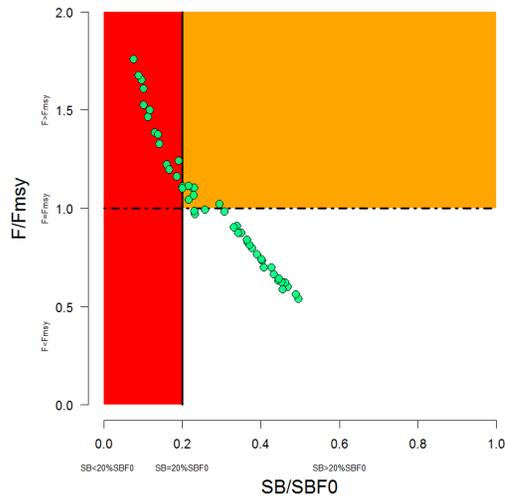
(a) All grid models SB_{latest}
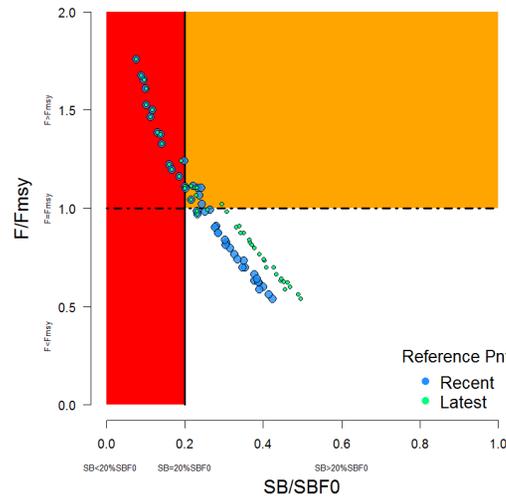
(b) All grid models SB_{latest} and SB_{recent}
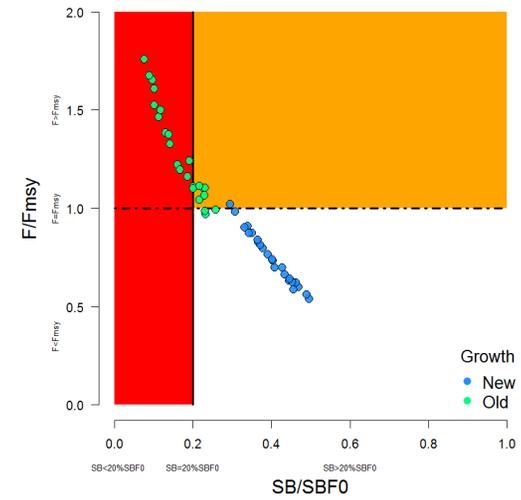
(c) Old and new growth
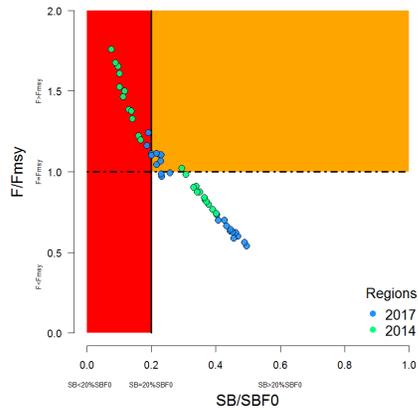
(d) 2014 and 2017 regions
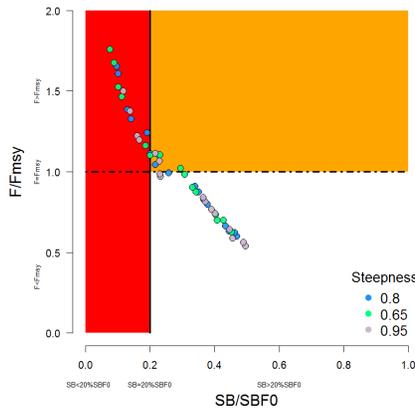
(e) Steepness
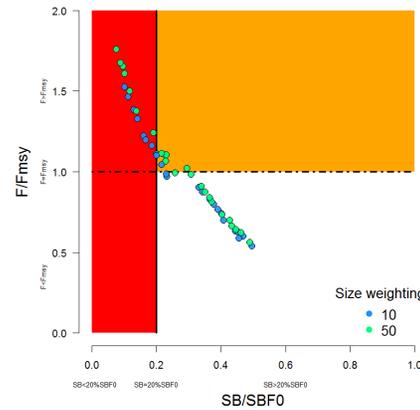
(f) Weighting of size data
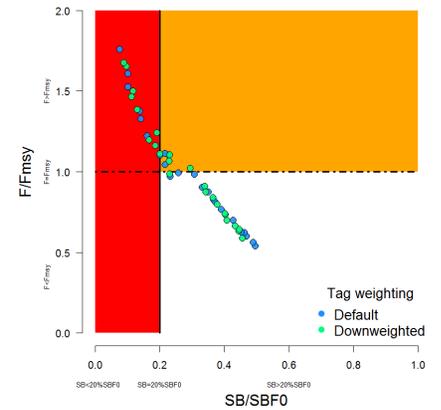
(g) Tagging data weighting

