SCIENTIFIC COMMITTEE
FIFTEENTH REGULAR SESSION
Pohnpei, Federated States of Micronesia
12–20 August 2019

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

WCPFC-SC15-2019/SA-WP-05-Rev2

M. T. Vincent[1], **G. M. Pilling** [1], **J Hampton** [1]

[1]Oceanic Fisheries Programme, The Pacific Community

Contents

Revision 1: August 1^{st} 2019

The previous version had 1.) selected the wrong output files for a few models in the 8 region model structural uncertainty grid. The differences in these models can only been seen in the calculated mean of reference points for the 8 region model, but the median and 80^{th} percentile ranges on which stock status advice is based are unchanged (Table 6); 2.) presented the "Recent average $F_{recent}/F_{\text{MSY}}$" but the median is reported, thus we changed the wording to "Recent median".

Revision 2: August 7^{th} 2019

The previous version had the projection plots for the 8 Region and 5 Region model switched (Section A.2 and Figure A9).

1 Executive Summary

This paper describes the 2019 stock assessment of skipjack tuna *Katsuwonus pelamis* in the western and central Pacific Ocean. An additional three years of data were available since the previous assessment in 2016, and the model extends through to the end of 2018. New developments to the stock assessment including addressing the recommendations of the 2016 stock assessment report (McKechnie et al., 2016), revision and incorporation of new data sources such as maturity-at-length, creation of an additional spatial structure, exploration of model uncertainty, and improving the diagnostics of previous assessments.

This assessment is supported by the analysis of catch-per-unit-effort data for pole-and-line (Kiyofuji et al., 2019b; Vincent et al., 2019) and purse seine fisheries (Bigelow et al., 2019; Vidal et al., 2019), recently collected biological data (Ohashi et al., 2019), and summaries of the input data used within the stock assessment (Vincent et al., 2019).

Changes that were influential in the progression from the 2016 to 2019 diagnostic models include:

1. update of data through the end of 2018 including updating the length-weight relationship
2. incorporation of Japanese tag releases that did not have release length from 1989 onward
3. methods to obtain the growth estimates used in the structural uncertainty grid
4. estimation of tag overdispersion parameters
5. incorporation of newly available maturity-at-length through application of recent developments in MULTIFAN-CL

6. reduction in the influence of length composition data (i.e., increasing the length composition scalar)

7. creation of an alternative 8-region model structure

In addition to the diagnostic model, we report the results of one-off sensitivity models to explore the impact of key data and model assumptions for the diagnostic 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 those areas of uncertainty from the one-off models were included.

It is recommended that management advice is formulated from the results of the structural uncertainty grid. We have provided separate grids for the two regional structures considered within the current assessment (8 region and 5 region) and recommend SC15 consider the provision of advice on a preferred spatial structure. We also invite SC15 to consider the relative weighting of models included in the structural uncertainty grid. The results below are based on equal weighting of all models.

Across the 54 models of the structural uncertainty grid run in this assessment within each regional structure, the most important factors when evaluating stock status were the assumed tag mixing period, and assumed steepness of the stock recruitment relationship. For tag mixing, assumption of a mixing period of 1 resulted in higher estimates of $F_{recent}/F_{\text{MSY}}$ and slightly more pessimistic stock status when compared to the assumption of a mixing period of 2. As expected, $F_{recent}/F_{\text{MSY}}$ increased as lower levels of steepness were assumed, and $SB_{recent}/SB_{F=0}$ became slightly more pessimistic.

The general conclusions of this assessment are as follows:

- Total biomass and spawning potential remained relatively stable, with fluctuations, until the mid 2000s, after which it declined. Estimated recruitment shows an increasing trend from 1980 to the recent period.

- Average fishing mortality rates for juvenile and adult age-classes increase throughout the period of the assessment.

- The 8 region model structure provided slightly more optimistic estimates of stock status when compared to the 5 region model structure. In both cases, the stock was assessed to be above the adopted LRP, and fished at rates below F_{MSY}, with 100% probability. We conclude the skipjack stock is not overfished, nor subject to overfishing.

- Overall median depletion over the recent period (2015-2018; $SB_{recent}/SB_{F=0}$) was 0.44 (80 percentile range 0.36-0.52) for the 8 region model, and 0.40 (80 percentile range 0.30-0.50) for the 5 region model.

- Results from both regional structures indicate a stock status currently on average below the

interim TRP for skipjack.

  - For the 8 region grid, 47 of the 54 models (85%) estimated $SB_{recent}/SB_{F=0}$ to be less than the TRP ($50\%SB_{F=0}$).
  - For the 5 region grid, 48 of the 54 models (87%) estimated $SB_{recent}/SB_{F=0}$ to be less than the TRP ($50\%SB_{F=0}$).

- Recent median fishing mortality (2014-2017; $F_{recent}/F_{\mathrm{MSY}}$) was 0.44 (80 percentile range 0.34-0.61) for the 8 region model, and 0.48 (80 percentile range 0.35-0.66) for the 5 region model.
- A number of key research needs have been identified in undertaking this assessment that should be investigated either internally or through directed research. These include: improving growth estimates through incorporation of tagging-based length-increment data; evaluation of alternative sources of CPUE time series; further evaluation of the tagging data and associated model settings; and consideration of time-varying movement functionality in MULTIFAN-CL.