Figure 41: Majuro plots summarising the results for each of the models in the structural uncertainty grid. The plots represent estimates of 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 points represent $SB_{latest}/SB_{F=0}$ for each model run except in panel (b) where $SB_{recent}/SB_{F=0}$ is also displayed. Panels (c)–(g) show the estimates for the different levels for the five axes of the grid.
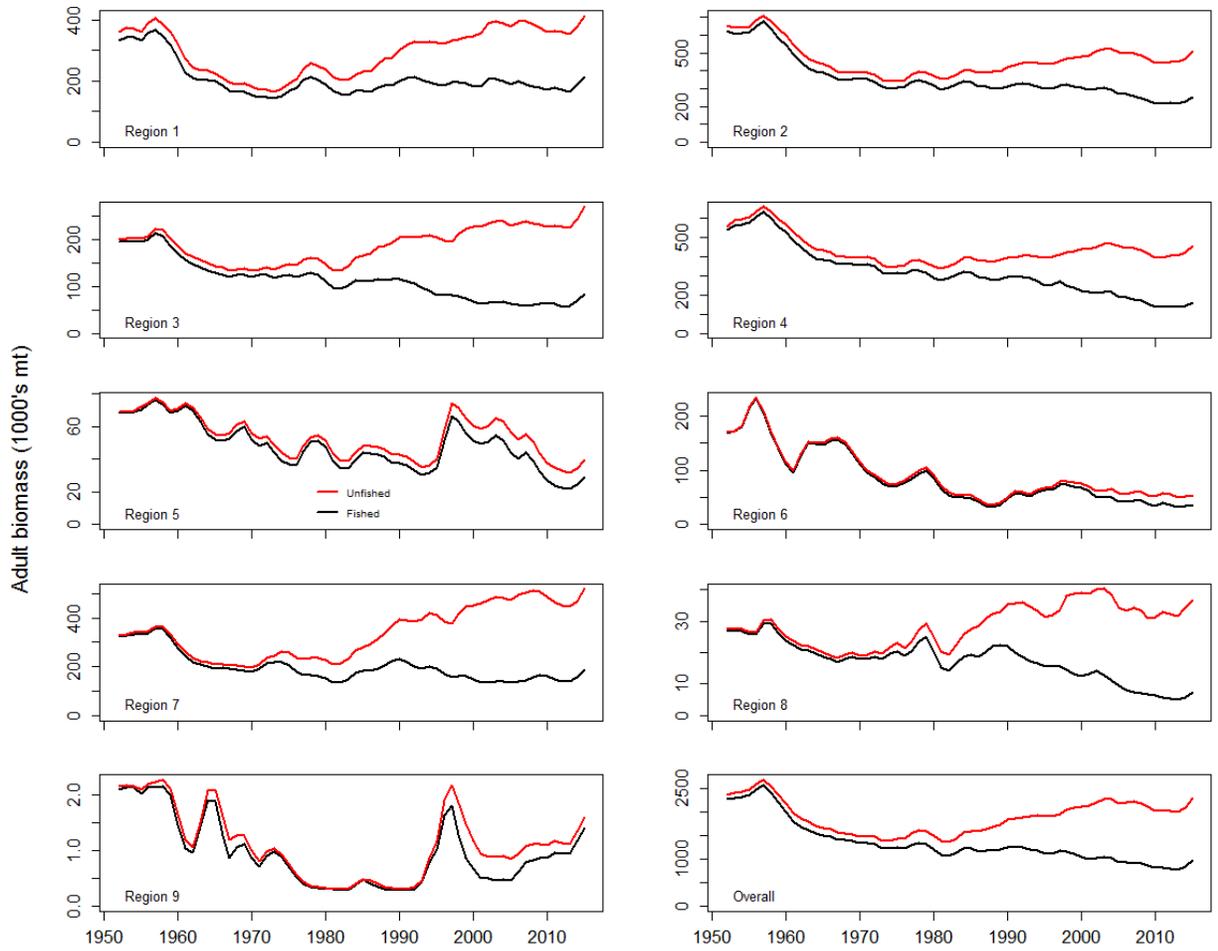
Figure 42: Comparison of the estimated annual spawning potential trajectories (lower solid black lines) with those trajectories that would have occurred in the absence of fishing (upper dashed red lines) for each region, and overall, for the diagnostic case model.
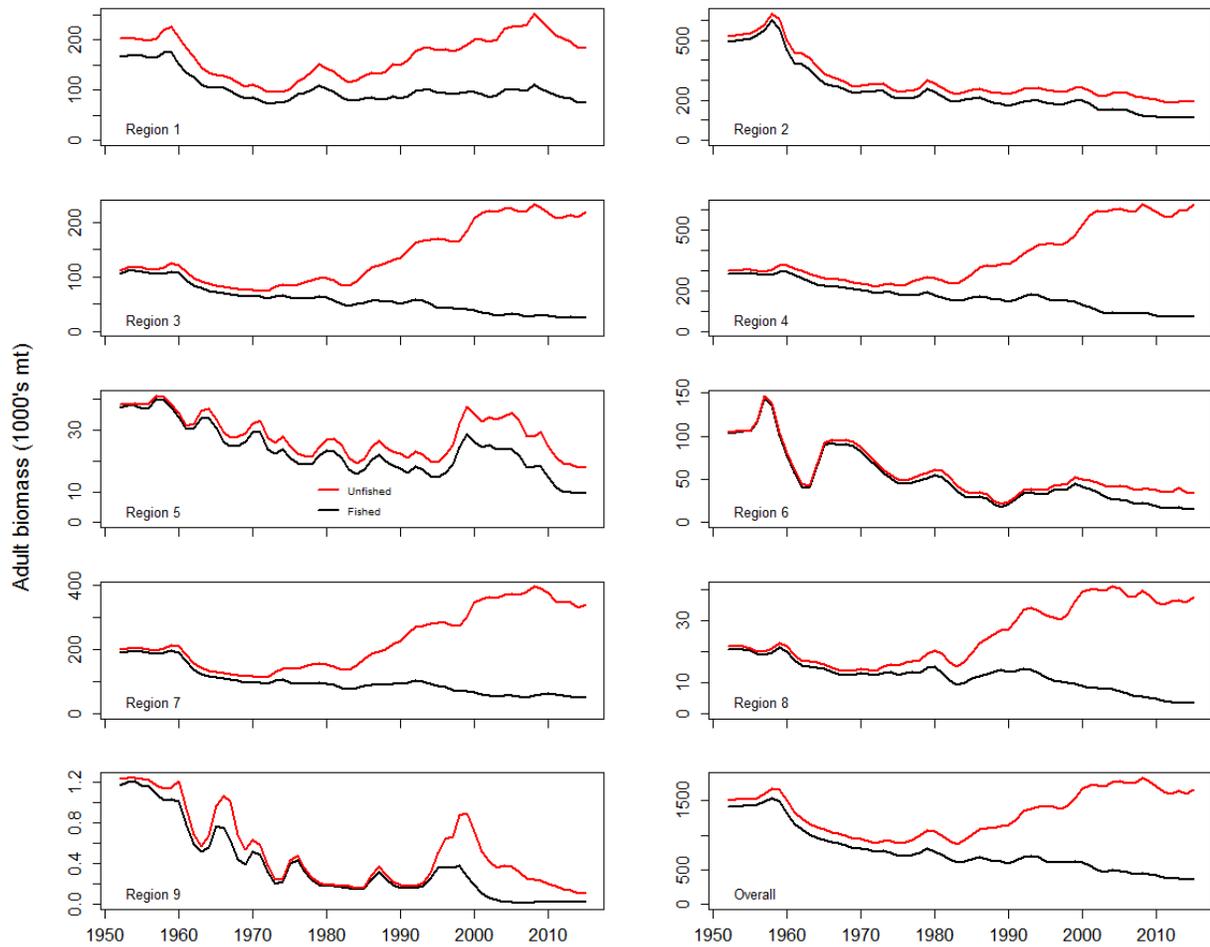
Figure 43: Comparison of the estimated annual spawning potential trajectories (lower solid black lines) with those trajectories that would have occurred in the absence of fishing (upper dashed red lines) for each region, and overall, for the sensitivity model *L2-184*.
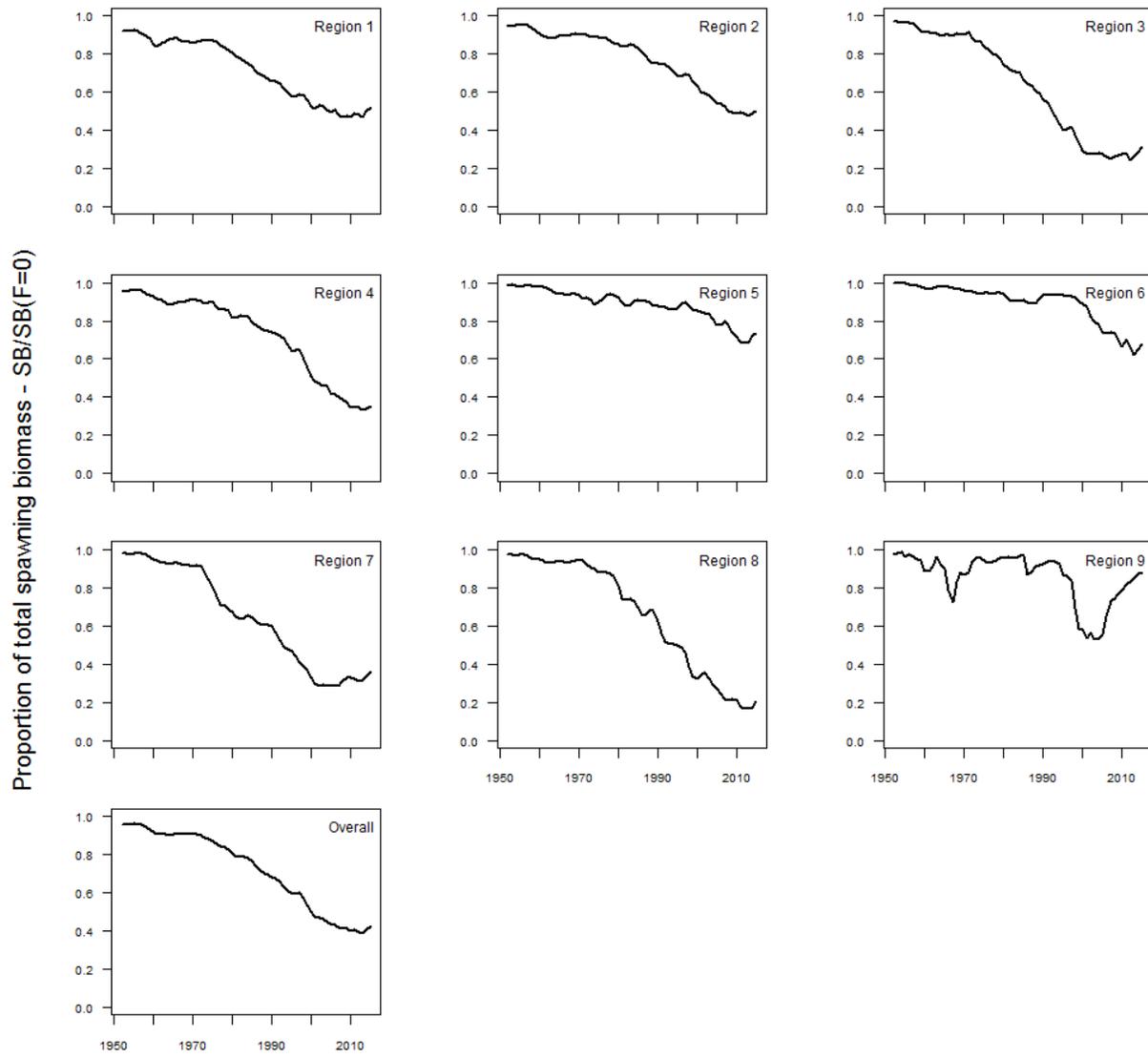
Figure 44: Ratio of exploited to unexploited spawning potential, $SB_{latest}/SB_{F=0}$, for each region, and overall, for the diagnostic case model.
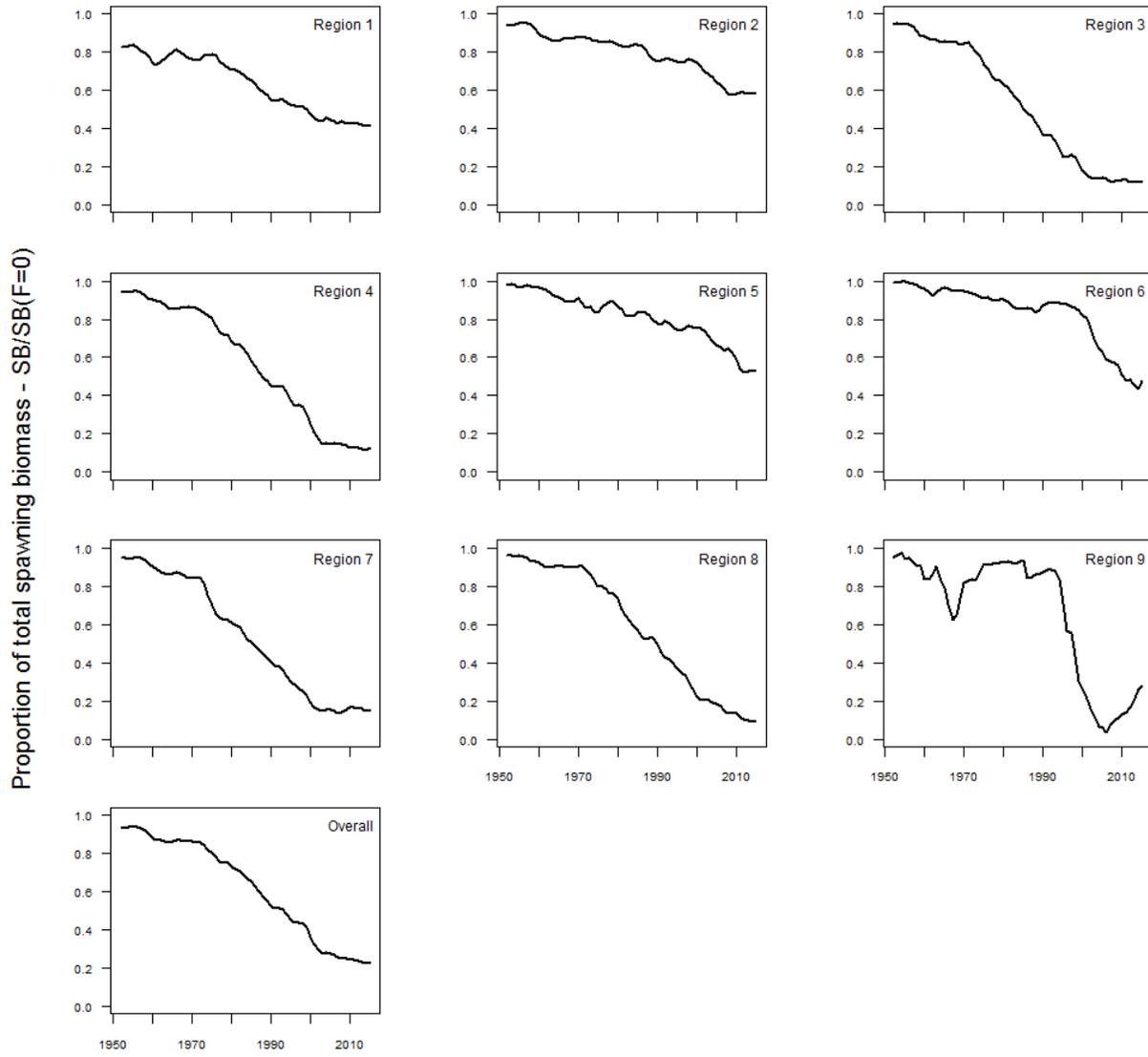
Figure 45: Ratio of exploited to unexploited spawning potential, $SB_{latest}/SB_{F=0}$, for each region, and overall, for the sensitivity model *L2-184*.
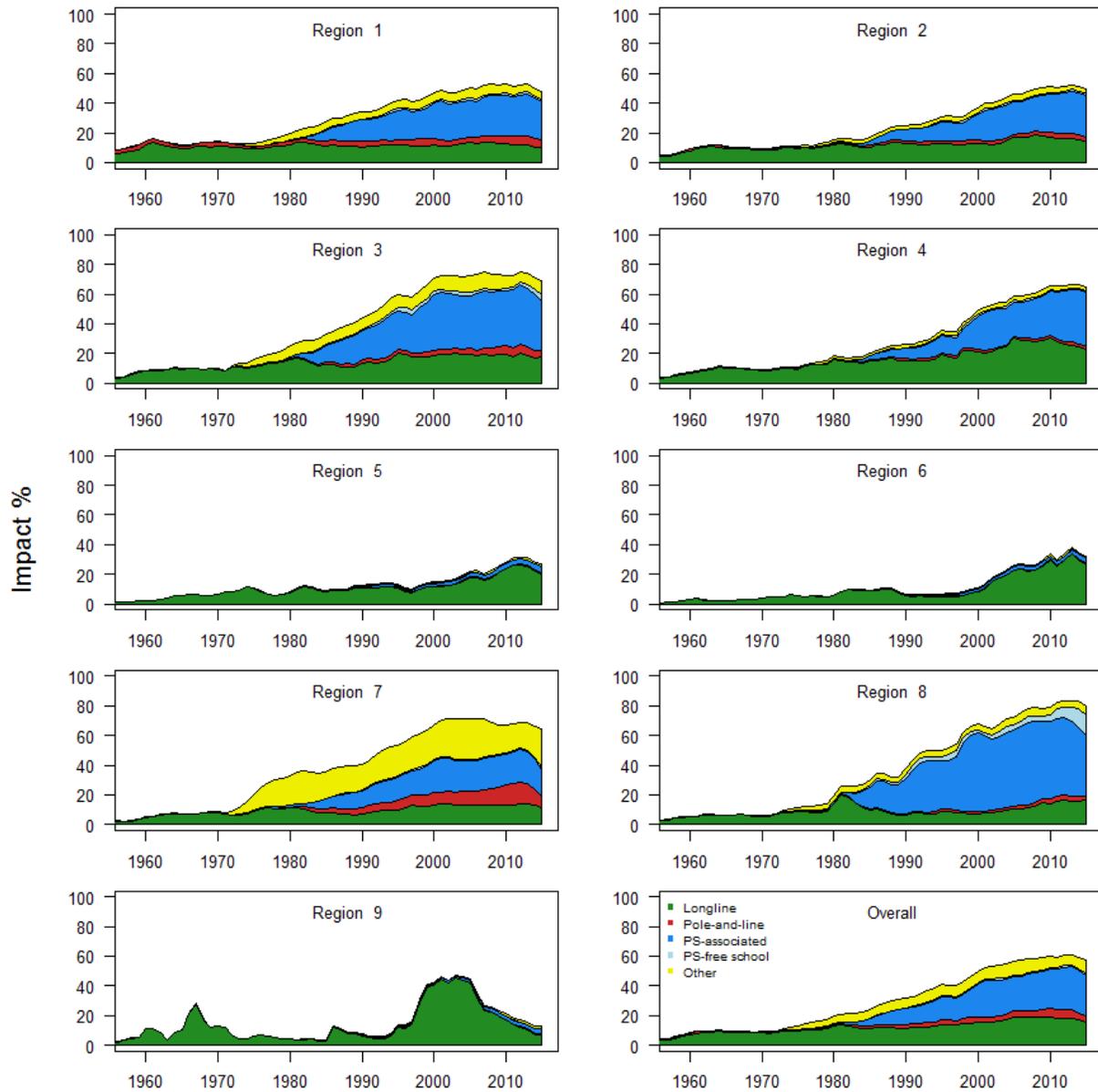
Figure 46: Estimates of reduction in spawning potential due to fishing (fishery impact = $1 - SB_{latest}/SB_{F=0}$) by region, and over all regions (lower right panel), attributed to various fishery groups for the diagnostic case model.
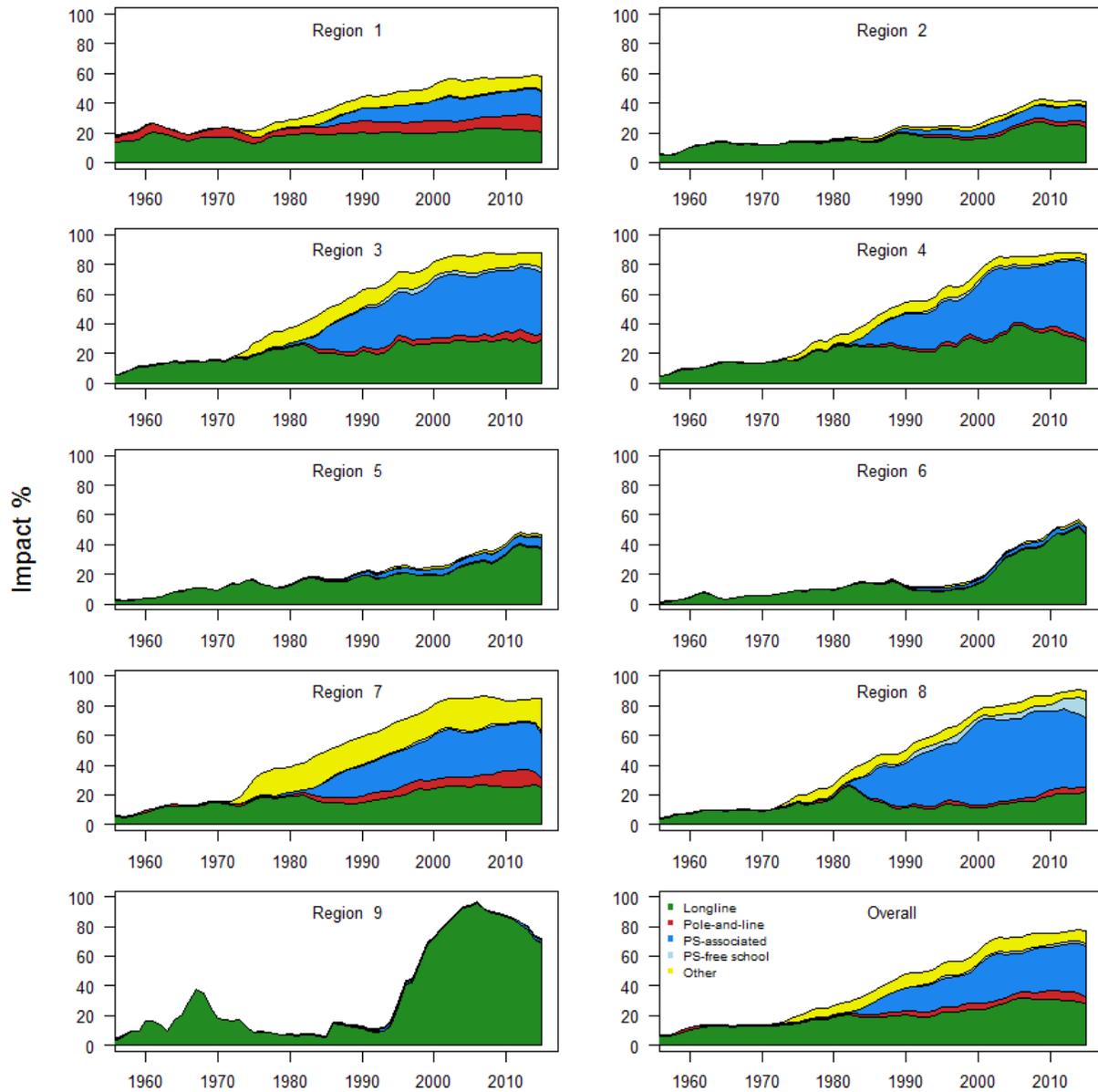
Figure 47: Estimates of reduction in spawning potential due to fishing (fishery impact $= 1 - SB_{latest}/SB_{F=0}$) by region, and over all regions (lower right panel), attributed to various fishery groups for the sensitivity model $L2\text{-}184$, which provides an example of fishery impacts for a model with old estimated growth and the 2017 regional structure.
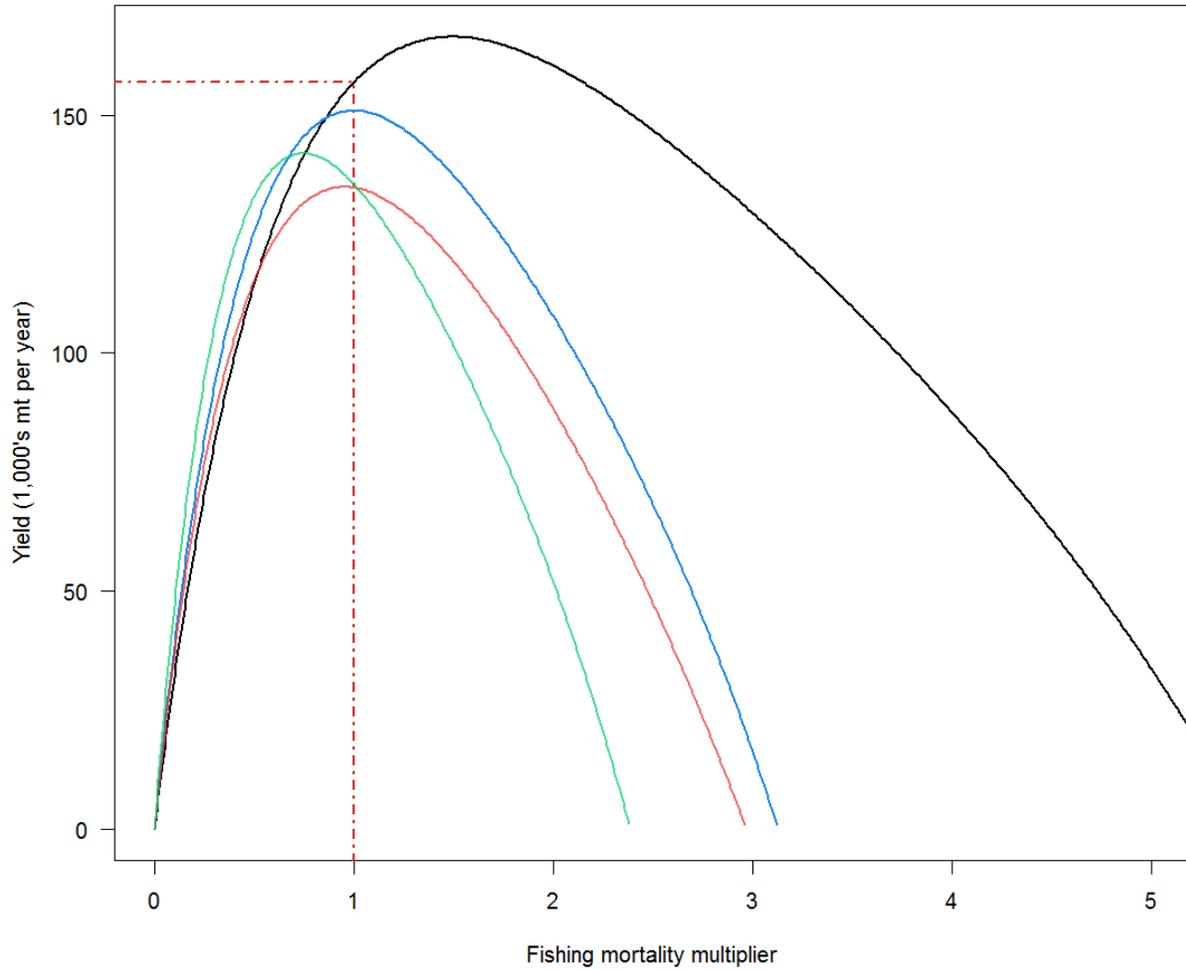
Figure 48: Estimated yield as a function of fishing mortality multiplier for four example models. The black line is the estimated yield curve for the diagnostic case model (representing new growth/2017 regions) and the red dashed line indicates the equilibrium yield at current fishing mortality. The other models displayed are sensitivity models *L2-184* (blue line; old growth/2017 regions), *2014Reg* (red line; new growth/2014 regions) and grid model *A0B1C0D0E1* (green line; old growth/2014 regions).
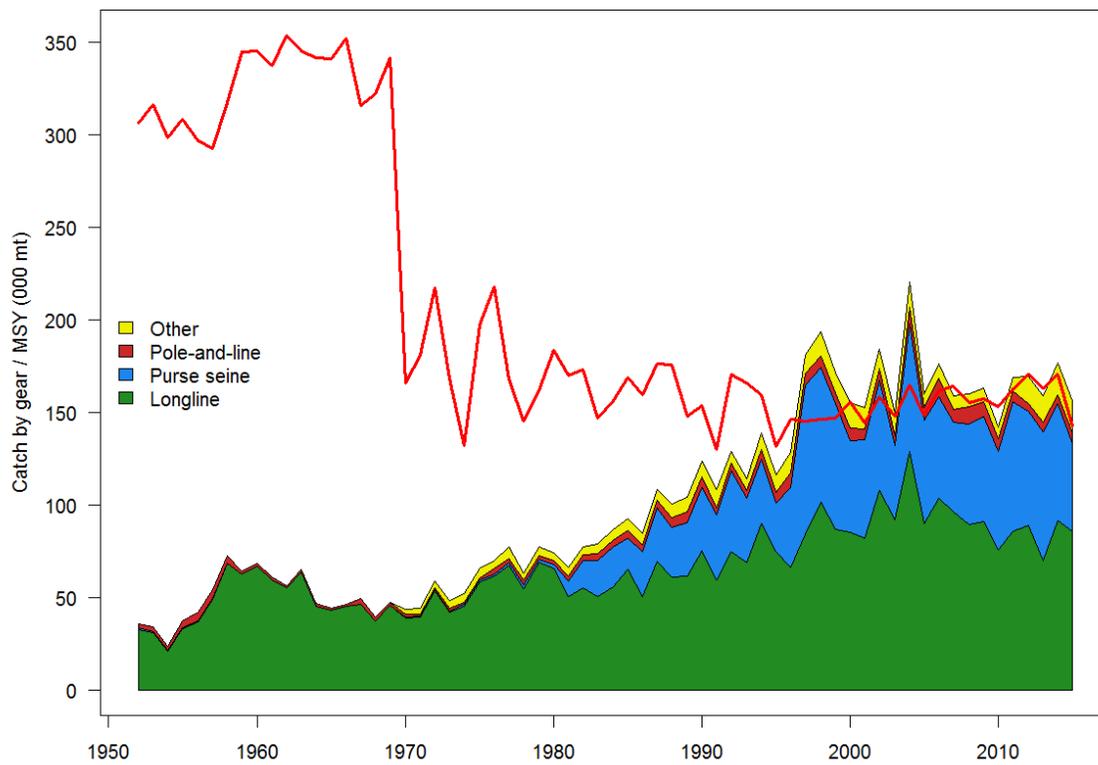
Figure 49: History of the annual estimates of MSY (red line) for the diagnostic case model compared with annual catch by the main gear types. Note that this is a "dynamic" MSY which is explained further in Section 5.7.4.
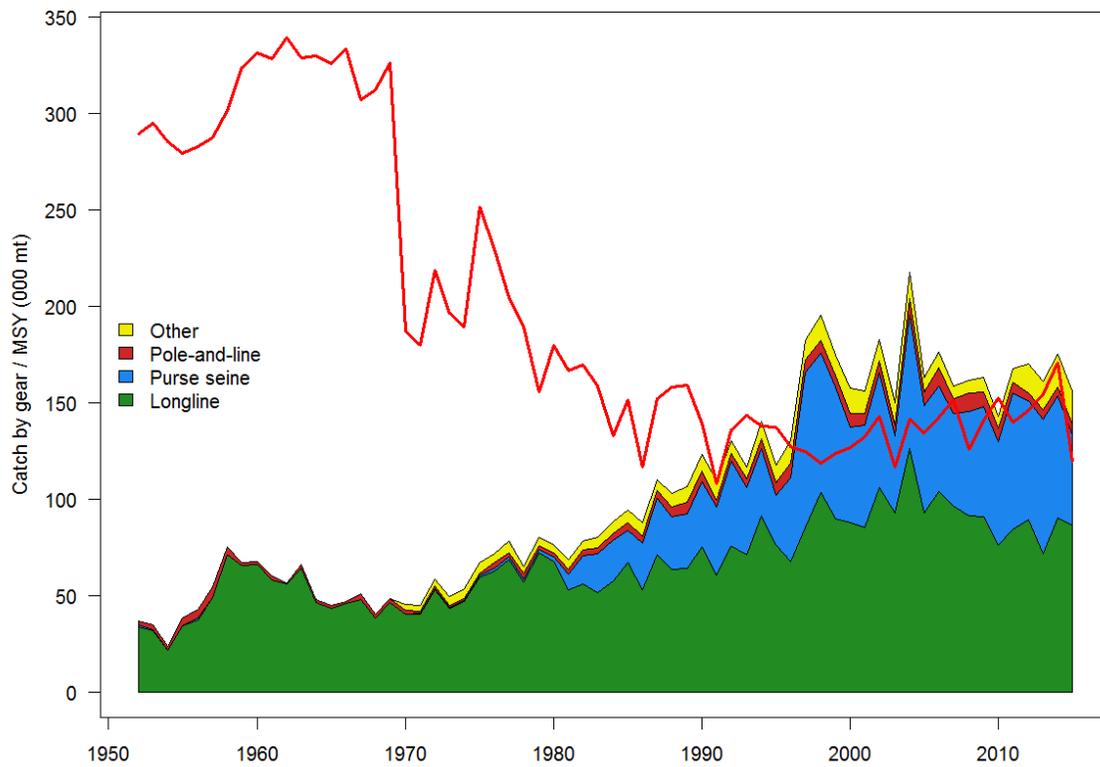
Figure 50: History of the annual estimates of MSY (red line) for the sensitivity model *L2-184* compared with annual catch by the main gear types. Note that this is a "dynamic" MSY which is explained further in Section 5.7.4.
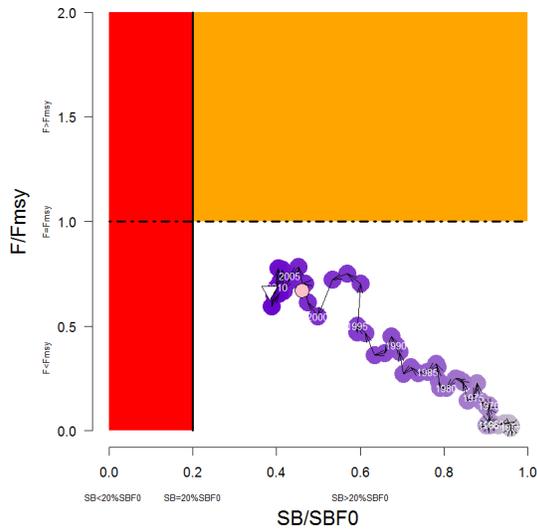
(a) Example; new growth/2017 regions (Diagnostic case)
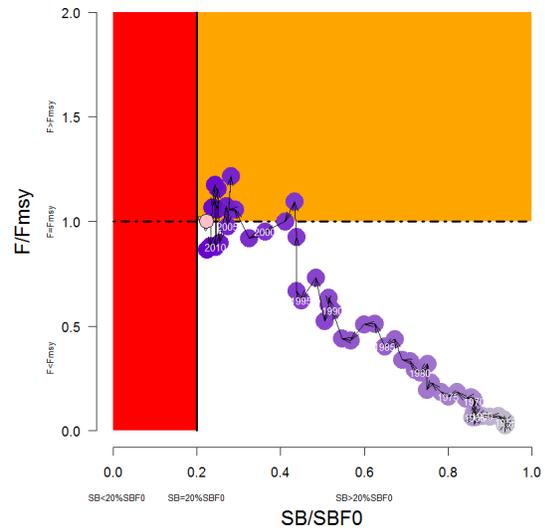
(b) Example; old growth/2017 regions (*L2-184*)
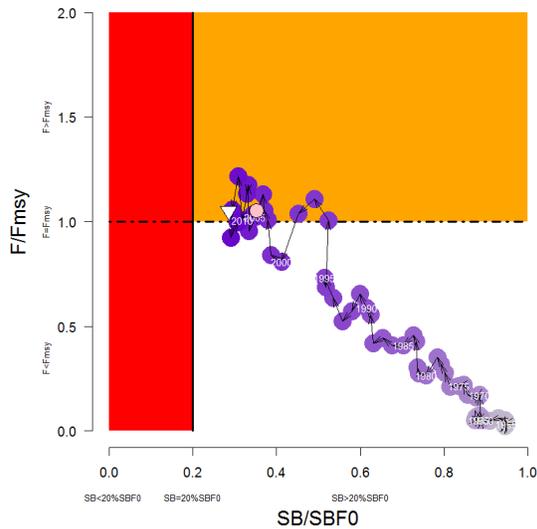
(c) Example; new growth/2014 regions (*2014Reg*)
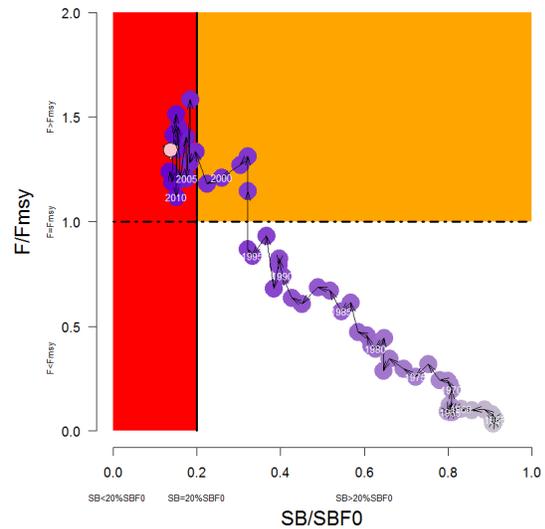
(d) Example; old growth/2014 regions (*A0B1C0D0E1*)

Figure 51: Estimated time-series (or "dynamic") Majuro plots for four example models from the assessment (one from each of the combinations of old/new growth and 2017/2014 regions). These plots are interpreted in the same manner as the description in Figure 41 except that they show the temporal change in stock status with respect to the reference points $F_{recent}/F_{\mathrm{MSY}}$ and $SB_{latest}/SB_{F=0}$, rather than the terminal estimates presented in previous figures. Note that the process of estimating a "dynamic" Majuro plot is explained further in Section 5.7.4.
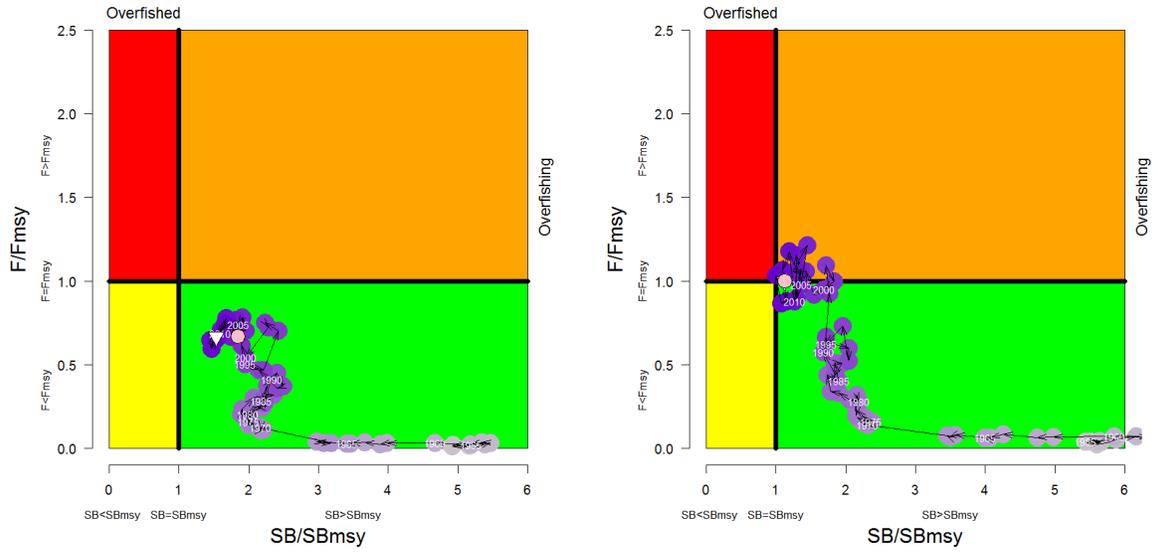
(a) Example; new growth/2017 regions (Diagnostic case)