2 Introduction

This paper presents the 2019 stock assessment of skipjack tuna (*Katsuwonus pelamis*; SKJ) in the western and central Pacific Ocean (WCPO; west of 150°W). Since 2000, an assessment for skipjack has been conducted regularly and the most recent assessments are documented in Hoyle et al. (2011), Rice et al. (2014), and McKechnie et al. (2016). Previous assessments incrementally improved the assessment model in light of the independent review of the 2011 bigeye assessment (Ianelli et al., 2012). Likewise, this assessment incrementally improved the assessment by review of data sources and incorporation of features resulting from ongoing development of the statistical stock assessment software, known as MULTIFAN-CL[2] (also known as MFCL; Fournier et al., 1998; Hampton and Fournier, 2001; Davies et al., 2018), that is used by the Pacific Community (SPC). Further developments to the stock assessment were undertaken to address the recommendations of the 2016 stock assessment report (McKechnie et al., 2016), to revise old input data and include additional years of data, to incorporate new data sources such as maturity-at-length, to explore model uncertainty, and to improve the diagnostics of previous assessments.

The objective of this assessment is to estimate population parameters for skipjack in the WCPO, such as time series of recruitment, biomass 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 modeling (Fournier and Archibald, 1982), which is carried out using MULTIFAN-CL, 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.

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 Bigelow et al. (2019); Vidal et al. (2019); Kiyofuji et al. (2019b) and Vincent et al. (2019). Finally, many of these issues were discussed in detail at the Pre-Assessment Workshop (PAW) held in Noumea over 2–5 April, 2019 (Pilling and Brouwer, 2019).

3 Background

3.1 Stock structure

Surface-schooling, adult skipjack tuna (greater than 40 cm fork length; FL) are highly abundant in tropical and subtropical waters of the Pacific Ocean (Figure 1). Skipjack in the WCPO are considered a single stock for assessment purposes (Wild and Hampton, 1994). In the western Pacific, warm, pole-ward-flowing currents near northern Japan and southern Australia seasonally extend

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

their distribution to about 40°N and 40°S. These limits roughly correspond to the 20°C surface isotherm.

A substantial amount of information on skipjack movement is available from tagging programs, which have documented some large-scale movement within the Pacific (Figure 2). In general, skipjack movement is highly variable (Sibert et al., 1999) and is thought to be influenced by large-scale oceanographic variability (Lehodey et al., 1997).

3.2 Biological characteristics

Skipjack growth is rapid compared to yellowfin and bigeye tuna. Approximate age estimates from counting daily rings on otoliths suggest that growth may vary between areas of the Pacific. At 150, 200, 300 and 400 days, fork lengths (FLs) of 30, 33, 40, and 46 cm were estimated for fish sampled mostly in the north Pacific (Tanabe et al., 2003), but growth estimates were faster (42, 47, 55, and 60 cm) for fish sampled close to the equator (Leroy, 2000). Growth has been found to vary spatially in the eastern Pacific (Maunder, 2001) and in the Atlantic (Gaertner et al., 2008), based on analyses of tagging data.

Estimates of natural mortality rate have been obtained using a size-structured tag attrition model (Hampton, 2000), which indicated that natural mortality was substantially larger for small skipjack (21–30 cm FL, M=0.8 mo^{-1}) compared to larger skipjack (51–70 cm FL, M=0.12–0.15 mo^{-1}). The longest period at liberty for a tagged skipjack to date has been 4.5 years.

Commensurate with their fast growth and relatively short lifespan, skipjack mature at an early age, when they reach a length of around 50 cm FL (Ohashi et al., 2019).

3.3 Fisheries

Skipjack tuna comprise the largest component of the tuna fisheries throughout the WCPO and are caught using a wide variety of fishing gears. Fisheries can be broadly classified in the Japanese pole-and-line fleets (both distant-water and offshore); domestic pole-and line fleets based in Pacific Island countries; artisanal fleets fishing a wide range of gears based in the Philippines (PH), Indonesia (ID), Vietnam (VN), and the Pacific Islands; and distant-water (DW) and Pacific-Island based purse seine fleets.

The Japanese pole-and-line fleets operate over a large area of the WCPO, although effort and the spatial extent of this fishery has gradually declined since the 1980's. A domestic pole-and-line fishery occurred in Papua New Guinea (PNG) from 1970 to 1985 and active fisheries have occurred in Fiji since 1974 (now discontinued), and the Solomon Islands since 1971 (now operating at a very low level).

A variety of gear types (e.g. gillnet, hook and line, longline, purse seine, ring net, pole-and-line and unclassified) capture a significant amount of skipjack in the Philippines, Indonesia, and Vietnam. Small, but locally important artisanal fisheries for skipjack and other tuna (mainly using traditional methods and trolling) also occur in many of the Pacific Islands.

Purse seine fleets usually operate in equatorial waters from 10°N to 10°S; although a Japanese offshore purse seine fleet operates in the temperate North Pacific. The distant-water fleets from Japan, Korea, Chinese Taipei, and the USA capture most of the skipjack in the WCPO, although catches by fleets flagged to or chartered by Pacific Island countries have increased considerably in recent years. The purse seine fishery is usually classified by set type categories - sets on floating objects such as logs and fish aggregation devices (FADs), which are termed "associated sets" and sets on free-swimming schools, termed "unassociated sets". These different set types have somewhat different spatial distributions, catch per unit effort (CPUE), and catch different sizes of skipjack and other tuna.

Skipjack tuna catches in the WCPO increased steadily after 1970, approximately doubling during the 1980s (Figure 3). The catch was then relatively stable during the early 1990s, approaching 1 million mt per annum. Catches increased again from the late 1990s and have varied between about 1.5 and 2 million mt since 2007, with a record catch of just over 2 million mt taken in 2014.

Pole-and-line fleets, primarily Japanese, initially dominated the fishery, with the catch peaking at 380,000 mt in 1984, but the relative importance of this fishery has declined steadily for economic reasons. Annual skipjack tuna catches increased during the 1980s due to growth in the international purse seine fleet, combined with increased catches by domestic fleets from the Indonesia, Philippines, and Vietnam.