(b) Example; old growth/2017 regions (*L2-184*)
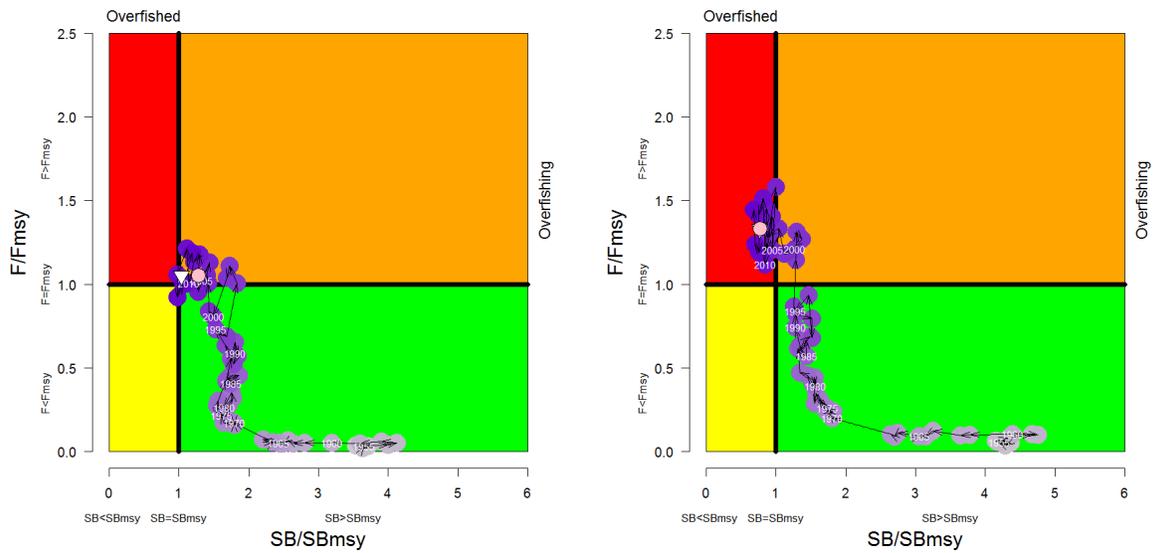
(c) Example; new growth/2014 regions (*2014Reg*)

(d) Example; old growth/2014 regions (*A0B1C0D0E1*)

Figure 52: Estimated time-series (or "dynamic") Kobe plots for four example models from the assessment (one from each of the combinations of old/new growth and 2017/2014 regions). Note that the process of estimating a "dynamic" Kobe plot is explained further in Section 5.7.4.
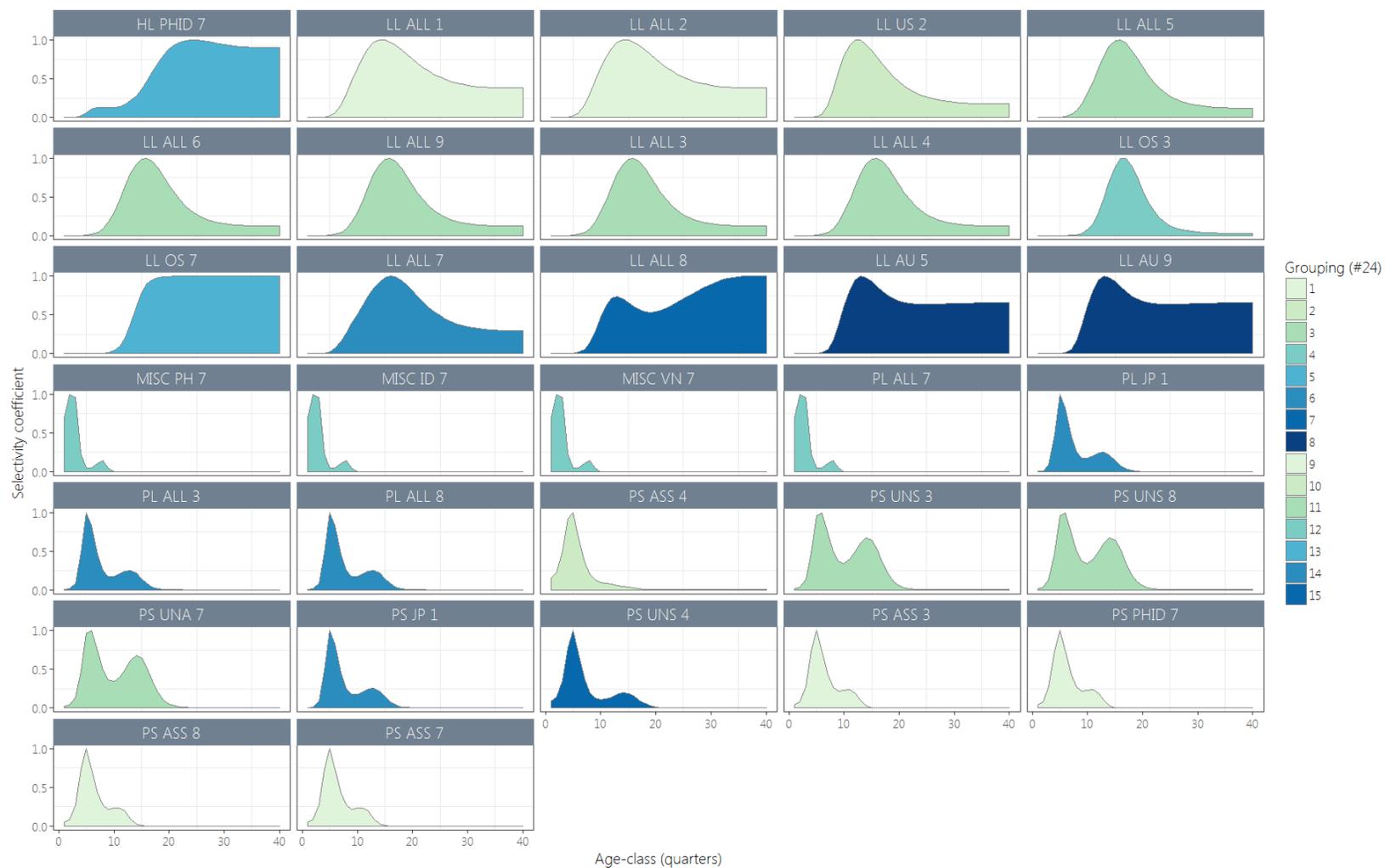
Figure 53: Age-specific selectivity coefficients by fishery for the sensitivity model *L2-184*. The colours indicate the groupings of the fisheries with respect to selectivity, such that fisheries sharing the same colour in the plot also share selectivity functions in the model.

11   Appendix

11.1   Likelihood profile

The approach for calculating a likelihood profile of the derived parameter, mean total
biomass over the assessment period (to represent scale of the stock size) is outlined
in Section 5.6. The profile was constructed by sequentially moving from the MLE in
either direction while progressively penalising the mean total biomass at increasingly
high and low values until it was determined that the minimum value had been reached
for all data components. The profile reflects the loss of fit over all the data, i.e.
the overall objective function value, and the individual data components, caused by
changing the population size from that of the maximum likelihood estimated value.
The change in likelihood relative to the maximum likelihood estimate is shown for
the total likelihood (black line) and the individual data components (coloured lines)
in Figure 54 and displays significant declines in the parameter moves further away
from the maximum value of the diagnostic case model, although the curves for the
individual components display different values of support for the mean total biomass.
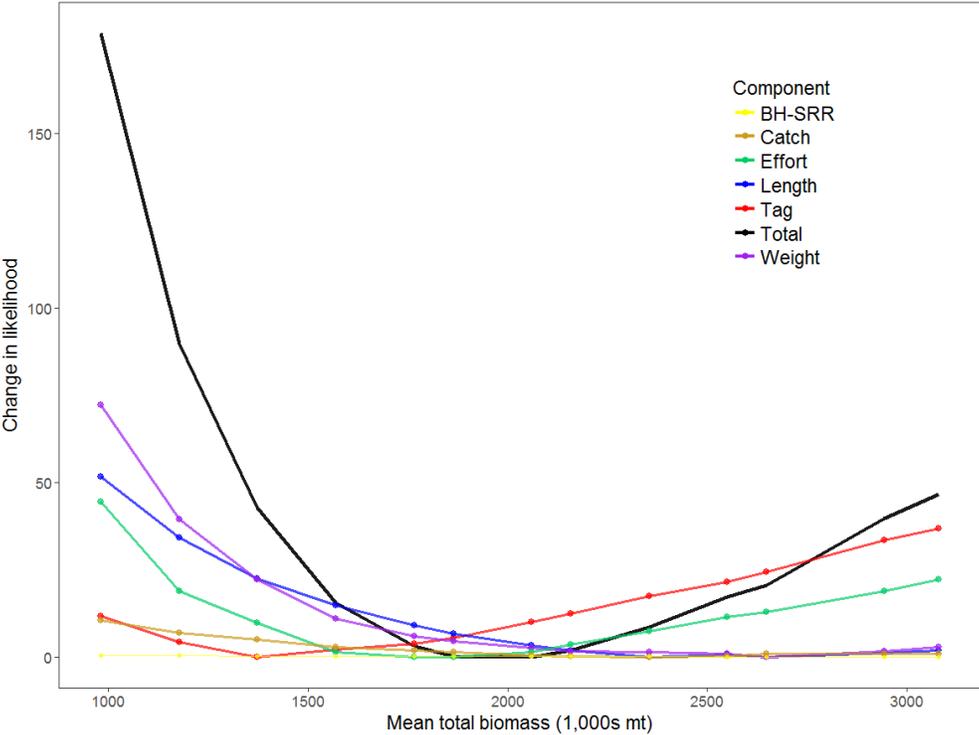
Figure 54: Change in the total, and individual data component log-likelihoods with respect to the
derived parameter, mean total biomass over the assessment period, across a range of values at which
this parameter was penalised to fit, for the diagnostic case model.

11.2 Retrospective analyses

Retrospective analyses involve rerunning the selected model by consecutively removing successive years of data to estimate model bias (Cadrin and Vaughan, 1997; Cadigan and Farrell, 2005). A series of five additional models were fitted starting with the full data-set (through 2015), followed by models with the retrospective removal of all input data for the years 2015–2011 sequentially. The models are named below by the final year of data included (e.g., 2010–2015). A comparison of the spawning potential, recruitment and depletion trajectories are shown in Figure 55.
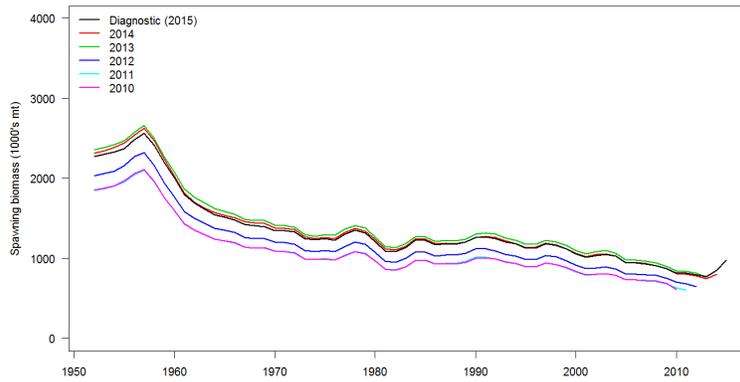
(a) Estimated spawning potential.
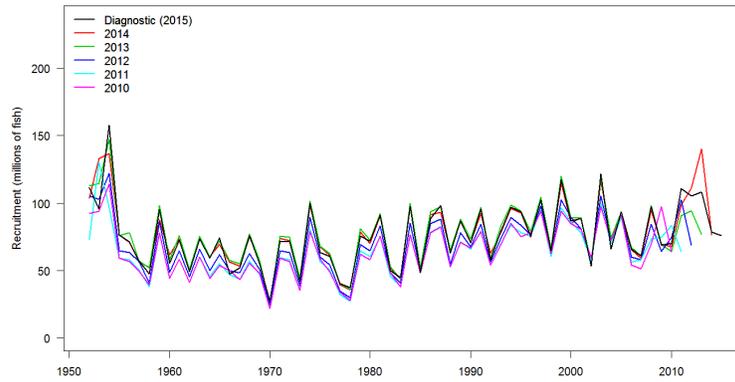
(b) Estimated recruitment.
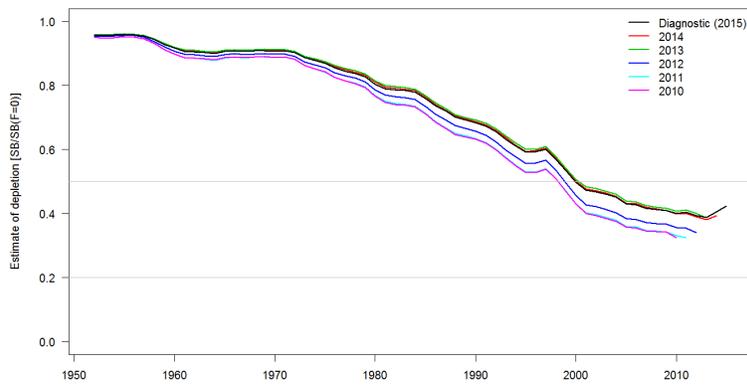
(c) Estimated fishing depletion.

Figure 55: Estimated spawning potential, recruitment and fishery depletion ($SB/SB_{F=0}$) for each of the retrospective models.

137

11.3 Sensitivity analyses reference points and likelihood values

In the current assessment, inferences about stock status and recommendations for management advice are bassed on the structural uncertainty grid, rather than the diagnostic case model and the one-off sensitivity model runs, which have received more focus in recent assessments. To this end, the estimates of reference points for the one-off sensitivity model runs are presented here in the appendix for relative comparisons against the diagnostic case model and among these models, rather than focusing on the absolute estimates that they provide. The set of focal reference points for these models are presented below in three sets, along with those of the diagnostic case model:

Table 11: Reference points for the diagnostic case model and the one-off sensitivity models

Quantity	Diagnostic	h0.65	h0.95	AL	Lorenzen	Mix-CS	L2-184
C_{latest}	156,088	155,468	155,484	153,424	156,020	155,619	156,151
MSY	166,480	179,160	162,160	146,840	161,880	168,960	150,960
$Y_{F_{recent}}$	157,000	171,800	149,920	146,800	158,240	158,520	150,960
f_{mult}	1.49	1.38	1.63	1.03	1.27	1.52	1.00
F_{MSY}	0.05	0.04	0.05	0.05	0.04	0.05	0.05
$F_{recent}/F_{\text{MSY}}$	0.67	0.73	0.61	0.97	0.78	0.66	1.00
SB_{MSY}	526,600	618,400	452,400	454,400	572,800	521,300	336,500
SB_0	1,961,000	2,166,000	1,813,000	1,801,000	2,215,000	1,957,000	1,416,000
SB_{MSY}/SB_0	0.27	0.29	0.25	0.25	0.26	0.27	0.24
$SB_{F=0}$	2,110,332	2,265,101	1,998,004	1,973,603	2,447,476	2,104,593	1,709,094
$SB_{MSY}/SB_{F=0}$	0.25	0.27	0.23	0.23	0.23	0.25	0.20
SB_{latest}/SB_0	0.50	0.44	0.53	0.45	0.39	0.49	0.27
$SB_{latest}/SB_{F=0}$	0.46	0.42	0.48	0.41	0.36	0.46	0.22
$SB_{latest}/SB_{\text{MSY}}$	1.85	1.55	2.12	1.80	1.52	1.86	1.14
$SB_{recent}/SB_{F=0}$	0.38	0.35	0.40	0.34	0.31	0.38	0.22
$SB_{recent}/SB_{\text{MSY}}$	1.54	1.28	1.75	1.48	1.34	1.54	1.12

Table 12: Reference points for the diagnostic case model and the one-off sensitivity models

Quantity	Diagnostic	Dirichlet	2014Reg	Size10	Size50	JP	Mix1	RR0.99
C_{latest}	156,088	155,636	146,426	155,993	131,678	144,194	144,037	155,424
MSY	166,480	216,120	134,920	190,400	162,920	154,760	154,440	169,840
$Y_{F_{recent}}$	157,000	176,080	134,760	171,720	148,880	149,720	148,880	158,720
f_{mult}	1.49	2.16	0.95	1.71	1.64	1.34	1.37	1.54
F_{MSY}	0.05	0.05	0.04	0.05	0.04	0.05	0.05	0.05
$F_{recent}/F_{\text{MSY}}$	0.67	0.46	1.05	0.58	0.61	0.74	0.73	0.65
SB_{MSY}	526,600	680,800	515,200	589,300	550,500	487,800	500,700	526,000
SB_0	1,961,000	2,610,000	1,669,000	2,233,000	2,087,000	1,849,000	1,893,000	1,973,000
SB_{MSY}/SB_0	0.27	0.26	0.31	0.26	0.26	0.26	0.26	0.27
$SB_{F=0}$	2,110,332	2,565,529	1,874,589	2,277,149	2,070,645	1,975,674	2,011,499	2,114,282
$SB_{MSY}/SB_{F=0}$	0.25	0.27	0.27	0.26	0.27	0.25	0.25	0.25
SB_{latest}/SB_0	0.50	0.57	0.40	0.49	0.46	0.43	0.45	0.50
$SB_{latest}/SB_{F=0}$	0.46	0.58	0.35	0.48	0.47	0.41	0.43	0.47
$SB_{latest}/SB_{\text{MSY}}$	1.85	2.18	1.29	1.84	1.76	1.64	1.71	1.89
$SB_{recent}/SB_{F=0}$	0.38	0.49	0.29	0.41	0.39	0.34	0.35	0.39
$SB_{recent}/SB_{\text{MSY}}$	1.54	1.84	1.04	1.58	1.48	1.36	1.40	1.55