Historically, most of the catch has been taken from the equatorial Pacific (Regions 5, 6, 7, and 8; Figure 4). During the 1990s, combined annual catches from this region fluctuated around 500,000–800,000 mt before increasing sharply to approximately 1.2 million mt in 2007–2009 (Figure 4). Since the late 1990s, there has been a large increase in the purse-seine fishery in the eastern equatorial region of the WCPO (Region 8), although catches from this region have been highly variable among years. Since 2014, catches in the eastern equatorial region have been particularly high, under the influence of strong El Niño conditions, thought to drive eastward displacement of skipjack tuna (Lehodey et al., 1997), that occurred in 2014–2016.

4 Data compilation

4.1 General notes

Data used in the stock assessment of skipjack tuna using MULTIFAN-CL consist of catch, effort and length-frequency data for the fisheries defined in the analysis, and tag-recapture data Figure 5.

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 Vincent et al. (2019) (length-weight relationship, tagging data, length compositions for purse seine fisheries), Vincent (2019) (fisheries definitions), Bigelow et al. (2019), Kinoshita et al. (2019) and Vidal et al. (2019) (CPUE standardizations). 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 assessment corresponds to the WCPO (from 50°N to 20°S between 120°E and 150°W) and oceanic waters adjacent to the east Asian coast (110°E between 20°N and 20°S). Two spatial structures were considered in this stock assessment: one with 8 regions (Figure 1a) and another with 5 regions (Figure 1b). The 8 region model was created in an attempt to capture the seasonal movement dynamics and differences in size composition observed in the Japanese pole-and-line fishery in this region (Kiyofuji et al., 2019a). Additional consideration of the region boundaries was given in an attempt to ensure there was sufficient tag releases within each region to estimate the size of biomass within that region. Equatorial Regions 7 and 8 are similar to the 2016 skipjack tuna assessment Regions 2 and 3 but differed by the exclusion of 10°S to 20°S from the equatorial region and inclusion in a newly created region. Region 5 is similar to Region 4 of the 5 region model but excludes a 10°x10°area from the northeast corner.

Throughout this report we will primarily focus on descriptions of the 8 region model, where descriptions of the fisheries and regions will be to this spatial stratification unless otherwise described. We present the 8 region model as the diagnostic model and present the evaluation of the fit to the data for this spatial stratification. However, we conducted assessments for each of the spatial stratifications and present the results of these assessments separately in the structural uncertainty analysis.

4.3 Temporal stratification

The time period covered by the assessment is 1972–2018. Within this period, data were compiled into quarters (1; Jan–Mar, 2; Apr–Jun, 3; Jul–Sep, 4; Oct–Dec). This assessment included data from the most recent full calendar year (2018), which are finalized very late and are considered to be preliminary.

4.4 Definition of fisheries

MULTIFAN-CL requires the definition of "fisheries" 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 some allowance can be made for time series variation. For most pelagic fisheries assessments, fisheries are defined according to gear type, fishing method, region, and sometimes by vessel flag or fleet.

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. A purse seine fishery for the Japanese purse seine fleet for all set types was assumed in Regions 1–3. Purse seine catch in Region 4 by all nationalities was insufficient (less than 1,000 metric tons) to warrant creating a fishery. Additional fisheries were defined for pole-and-line fisheries in each region and miscellaneous fisheries (gillnets, ringnets, handlines etc.) in the western equatorial area. A longline fishery was defined in each region to hold the long time series of skipjack length composition data from Japanese research longline cruises in the WCPO and more recently, observer-measured length composition samples. Catch data from the Japanese troll fishery in Region 1 was combined with the pole-and-line fishery in this region. Miscellaneous gears from Indonesia, Philippines, and Vietnam were aggregated by nationality, with the exception of a fishery created for the domestic Indonesia and Philippines purse-seine fleets.

Several changes were made to the 2016 fisheries definitions due to the transition to an 8 region model and these are presented in detail by Vincent (2019). The number of fisheries increased from 23 to 31 fisheries to account for the addition of 3 regions (see Table 1 for fishery descriptions).

4.5 Catch and effort data

4.5.1 General characteristics

Catch and effort data were compiled by year and quarter according to the fisheries defined in Vincent (2019). The catches of all fisheries were expressed in weight of fish, with the exception of the longline fishery (the catches of which are very small and set at a nominal level) which were expressed in numbers of fish.

Total annual catches by major gear categories for the WCPO are shown in Figure 3 and a regional breakdown is provided in Figure 4. The spatial distribution of catches over the past ten years is provided in Figure 6. Discarded catches are estimated to be minor and were not included in the analysis. Catches in the northern region are highly seasonal.

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

- The development of the Japanese off-shore purse-seine fishery in Region 1, 2, and 3 since the mid-1990s.
- The virtual cessation of the domestic pole-and-line fisheries in PNG and Fiji and the recent low catches from the Solomon Islands fishery;
- The general decline in the Japanese distant-water pole-and-line fisheries in the equatorial regions, particularly in the eastern region.
- 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 to the east where the majority of recent harvest has occurred.
- 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, Philippines, and Vietnam.

4.5.2 Purse seine

Purse seine catch by species within each set type (associated or unassociated) is estimated by applying estimates of species composition from observer-collected samples to total catches estimated from raised logsheet data (Hampton and Williams, 2016). For the Japanese (JP) fleet for which there is greater confidence in species-based reporting, we use reported catch by species rather than estimating it.

Unlike previous assessments, effort data units for purse seine fisheries are defined as number of sets, specified by set type (associated or unassociated), to be consistent with proposed management strategy evaluation models. The change from days fishing and/or searching to number of sets did not affect model results but avoids the issue of effort creep from changes in reporting practices that attribute fewer days as searching (SPC-OFP, 2013). Time-varying changes in catchability and a low weight to the effort deviations were assumed for these fisheries within the stock assessment model.

In Regions 5 and 6 where effort of the JP pole-and-line fleet was not adequate for conducting CPUE standardizations, purse seine fisheries instead received standardized CPUE. CPUE for the Philippines domestic purse seine was analyzed using generalized linear models (GLMs) by Bigelow et al. (2019), who produced a standardized index of abundance (Figure 7). These indices were applied to the catches of the S-ID.PH-5 fishery for the years 2005–2018. Estimates of time-variant precision were implemented as time-variant effort deviation penalties in MULTIFAN-CL.

CPUE data for the purse seine fishery operating largely within the PNG archipelagic waters (Region 6) was also analyzed using GLMs (Vidal et al., 2019; Figure 7). These indices were applied to the catches of the SA-All-6 fishery for the years 1997-2018 and time-variant effort deviation penalties were again estimated and used.

4.5.3 Longline

Longline fisheries take a negligible proportion of the total skipjack catch. These fisheries were included in the model solely to utilize the available size frequency data and to attribute tag returns. Catches are set at very low, arbitrary levels with low fitting penalties and this is detailed later in the document.

4.5.4 Pole-and-line

Standardized CPUE indices of abundance applied to fisheries 1, 4, 7, 22, 24, and 28 were based on Japanese pole-and-line fisheries in Regions 1, 2, 3, 4, 7, and 8, respectively. These standardized indices were estimated using GLMs fitted to operational catch and effort data (Kinoshita et al., 2019). The 2016 skipjack tuna assessment used a traditional GLM approach where separate analyses were conducted for each region on either the distant-water (DW) or off-shore (OS) pole-and-line fleet, which was repeated for the current assessment with three additional years of data. In addition, a geostatical CPUE model was fit to the DW and OS data combined and included spatial and environmental components. The traditional GLM approach was used in preliminary steps toward the development of the diagnostic case model for 2019. The geostatistical CPUE data were used in the diagnostic model and the structural uncertainty grids (Figure 7). The uncertainty in each pole-and-line CPUE estimate from the GLM, by fishery and time, was included in the model by way of time-variant penalty weights for the effort deviations (Figure 8).

Nominal fishing-vessel-day was used as the unit of effort for the pole-and-line fisheries of PNG and Solomon Islands in Region 6, and the mixed flag (but mainly ID) pole-and-line fishery in Region 5.

4.5.5 Other fisheries

Effort data for the Indonesian, Philippines, and Vietnam surface fisheries and research longline fisheries were unavailable. Where effort data are absent, the model directly computes fishing mortality consistent with the observed catch using a Newton-Raphson procedure. This is covered in more detail in Section 5.3.3.

4.6 Size data

4.6.1 Purse seine

Only length frequency samples are used in the skipjack assessment. The methods developed for the 2014 skipjack stock assessment were conducted to create the length frequency samples for the purse seine fisheries. These methods and results are described in detail in other papers (Abascal et al., 2014; Vincent et al., 2019), so will only be briefly explained here. The length frequency data comprise

samples derived from long-term port sampling data of primarily US purse seiners in Pago Pago, and samples corrected for grab-sample bias (Lawson, 2011, 2013) taken by at-sea observers from most large purse seine vessels operating in the fishery since 2010. All samples were spatially weighted by the catch at the set type and spatial strata levels (5°x5° squares), with thresholds applied to ensure small samples were not given too much weight. Similar methodologies were conducted for the longline fisheries in the bigeye and yellowfin tuna assessments in 2017 (McKechnie et al., 2017; Tremblay-Boyer et al., 2017). The number of length measurements capped at a maximum of 1000 fish per fishery incident are shown in Figure 9.

4.6.2 Longline

Longline fisheries typically do not target skipjack but do catch small numbers of large skipjack within the 50–90 length range as bycatch, and the catch is usually discarded. Japanese research vessels have routinely collected measurements of the length of skipjack caught since the start of the assessment time period. Japanese research data is only available sporadically in several regions and sample sizes have decreased or ceased to exist in areas where JP longline fishery effort has declined. For the 2019 stock assessment, data for other flags conducting longline fishing were added to these fisheries, resulting in substantial increases in sample size and temporal coverage. Most of these samples were collected by observer programs (since the 2000's) and the datasets for these fisheries are covered by Vincent (2019).

These data are included to provide information to the model on the existence of these larger sized skipjack rarely caught in purse seine or pole-and-line fisheries. The size data from the longline fishery is an important component of the model because it provides information regarding the larger-sized skipjack that are not typically caught by the purse seine and pole-and-line surface fisheries. Without this information, the model would have difficulty estimating this "cryptic biomass" component of the population.

4.6.3 Pole-and-line

Size composition for pole-and-line fisheries primarily come from observer data, with the exception of Regions 1, 2, 3, 4, and 7 where length data are available from the Japanese offshore and distant-water fleet from the beginning of the model period until 2018 (Vincent, 2019). The length frequency for the Japanese pole-and-line fishery in the database held by SPC were revised based on the methods described in Kiyofuji et al. (2019b). Length data for the equatorial pole-and-line fisheries are available from both the Japanese distant-water fleet and domestic fleets.