Table 13: Reference points for the diagnostic case model and the one-off sensitivity models

Quantity	Diagnostic	OD2	SRR-qtr	C-Uncert	L2-184-NewMat	CPUE-Proxy	CPUE-Geostat
C_{latest}	156,088	150,437	155,516	163,605	157,536	133,020	151,890
MSY	166,480	150,800	156,360	170,600	155,840	149,760	157,240
$Y_{F_{recent}}$	157,000	149,400	151,960	165,240	155,800	144,760	153,320
f_{mult}	1.49	1.17	1.29	1.34	1.03	1.35	1.30
F_{MSY}	0.05	0.05	0.04	0.05	0.05	0.05	0.05
$F_{recent}/F_{\text{MSY}}$	0.67	0.85	0.78	0.75	0.97	0.74	0.77
SB_{MSY}	526,600	429,800	588,400	536,900	465,200	495,300	448,600
SB_0	1,961,000	1,608,000	1,968,000	1,964,000	1,888,000	1,905,000	1,697,000
$SB_{F=0}$	2,110,332	1,841,354	2,079,074	2,087,500	2,270,376	1,938,003	1,948,445
$SB_{MSY}/SB_{F=0}$	0.25	0.23	0.28	0.26	0.20	0.26	0.23
SB_{MSY}/SB_0	0.27	0.27	0.30	0.27	0.25	0.26	0.26
SB_{latest}/SB_0	0.50	0.41	0.49	0.46	0.31	0.41	0.42
$SB_{latest}/SB_{F=0}$	0.46	0.36	0.46	0.43	0.25	0.41	0.37
$SB_{latest}/SB_{\text{MSY}}$	1.85	1.53	1.64	1.68	1.24	1.59	1.59
$SB_{recent}/SB_{F=0}$	0.38	0.29	0.38	0.36	0.23	0.34	0.32
$SB_{recent}/SB_{\text{MSY}}$	1.54	1.26	1.36	1.41	1.12	1.34	1.38

Table 14: Likelihood components for the diagnostic case model and the one-off sensitivity models

Component	Diagnostic	h0.65	h0.95	AL	Lorenzen	Mix-CS	L2-184
Beverton Holt	0.6	0.6	0.6	0.8	0.7	0.6	0.5
Effort devs	2935.0	2929.5	2928.8	2889.9	3499.2	2926.9	2723.5
Catch devs	179.9	177.0	177.0	172.1	173.2	178.5	182.0
Length comps	-173615.6	-173624.6	-173624.1	-173239.9	-197570.7	-173624.9	-173468.7
Wgt comps	-921051.1	-921034.7	-921034.5	-921214.6	-1021983.9	-921039.4	-922173.9
Tagging	4762.1	4750.3	4749.5	4696.3	4904.9	4632.4	4675.8
Age-Length	0.0	0.0	0.0	1983.1	0.0	0.0	0.0
Total	-1086513.9	-1086524.7	-1086525.9	-1084385.9	-1210605.1	-1086646.8	-1087746.4

Table 15: Likelihood components for the diagnostic case model and the one-off sensitivity models

Component	Diagnostic	Dirichlet	2014Reg	Size10	Size50	JP	Mix1	RR0.99
Beverton Holt	0.6	0.6	0.6	0.6	0.4	0.5	0.5	0.6
Effort devs	2935.0	3579.3	3028.3	3189.9	2196.4	2754.7	2765.7	2912.3
Catch devs	179.9	200.7	181.5	181.9	178.4	181.1	175.4	177.5
Length comps	-173615.6	108752.7	-192637.9	-197615.4	-139438.8	-173540.1	-173534.5	-173628.7
Wgt comps	-921051.1	314651.4	-945520.2	-1022022.8	-783441.5	-920943.9	-920929.8	-921033.2
Tagging	4762.1	5162.0	4915.4	4918.6	4707.2	5565.4	5154.5	4727.3
Total	-1086513.9	221167.1	-1129718.7	-1210965.4	-915547.6	-1085689.7	-1086058.0	-1086562.9

Table 16: Likelihood components for the diagnostic case model and the one-off sensitivity models

Component	Diagnostic	OD2	SRR-qtr	C-Uncert	L2-184-NewMat	CPUE-Proxy	CPUE-Geostat
Beverton Holt	0.6	0.8	12.3	0.5	0.5	0.5	0.7
Effort devs	2935.0	2837.6	2925.0	2750.2	2699.0	2398.8	2185.7
Catch devs	179.9	173.9	177.7	183.5	188.7	196.9	181.9
Length comps	-173615.6	-173555.2	-173622.5	-173573.9	-173433.4	-173459.2	-173541.8
Wgt comps	-921051.1	-921122.1	-921033.1	-920936.6	-922034.0	-920510.9	-921184.6
Tagging	4762.1	5413.8	4751.0	4793.1	4613.8	4821.7	4828.6
Total	-1086513.9	-1085978.1	-1086513.7	-1086492.1	-1087629.0	-1086228.3	-1087265.2

11.4 Summaries of the structural uncertainty grid with the addition of an extra level of size data weighting

This section presents an update of the structural uncertainty grid for extra models run subsequent to the submission of the previous version of the bigeye stock assessment report uploaded to the WCPFC SC website. These relate to the addition of an extra level to the size data weighting axis in the structural uncertainty grid. Due to time constraints, only the two bounds of the size data weightings (divisors of 10 and 50) that were investigated in the one-off sensitivity model runs were originally included in the grid, due to the large computational overhead, limited computational resources and time constraints imposed on the assessments. Subsequent to the report deadline, we have completed the model runs for this extra level and so modify the stock assessment figures and tables to include summaries from these extra runs. These are briefly presented in this section in Figures A39–A41 and Tables A6–A10 and the two extra tables are presented that summarise this new grid over all models assuming the new growth function (Table A11) and the old growth function (Table A12), both of which comprise 36 individual model runs. These figures and tables are directly comparable to the original figures and tables via their consistent numbering. As this section presents the summaries of the most complete structural uncertainty grid, we recommend referring to the following figures and tables when formulating recommendations from this stock assessment. The original figures, tables and main body text are left unmodified only to prevent confusion for those that have previously downloaded the stock assessment report.

The addition of the extra level of the size data weighting uncertainty axis had minimal effects on the estimates of stock status and model outputs of the assessment. Because the divisor of 20 falls between the more extreme levels (10 and 50) originally included, the new models almost exclusively produce estimates less extreme than those in the original figures and tables. Therefore, the bounds of the previously estimated grid and subsets of it (new growth/new regions etc.) tend to remain constant compared to Section 7.5.2. Minor changes occur with respect to estimates of central tendencies and the quantiles owing to the sometimes small sample sizes of models in a certain subset (previously 12, now 18, for the growth/regional structure subsets), but in no case are the conclusions about stock status modified from those originally presented.
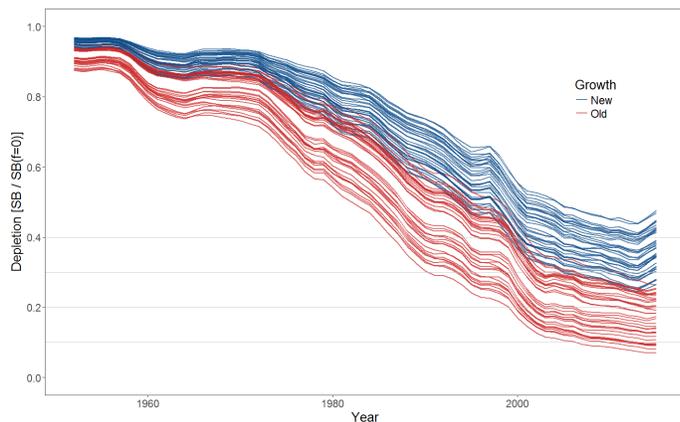
(a) Old and new growth
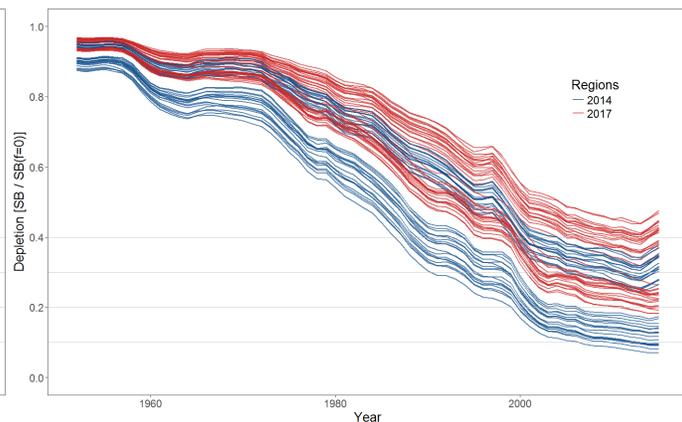
(b) 2014 and 2017 regions
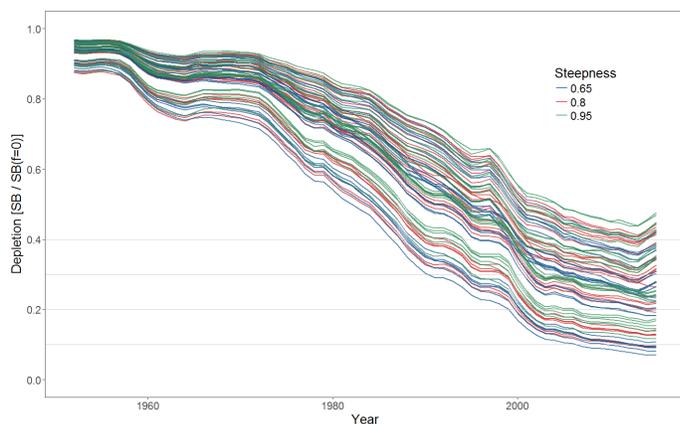
(c) Steepness
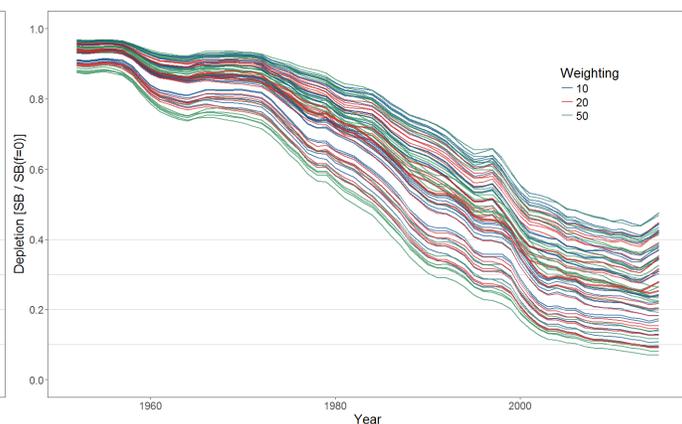
(d) Weighting of size data
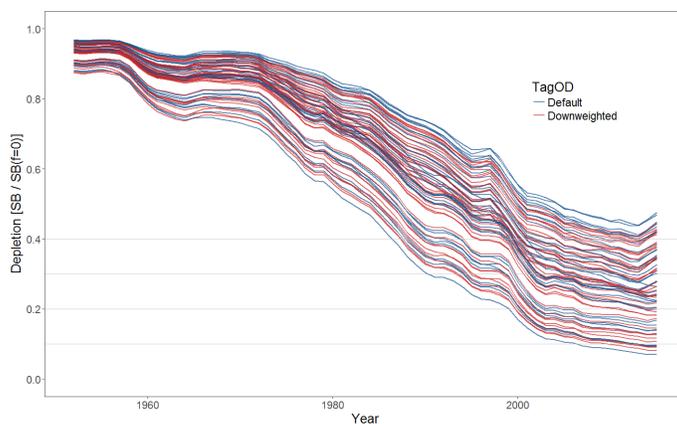
(e) Weighting of tagging data