The data from the pole-and-line fishery in Region 8 (P-ALL-8) was dominated by data for the Japanese fleets (1974–2004) with additional data from Fiji in the 1990's. Length data from the pole-and-line fishery in Region 5 (P-All-5) was filtered to only include data from Indonesia because

it constituted the majority of the catch. Data for Indonesia are sporadically available early in the time series, which precluded weighting samples from other countries based on the catch, but were readily available in 2009–2018. The data from the pole-and-line fishery in Region 6 (P-ALL-6) are a large dataset for multiple countries with the Solomon Island and PNG contributing the majority of length composition samples. The pole-and-line fisheries in the northern regions generally catch smaller fish than the equatorial fisheries (Regions 5–8), although over the model period, there was a slight general increase in the length of fish sampled from the pole-and-line fisheries in the four northern regions, with a substantial change in Region 3 during the early 2000s (Vincent, 2019). No systematic trend in the length composition was evident in Regions 5–8.

4.6.4 Other fisheries

Size composition data for the Philippines domestic fisheries (Z-PH-4) were collected by a sampling program conducted in the Philippines in 1993–95 and augmented with data from the 1980s. In addition, data collected during 1997–2006 under the National Stock Assessment Project, and in more recent years under the GEF-WPEA project, were included in the current assessment. Unlike the assessment in 2016, the Indonesia domestic fishery (Z-ID-4) was given its own selectivity function as a result of the addition of three years of data and the large number of measurements from recent sampling under the GEF-WPEA project. Similarly, data for the Vietnam domestic fishery (Z-VN-4) numbers nearly a hundred thousand samples in recent years and was given a separate selectivity in the model (Table 2).

4.7 Tagging data

A large amount of tagging data was available for incorporation into the assessment and a summary of its characteristics and the process of constructing the MULTIFAN-CL tagging files is presented in Vincent et al. (2019). The data were available from SPC's Skipjack Survey and Assessment Program (SSAP) carried out during 1977–80, the Regional Tuna Tagging Project (RTTP) during 1989–92 (including affiliated in-country projects in the Solomon Islands, Kiribati, Fiji and the Philippines), and the Pacific Tuna Tagging Program (PTTP) which has been ongoing since 2006 (Figure 10). Tags were released using standard tuna tagging equipment and techniques by trained scientists and technicians. Tags have been returned mostly from purse seine vessels via processing and unloading facilities throughout the Asia-Pacific region.

Tagging data from regular Japanese research tagging cruises were available for the period 1989–2017. Tag releases prior to 1998 were not included in the 2016 assessment because the tag releases were not measured at the time of tagging. However since the number of pre-1998 releases is substantial and potentially informative, we evaluated several methods for assigning release lengths to these early data (Vincent et al., 2019). Ultimately, we opted to assign missing release lengths by sampling from the measured release lengths for the Japanese tagging program.

As in recent tropical tuna assessments, the numbers of tag releases were adjusted for a number of sources of tag loss, e.g., 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. The procedures used in re-scaling the releases are described in detail in Berger et al. (2014), McKechnie et al. (2016) and Vincent et al. (2019). 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.

There is a delay between tagged fish being caught, the tag being reported, and the data being entered into the tagging database. 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. This was a particular issue for the 2017 PTTP tag releases, for which considerable numbers of tags were known to have been recovered in fishing ports, but not yet returned to SPC. Therefore, these releases were excluded from the assessment (Vincent et al., 2019). For the Japanese tagging program, tags are generally returned more promptly; thus it was possible to include tag releases to the end of 2017 in the assessment.

For incorporation into the assessment, tag releases were stratified by release region, year/quarter of release, and length at release using the same size bins as the length-frequency data (Figure 11). Tag release events that had less than 100 tags released per event were removed from the analysis to reduce the computational load otherwise created by the many small release events, and allow for better model convergence. A total of 329,811 effective releases were classified into 269 tag release groups (Table 3). The returns from each size-class of each tag release group (56,092 total usable tag returns) were then classified by recapture fishery and time period (quarter). A summary of tags recaptured by release and recapture regions is presented in Figure 12. Tag return data were aggregated across set types for the purse seine fisheries in each region because tag returns by purse seiners were often not accompanied by information concerning the set type. The assessment model was configured to predict tag recaptures by these fisheries grouped together by region.

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. (2019), but a brief summary will be given below. In addition, we describe the procedures followed for estimating the parameters of the model and the way in which stock status conclusions

are drawn using a series of reference points.

5.2 Population dynamics

The model partitions the population into eight spatial regions and 16 quarterly age-classes. The last age-class comprises a "plus group" in which mortality and other vital rates are assumed to be constant. The population is "monitored" in the model at quarterly time steps, extending through a time window of 1972–2018. 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 $\sim$ 23 cm given the current growth curve) in the population. Tropical tuna spawning does not always follow a clear seasonal pattern but occurs sporadically when food supplies are plentiful (Itano, 2000). The assessment model 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.

The proportion of total recruitment occurring in each region was initially set relative to the variation in average regional catch and then estimated during the later phases of the fitting procedure. The distribution of recruitment among the model regions was estimated within the model and allowed to vary over time in a relatively unconstrained fashion.

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 to prevent the instability that has been detected by retrospective analyses. This approach has been continued here with the terminal two recruitments fixed at the geometric mean, which is appropriate for a log-normally distributed random variable.

Spatially-aggregated recruitment was assumed to have a weak relationship (CV of log-recruitment deviates set to 2.2) with total spawning potential in the preceding quarter, according to 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). As has been the practice in other tuna stock assessments, h was fixed at 0.80 in the diagnostic model, and values of 0.65 and 0.95 included in the structural uncertainty grid (ISSF, 2011).

The high CV (2.2) of the log-recruitment deviates, computed annually, ensured that the SRR had negligible impact on the estimation of recruitment and other model parameters, as recommended by Ianelli et al. (2012). The SRR was estimated over the period 1982–2017 to prevent the earlier

recruitments, which appear to be part of a less productive recruitment regime, from influencing the relationship.

5.2.2 Initial population

The population age structure in the initial time period in each region was assumed to be at equilibrium and determined as a function of the average total mortality during the first 20 quarters. This assumption avoids having to estimate independent parameters for each age and region, which are generally poorly determined.

5.2.3 Growth

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

The estimation of growth within the skipjack stock assessment model was dependent upon the influence assigned to the length frequency data. Due to the conflict with other data sources, the influence of the length composition in the model was decreased (i.e., high scalar value), which made the resulting estimates of growth unreasonable. A multi-step approach was taken to estimate growth curves that were fixed for models in the structural uncertainty grid. Estimates from a gridded search over each growth parameter (e.g., length at first age-class, length at last age-class) that provided the best likelihood within the assessment model were used as the starting values of growth parameters in subsequent models. We determined that the length frequency scalars influenced the estimated growth curve parameters. Therefore, three assessment models with different length frequency scalars (1, 20 and 40), which were otherwise similar to the diagnostic model, were used to estimate the growth parameters including the standard deviation parameters. These values of the length frequency scalars were chosen to incorporate the entire range of scalars where estimates of growth parameters were not hitting the bounds. The estimated growth curve that resulted from the model with a length frequency scalar of 20 was used as the diagnostic case. The two remaining growth curves were used as alternative options in the structural uncertainty grid for the axis of growth.

5.2.4 Movement

Movement was assumed to occur instantaneously at the beginning of each quarter. Parameters were estimated for regions that shared a common boundary, but 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., 2019 for details). Movement is parameterized 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. Movement was assumed to be constant across ages, dissimilar to the previous assessment, because the increased number of parameters for the additional regions did not yield sufficient improvement in model fit to the data to warrant their inclusion. Alternative models with age-dependent movement were considered, but are not presented. The seasonal pattern of movement persists from year to year with no allowance for inter annual variation in movement. A prior mean of 0 was assumed for all movement coefficients, but the penalty was weak for deviations from the mean. Starting parameter estimates for the movement coefficients were specified based on the proportion of tags returned within each region and quarter relative to total tags returned for releases in the same region (Figure 12).

5.2.5 Natural mortality

Natural mortality was estimated internally in the assessment model using a cubic spline with 5 nodes and was assumed to be age-specific, but invariant over time and region.

5.2.6 Sexual maturity

Age-specific sexual maturity was computed internally from a specified maturity-at-length ogive and the growth parameters, which were also specified in this assessment. At the 2019 PAW (Pilling and Brouwer, 2019), there was discussion surrounding the appropriate use of the maturity-at-length data within the stock assessment. The majority of samples from the biological study came from the temperate and subtropical regions, but the majority of the population is believed to occur in the tropical region (Ohashi et al., 2019). The PAW decided to use the maturity-at-length data from the tropical region only for the diagnostic model and all models in the structural uncertainty grid. The maturity-at-length ogive for skipjack tuna sampled from tropical waters (Ohashi et al., 2019) was modeled in the assessment with the equation:

$$P_l = \frac{1}{1 + exp(7.414 - 0.148 * FL)} \tag{1}$$

where the proportion mature for each length bin was calculated using Equation (1) and the center of the length bin (Figure 13).

The PAW also recommended use of the biomass estimate in the tropical, subtropical, and temperate regions to weight the corresponding maturity data in a one-off sensitivity model. Therefore, a maturity ogive that weighted the samples according to the estimated proportion in the temperate (5%), subtropical (27%), and tropical (68%) regions was created by resampling the data 100 times according to these percentages. A GLM to estimate the proportion mature at length was fit to these datasets and then the average of the parameters was taken. The equation of the regionally weighted maturity ogive is:

$$P_l = \frac{1}{1 + exp(6.699085 - 0.132 * FL)} \tag{2}$$

This yielded a maturity-at-length estimate similar to the maturity ogive from the tropical region only (Figure 13), which was used in the one-off sensitivity model.

Unlike *Thunnus* species, the sex ratio for skipjack does not appear to vary with size. Sex ratio and fecundity at size were not included in the maturity parameter, so in this assessment the term "spawning potential" refers to the biomass of adult fish, rather than female spawning potential as in the yellowfin, bigeye, and albacore stock assessments.

5.3 Fishery dynamics

5.3.1 Selectivity

Selectivity is often modeled as a functional relationship with age to reduce the number of parameters estimated within the stock assessment model. Examples include a logistic curve to model monotonically increasing selectivity and various dome-shaped curves to model fisheries that select neither the youngest nor oldest fish. Modeling 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. Instead, we have used a method based on a cubic spline interpolation. 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 characterize selectivity over the age range. The number of nodes varied by fishery and were selected because they provided an improvement in Akaike Information Criterion (AIC), increased model stability by reducing the number of parameters, or removed unreasonable trends in the selectivity-at-age.

The specific selectivity function for each fishery is presented in Table 2. In all cases, selectivity was assumed to be time-invariant and fishery-specific. However, a single selectivity function could be "shared" among a group of fisheries that have similar length compositions or were assumed to operate in a similar manner. This grouping facilitates a reduction in the number of parameters estimated and can provide insight into the regional abundance of fish of specific sizes.

Selectivities of the longline fisheries were assumed to be logistic functions of age and were grouped

into three shared functions due to concerns regarding the limited number and temporal distribution of available samples. The primarily Japanese pole-and-line fisheries in Regions 1-4, 7, and 8 were assumed to share selectivity (4 node cubic spline function). Selectivity for the equatorial purse seine fisheries was grouped by set type, i.e., selectivity was shared among regions separately for associated and unassociated sets fisheries. The Indonesian, Philippines and Vietnam domestic fisheries in Region 4 (Z-ID-4, Z-PH-4, and Z-VN-4) were given separate selectivity functions with differing numbers of nodes (Table 2).