Figure A39: Plots showing the trajectories of fishing depletion (of spawning potential) for the model runs included in the structural uncertainty grid (see Section 6.2 for details of the structure of the grid models). The five panels show the models separated on the basis of the five axes used in the grid, with the colour denoting the level within the axes for each model.
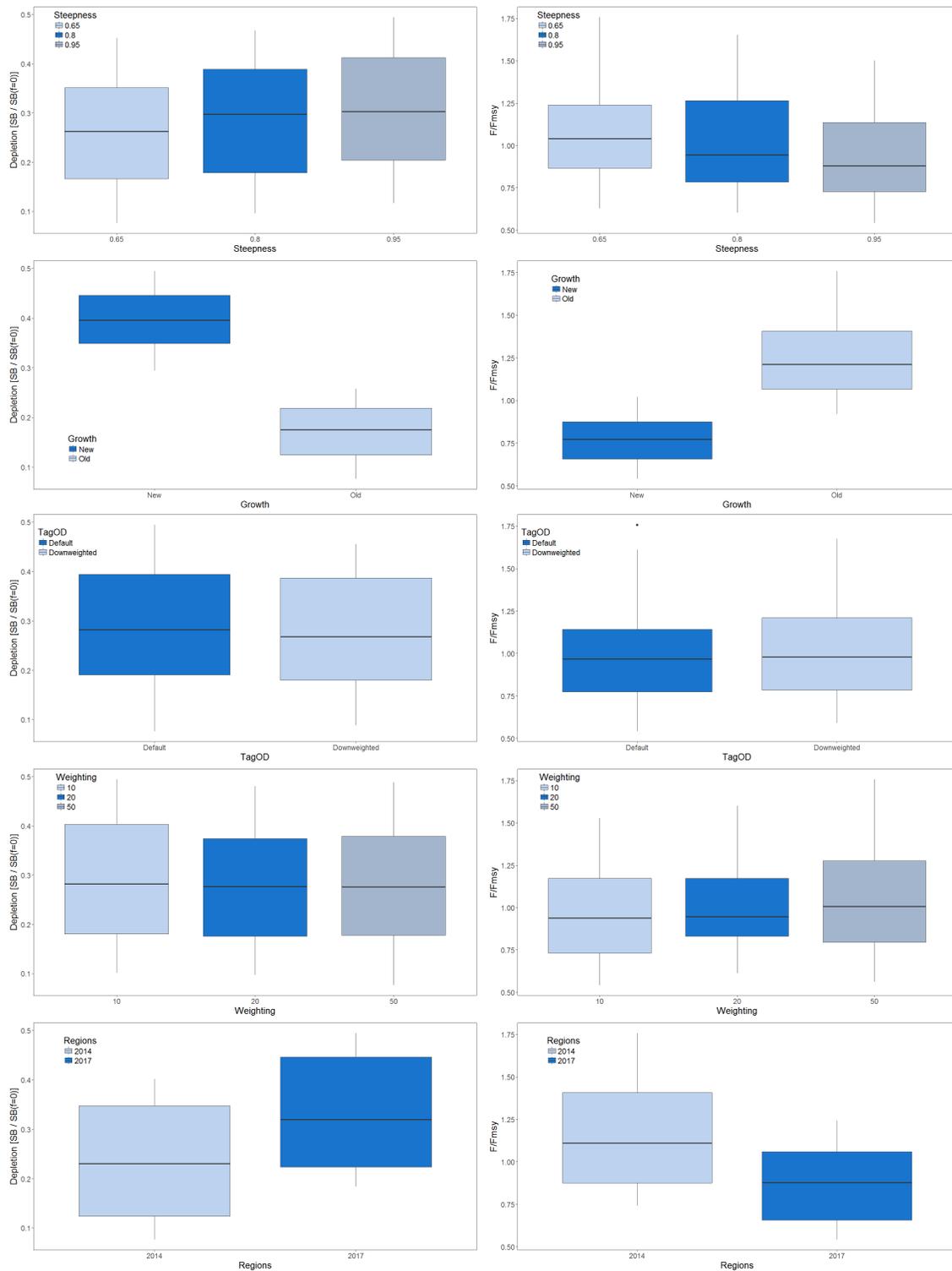
Figure A40: Boxplots summarising the results of the structural uncertainty grid with respect to the spawning potential reference point (left panels), and the fishing mortality reference point F_{recent}/F_{MSY} (right panels). The colours indicate the level of the model with respect to each uncertainty axis.
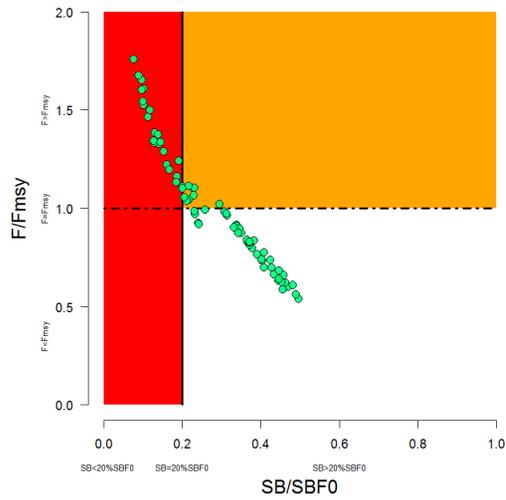
(a) All grid models SB_{latest}
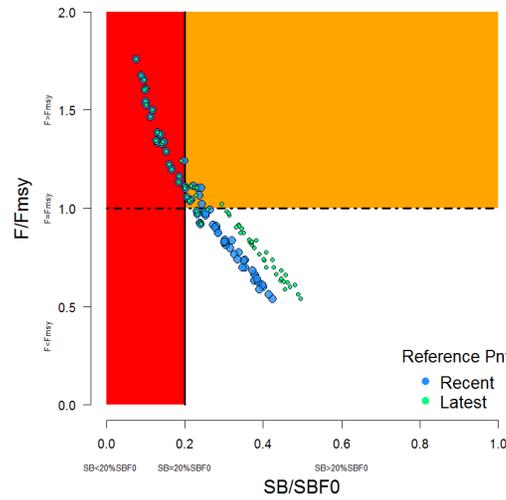
(b) All grid models SB_{latest} and SB_{recent}
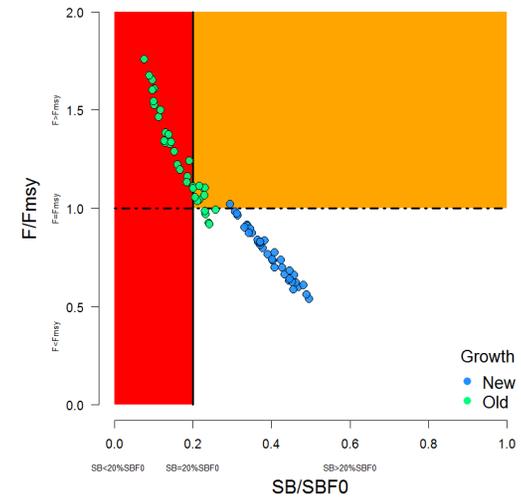
(c) Old and new growth
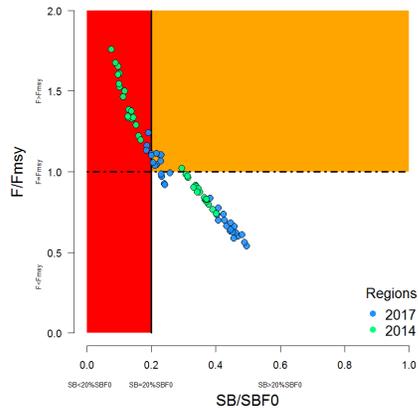
(d) 2014 and 2017 regions
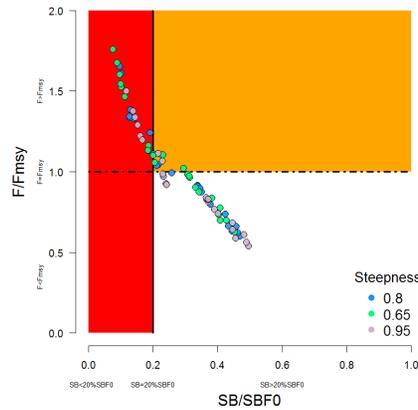
(e) Steepness
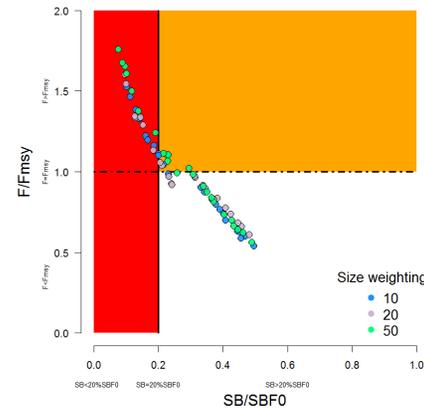
(f) Weighting of size data
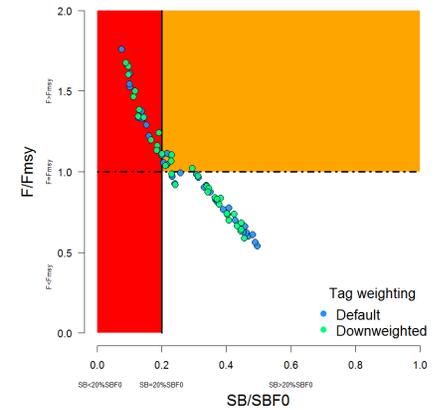
(g) Tagging data weighting

Figure A41: Majuro plots summarising the results for each of the models in the structural uncertainty grid. The plots represent estimates of 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 points represent $SB_{latest}/SB_{F=0}$ for each model run except in panel (b) where $SB_{recent}/SB_{F=0}$ is also displayed. Panels (c)–(g) show the estimates for the different levels for the five axes of the grid.

Table A5: Description of the structural sensitivity grid used to characterise uncertainty in the assessment. Note that this is for the full updated grid consisting of 72 models.

Axis	Levels	Option
Steepness	3	0.65, 0.80, or 0.95
Growth	2	Low growth L2 = 152, High growth L2 = 184
Tagging overdispersion	2	Default level, Fixed (moderate) level
Size frequency weighting	3	sample sizes divided by 10, 20, or 50
Regional structure	2	2017 regions, 2014 regions

Table A6: Summary of reference points over all 72 models in the structural uncertainty grid

	Mean	Median	Min	25%	75%	Max
C_{latest}	150,843	154,694	130,903	147,434	155,763	157,725
MSY	153,152	151,620	124,120	140,700	164,340	204,040
$Y_{F_{recent}}$	146,572	146,480	118,000	137,130	156,710	187,240
f_{mult}	1.08	1.03	0.57	0.83	1.29	1.85
F_{MSY}	0.05	0.05	0.04	0.04	0.05	0.06
F_{recent}/F_{MSY}	1.01	0.97	0.54	0.77	1.21	1.76
SB_{MSY}	413,774	419,500	219,500	312,825	490,850	710,000
SB_0	1,604,986	1,574,000	1,009,000	1,323,750	1,818,500	2,509,000
SB_{MSY}/SB_0	0.25	0.26	0.22	0.24	0.27	0.29
$SB_{F=0}$	1,842,745	1,904,313	1,317,336	1,624,408	2,012,394	2,460,411
$SB_{MSY}/SB_{F=0}$	0.22	0.22	0.17	0.19	0.25	0.29
SB_{latest}/SB_0	0.32	0.31	0.11	0.22	0.42	0.53
$SB_{latest}/SB_{F=0}$	0.28	0.28	0.08	0.18	0.39	0.49
SB_{latest}/SB_{MSY}	1.24	1.21	0.42	0.91	1.58	2.12
$SB_{recent}/SB_{F=0}$	0.25	0.25	0.08	0.18	0.33	0.42
SB_{recent}/SB_{MSY}	1.05	1.04	0.32	0.74	1.33	1.86

Table A7: Summary of reference points over the 18 models in the structural uncertainty grid within the subset of new growth/ 2017 regions

	Mean	Median	Min	25%	75%	Max
C_{latest}	146,399	154,694	130,903	131,593	155,418	156,173
MSY	169,347	165,320	144,600	159,260	180,860	204,040
$Y_{F_{recent}}$	158,333	158,720	135,200	148,660	164,240	187,240
f_{mult}	1.53	1.53	1.20	1.43	1.63	1.85
F_{MSY}	0.05	0.05	0.04	0.04	0.05	0.05
F_{recent}/F_{MSY}	0.66	0.65	0.54	0.61	0.70	0.84
SB_{MSY}	539,494	526,100	422,400	484,300	599,525	710,000
SB_0	2,029,444	2,042,000	1,682,000	1,932,000	2,148,500	2,509,000
SB_{MSY}/SB_0	0.26	0.26	0.24	0.25	0.28	0.28
$SB_{F=0}$	2,102,800	2,107,094	1,812,463	1,999,665	2,172,111	2,460,411
$SB_{MSY}/SB_{F=0}$	0.26	0.25	0.22	0.24	0.27	0.29
SB_{latest}/SB_0	0.46	0.47	0.40	0.42	0.49	0.53
$SB_{latest}/SB_{F=0}$	0.44	0.45	0.38	0.42	0.46	0.49
SB_{latest}/SB_{MSY}	1.75	1.77	1.40	1.50	1.98	2.12
$SB_{recent}/SB_{F=0}$	0.38	0.38	0.32	0.35	0.39	0.42
SB_{recent}/SB_{MSY}	1.56	1.57	1.22	1.45	1.66	1.86

Table A8: Summary of reference points over the 18 models in the structural uncertainty grid within the subset of new growth/ 2014 regions

	Mean	Median	Min	25%	75%	Max
C_{latest}	148,628	151,211	141,388	141,730	153,131	153,153
MSY	151,411	151,380	134,800	141,930	158,400	170,760
$Y_{F_{recent}}$	150,051	147,820	133,400	141,230	156,040	170,000
f_{mult}	1.15	1.14	0.98	1.09	1.21	1.35
F_{MSY}	0.05	0.05	0.04	0.05	0.05	0.05
F_{recent}/F_{MSY}	0.88	0.88	0.74	0.83	0.92	1.02
SB_{MSY}	461,606	441,700	372,400	418,350	519,800	579,800
SB_0	1,682,222	1,658,000	1,449,000	1,563,250	1,787,750	1,999,000
SB_{MSY}/SB_0	0.27	0.27	0.25	0.26	0.29	0.29
$SB_{F=0}$	1,872,444	1,825,564	1,637,143	1,767,568	1,977,536	2,144,857
$SB_{MSY}/SB_{F=0}$	0.25	0.25	0.22	0.23	0.26	0.27
SB_{latest}/SB_0	0.39	0.39	0.32	0.36	0.41	0.45
$SB_{latest}/SB_{F=0}$	0.35	0.35	0.29	0.33	0.37	0.40
SB_{latest}/SB_{MSY}	1.43	1.44	1.12	1.24	1.61	1.75
$SB_{recent}/SB_{F=0}$	0.29	0.28	0.24	0.27	0.30	0.33
SB_{recent}/SB_{MSY}	1.17	1.16	0.97	1.11	1.24	1.39

Table A9: Summary of reference points over the 18 models in the structural uncertainty grid within the subset of old growth/ 2017 regions

	Mean	Median	Min	25%	75%	Max
C_{latest}	152,613	155,681	144,501	147,799	156,317	156,440
MSY	146,518	145,840	124,120	138,840	151,300	168,040
$Y_{F_{recent}}$	145,824	145,620	123,880	138,830	151,030	167,760
f_{mult}	0.95	0.94	0.80	0.90	1.01	1.09
F_{MSY}	0.04	0.05	0.04	0.04	0.05	0.05
F_{recent}/F_{MSY}	1.06	1.06	0.92	0.99	1.11	1.24
SB_{MSY}	354,689	333,700	278,100	285,400	422,200	470,400
SB_0	1,459,167	1,395,000	1,256,000	1,296,000	1,623,750	1,767,000
SB_{MSY}/SB_0	0.24	0.24	0.22	0.23	0.26	0.27
$SB_{F=0}$	1,722,699	1,698,361	1,474,692	1,570,332	1,924,588	2,001,488
$SB_{MSY}/SB_{F=0}$	0.20	0.20	0.18	0.19	0.22	0.24
SB_{latest}/SB_0	0.26	0.25	0.22	0.24	0.28	0.29
$SB_{latest}/SB_{F=0}$	0.22	0.22	0.18	0.20	0.23	0.26
SB_{latest}/SB_{MSY}	1.08	1.09	0.86	0.93	1.19	1.35
$SB_{recent}/SB_{F=0}$	0.22	0.23	0.18	0.20	0.24	0.26
SB_{recent}/SB_{MSY}	0.93	0.91	0.70	0.85	1.02	1.13

Table A10: Summary of reference points over the 18 models in the structural uncertainty grid within the subset of old growth/ 2014 regions

	Mean	Median	Min	25%	75%	Max
C_{latest}	155,731	155,476	154,089	155,204	156,361	157,725
MSY	145,331	141,580	125,120	131,350	163,890	170,080
$Y_{F_{recent}}$	132,080	132,860	118,000	122,280	138,600	149,400
f_{mult}	0.70	0.70	0.57	0.63	0.75	0.83
F_{MSY}	0.05	0.05	0.04	0.05	0.05	0.06
F_{recent}/F_{MSY}	1.45	1.43	1.20	1.34	1.59	1.76
SB_{MSY}	299,306	278,200	219,500	244,875	375,550	393,200
SB_0	1,249,111	1,200,000	1,009,000	1,107,500	1,461,000	1,519,000
SB_{MSY}/SB_0	0.24	0.24	0.22	0.22	0.26	0.26
$SB_{F=0}$	1,673,038	1,597,112	1,317,336	1,437,125	2,001,950	2,137,922
$SB_{MSY}/SB_{F=0}$	0.18	0.17	0.17	0.17	0.19	0.19
SB_{latest}/SB_0	0.16	0.16	0.11	0.14	0.18	0.22
$SB_{latest}/SB_{F=0}$	0.12	0.12	0.08	0.10	0.14	0.17
SB_{latest}/SB_{MSY}	0.69	0.71	0.42	0.54	0.81	0.98
$SB_{recent}/SB_{F=0}$	0.12	0.12	0.08	0.10	0.14	0.17
SB_{recent}/SB_{MSY}	0.52	0.54	0.32	0.42	0.60	0.74

Table A11: Summary of reference points over the 36 models in the structural uncertainty grid within the subset of all new growth models (both 2017 and 2014 regions)

	Mean	Median	Min	25%	75%	Max
C_{latest}	147,513	151,242	130,903	141,679	154,680	156,173
MSY	160,379	158,520	134,800	150,230	168,310	204,040
$Y_{F_{recent}}$	154,192	153,920	133,400	144,590	162,380	187,240
f_{mult}	1.34	1.30	0.98	1.14	1.52	1.85
F_{MSY}	0.05	0.05	0.04	0.04	0.05	0.05
F_{recent}/F_{MSY}	0.77	0.77	0.54	0.66	0.88	1.02
SB_{MSY}	500,550	497,100	372,400	428,975	549,825	710,000
SB_0	1,855,833	1,819,000	1,449,000	1,660,000	2,039,500	2,509,000
SB_{MSY}/SB_0	0.27	0.27	0.24	0.26	0.28	0.29
$SB_{F=0}$	1,987,622	1,996,851	1,637,143	1,820,761	2,131,877	2,460,411
$SB_{MSY}/SB_{F=0}$	0.25	0.25	0.22	0.24	0.27	0.29
SB_{latest}/SB_0	0.43	0.42	0.32	0.40	0.47	0.53
$SB_{latest}/SB_{F=0}$	0.40	0.40	0.29	0.35	0.45	0.49
SB_{latest}/SB_{MSY}	1.59	1.58	1.12	1.41	1.76	2.12
$SB_{recent}/SB_{F=0}$	0.33	0.33	0.24	0.29	0.38	0.42
SB_{recent}/SB_{MSY}	1.37	1.33	0.97	1.17	1.56	1.86

Table A12: Summary of reference points over the 36 models in the structural uncertainty grid within the subset of all old growth models (both 2017 and 2014 regions)

	Mean	Median	Min	25%	75%	Max
C_{latest}	154,172	155,568	144,501	154,782	156,335	157,725
MSY	145,924	144,020	124,120	136,600	161,400	170,080
$Y_{F_{recent}}$	138,952	138,780	118,000	129,430	147,530	167,760
f_{mult}	0.82	0.83	0.57	0.71	0.94	1.09
F_{MSY}	0.05	0.05	0.04	0.04	0.05	0.06
F_{recent}/F_{MSY}	1.26	1.21	0.92	1.07	1.41	1.76
SB_{MSY}	326,997	307,450	219,500	277,375	384,800	470,400
SB_0	1,354,139	1,315,500	1,009,000	1,207,000	1,506,000	1,767,000
SB_{MSY}/SB_0	0.24	0.24	0.22	0.22	0.26	0.27
$SB_{F=0}$	1,697,869	1,622,683	1,317,336	1,560,934	1,940,752	2,137,922
$SB_{MSY}/SB_{F=0}$	0.19	0.19	0.17	0.17	0.20	0.24
SB_{latest}/SB_0	0.21	0.22	0.11	0.16	0.25	0.29
$SB_{latest}/SB_{F=0}$	0.17	0.18	0.08	0.12	0.22	0.26
SB_{latest}/SB_{MSY}	0.88	0.90	0.42	0.72	1.09	1.35
$SB_{recent}/SB_{F=0}$	0.17	0.18	0.08	0.12	0.22	0.26
SB_{recent}/SB_{MSY}	0.73	0.73	0.32	0.55	0.91	1.13

11.5 Plot displaying the recruitment estimates for model *L2-184*

Subsequent to the submission of the original version of this document, a CCM requested that the recruitment estimates for the old growth model (sensitivity model *L2-184*) be presented to confirm the recent high recruitment event. This is shown in Figure A42 and we also present the total biomass and spawning potential for this model to show that these recruits led to increased total biomass in the following years but had not yet entered the spawning potential "population" by the terminal year of the assessment (2015). This is expected to occur in the years subsequent to 2015 and it is expected that an increase in spawning potential and slightly more optimistic depletion estimates will occur in a similar manner to the new growth models.
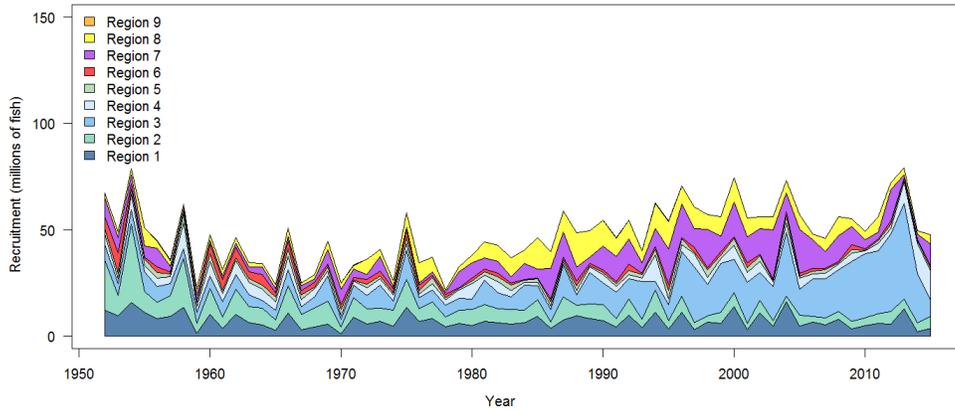
(a) Regional recruitment
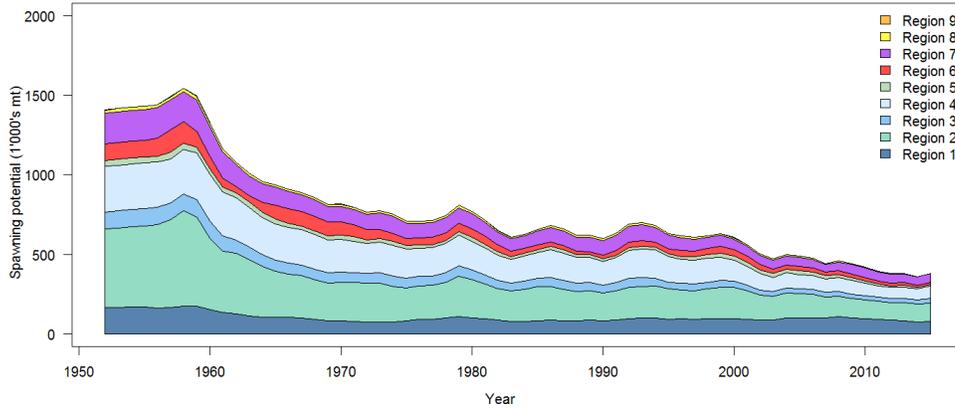
(b) Spawning potential
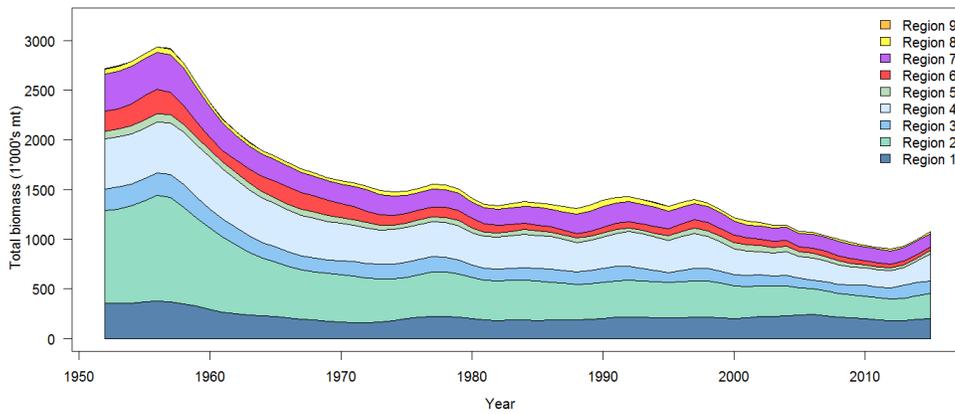
(c) Total biomass

Figure A42: Estimated annual recruitment, and average annual spawning potential and total biomass by model region for the sensitivity model *L2-184*, showing the relative sizes among regions.