5.3.2 Catchability

Seasonal but year-invariant catchability was estimated for fisheries providing standardized abundance indices (Table 2). This assumption is similar to assuming that the CPUE for these fisheries indexes the exploitable abundance over time. No grouping of catchability for these fisheries was employed, and therefore the relative level of CPUE was not used to scale the relative exploitable biomass amongst regions - the model relies on other data, size and most importantly tagging, to estimate the regional distribution of abundance. Significant quantities of tagging data exist in all 8 regions which provides information on region-specific fishing mortality and hence biomass. These data allow catchabilities for these fisheries to be estimated independently. The relative regional population sizes are estimated relatively freely by the model.

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 approximating a CV of 0.10. The catchability of fisheries that were not based on annual estimates of harvest were modeled as a sinusoidal seasonal function, including those fisheries providing a standardized CPUE index (Table 2). For fisheries having no available effort estimates (e.g. the Philippines, Indonesian, and Vietnamese domestic 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.

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 2. For fisheries which provided a standardized CPUE indices, the prior was set to mean zero and time-variant CV based on the variance estimates. The traditional GLMs used the canonical variance method of Francis (1999) to estimate the time-variant CV, whereas the geostatistical model estimated this quantity internally (Bigelow et al., 2019; Kinoshita et al., 2019; Vidal et al., 2019).

Effort estimates are not available for the ID/PH/VN miscellaneous fisheries or the longline fisheries; these fisheries were assigned missing effort for all time steps except the terminal four quarters, for which effort was arbitrarily specified as 1.0. This was done only to provide a basis for effort-based projections of these fisheries in management analyses.

For all other fisheries, nominal effort was provided assuming moderate effort deviation penalties, averaging 0.41 with individual penalties proportional to the square root of the effort (Table 2).

5.4 Dynamics of tagged fish

Tagged fish are modeled as discrete cohorts based on the region, year, quarter, and length at release for the first 12 quarters after release. The tags released are assigned to an age bin according to the length at release and estimated growth curve and its associated standard deviations. Subsequently, the tagged fish are pooled into a common group in order to limit memory and computational requirements. Increasing the number of quarters before tag pooling occurred did not have a significant impact on model estimates.

5.4.1 Tag reporting

In theory, tag-reporting rates can be estimated internally within the stock assessment model. In practice, experience has shown that independent information on tag-reporting rates for at least some fisheries tends to be required for reasonably precise estimates to be obtained. We provided reporting rate priors for all reporting groups that reflect independent estimates of the reporting rates and their variance. We also made some informed assumptions regarding fisheries that were similar to those with independent estimates but increased the prior variance. The values of the priors and rationale for these assumptions are presented in Vincent et al. (2019).

Previous assessments have assumed fishery-specific reporting rates are constant over time. This assumption was reasonable when most of the tag data were associated with a single tagging program. However, tag reporting rates may vary considerably between tagging programs due to changes in the composition and operation of individual fisheries, and different levels of publicity and follow-up. Consequently, fishery-specific tag reporting rates that are also specific to individual tagging programs were estimated. Tag recapture and reporting rate groupings are provided in Table 2, where the same groupings were assumed across all four tagging programs.

The estimation of the reporting rates included penalty terms in respect of pre-determined priors. These were derived from analyses of tag seeding experiments (Peatman, 2019). For the RTTP and PTTP, relatively informative priors were formulated for the equatorial purse seine fisheries given the larger extent of information available. Reporting rate priors from the RTTP were assumed for the SSAP with an increased variance. All reporting rates within a tagging program were assumed to be time-invariant.

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. It is therefore desirable to designate one or more time periods after release as "pre-mixed" and compute fishing mortality for the tagged fish based on the actual recaptures, corrected for tag reporting (see below), rather than use fishing mortalities based on the general population parameters. This, in effect, desensitizes 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 skipjack gradually mix with the untagged population at the regional level and that this mixing process is complete by the end of the quarter of release (mixing period of one quarter). We also tested the sensitivity of this assumption by increasing the mixing period to the end of the quarter after release (mixing period of two quarters).

5.5 Likelihood components

There are three data components that contribute to the log-likelihood function for the skipjack stock assessment - the total catch data, the length-frequency data, and the tagging data. The observed total catch data for non-longline fisheries are assumed to be unbiased and relatively precise, with the SD of residuals on the log scale being 0.007. For the longline fisheries for which arbitrary small catches were specified, the residual SD was set to 0.03.

The probability distributions for the length-frequency proportions are assumed to be approximated by robust normal distributions, with the variance determined by the effective sample size and the observed length-frequency proportion. Length frequency samples are assigned effective sample sizes lower than the number of fish measured. Reduction of the effective sample size recognizes 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.

The size data were considered to be moderately informative and were assigned moderate weight in the likelihood function. To prevent some time periods with large samples sizes from dominating the likelihood, the maximum number of fish sampled per fishery within a single time-step was set at 1,000 fish. These sampled fish were then assigned an effective sample size of 0.01 times the actual sample size, which corresponds to a maximum effective sample size of 10. This is lower than the previous assessment and resulted from investigation of the conflict between the size data, tagging data, and the CPUE data (see Section A.1 for further details). Alternative scalars for down-weighting the length composition data were explored in the sensitivity analyses (Section 6.2).

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 parameterization of the overdispersion parameter (τ) 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 recognize this. This should then provide a more realistic weighting of the tag return data in the overall log-likelihood and allow the variability in tag returns to impact the confidence intervals of estimated parameters. Therefore, we allowed the overdispersion parameter (τ) to be estimated by the assessment model and conducted a likelihood profile over a range of values. A complete derivation and description of the negative binomial likelihood function for tagging data is provided in Kleiber et al. (2019).

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 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 skj.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 model to obtain estimates of the covariance matrix. The covariance matrix is used in combination with the delta method to compute approximate confidence intervals for parameters of interest (for example, the total biomass and recruitment trajectories). For the second approach, a factorial grid of model runs was undertaken which incorporated many of the uncertainties explored in the one-off sensitivity

analyses. This structural uncertainty grid attempts to describe the main sources of structural and parameter uncertainty in the assessment. Previous experience has shown that overall uncertainty is dominated by the structural uncertainty grid. Because of the computational demands of computing the Hessian, it is not feasible to do this for all models in the grid. Therefore, the uncertainty in the main quantities of stock assessment interest are based only on the structural uncertainty grid.

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 undertaken of the marginal posterior likelihood in respect of population scaling, following the procedure outlined by McKechnie et al. (2017) and Tremblay-Boyer et al. (2017). This procedure is presented in the Appendix (Section A.1).

Retrospective analyses were conducted 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 2019 diagnostic model are presented in the Appendix (Section A.3.1).

5.7 Stock assessment interpretation methods

Several ancillary analyses using the converged model/s were conducted in order to interpret the results for stock assessment purposes. The methods involved are summarized below and the details can be found in Kleiber et al. (2019).

5.7.1 Yield analysis

The yield analysis consists of computing equilibrium catch (or yield) and spawning potential, 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. (2019), and is equivalent to the MSY. Similarly, the spawning potential at MSY (SB_{MSY}) can be determined from this analysis. The ratios of the current (or recent average) levels of fishing mortality and spawning potential to their respective levels at MSY are determined for all models of interest. This analysis was conducted for all models in the structural uncertainty grid and thus includes alternative values of steepness assumed for the SRR (see Section 6.2).

Fishing mortality-at-age (F_a) for the yield analysis was determined as the mean over a recent period of time (2014–2017). We do not include 2018 in the average as fishing mortality tends to have high uncertainty for the terminal data year of the analysis and the catch and effort data for this terminal

year are potentially incomplete. Additionally, recruitments for the last two quarters of the terminal year of the model are constrained to be the geometric mean across the entire time series, which affects the F for the youngest age classes.

MSY was also computed using the average annual F_a from each year included in the model (1972–2018). 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.

5.7.2 Depletion and fishery impact

Fishery depletion was calculated by computing the unexploited spawning potential time series (at the region level) using the estimated model parameters, but assuming that fishing mortality was zero. Both the estimated spawning potential SB_t (with fishing) and the unexploited spawning potential $SB_{F=0[t]}$ incorporate recruitment variability. Therefore, the ratio of these two quantities 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 spawning potential 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 equilibrium recruitment between the fished and unfished spawning potential estimates.

A similar approach was used to estimate depletion associated with specific fisheries or groups of fisheries. Here, fishery groups of interest (purse seine free-school sets, purse seine associated sets, purse seine unidentified, pole-and-line, longline, and miscellaneous fisheries), are removed in-turn in separate simulations. The changes in depletion observed in these runs are then indicative of the depletion caused by each of 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 (Section 5.2.1). 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 skipjack stock where $SB_{F=0}$ for this assessment is calculated as the average over the period 2008–2017. The recently adopted target reference point (TRP) for this stock is $50\%SB_{F=0}$ where $SB_{F=0}$ is again calculated over 2008–2017. Stock status was referenced against these points by calculating the reference points; $SB_{recent}/SB_{F=0}$ and $SB_{latest}/SB_{F=0}$ where $SB_{F=0}$ is calculated over 2008–2017 and SB_{latest} and SB_{recent} are the estimated spawning potential in 2018, and the mean over 2015–2018, respectively (Table 4).

The other key reference point, $F_{recent}/F_{\mathrm{MSY}}$, is the estimated average fishing mortality at the full assessment area scale over a recent period of time (F_{recent}; 2014–2017 for this stock assessment)

divided by the fishing mortality producing MSY which is a product of the yield analysis and was detailed in Section 5.7.1.

5.7.4 Kobe analysis and Majuro 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 (2014–2017 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 by repeating the yield analysis for each year in turn. This enabled temporal trends in the reference points to be estimated and a consideration of the differences in MSY levels under 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 stock assessments in recent years.

5.7.5 Stochastic projections from the structural uncertainty grid

Projections of stock assessment models can be conducted routinely within MULTIFAN-CL to ensure consistency between the fitted model and the simulated future dynamics, and the framework for performing this exercise is detailed in Pilling et al. (2016). One hundred future population trajectories were simulated for each model in both structural uncertainty grids. A summary of the results is given in the Appendix (Section A.2).

6 Model runs

6.1 Developments from the last assessment

The progression of model development from the 2016 assessment reference case model to the 2019 diagnostic model was incremental, to ascertain the consequence of each modification. Changes made to the previous assessment model include additional input data for the years 2016-2018, changes to the fisheries structures (Vincent (2019)), modified methods in producing the input files (Vincent et al. (2019); Vidal et al. (2019); Kinoshita et al. (2019); Bigelow et al. (2019)), incorporation of additional biological data, and implementation of additional or new features in MULTIFAN-CL (Davies et al., 2018). An outline of the model development progression is as follows:

1. The 2016 reference case model [*Ref2016*]
2. The 2016 reference case model with the new MULTIFAN-CL executable [*S1NewExe*]
3. Addition of updated data inputs (catch, CPUE, tagging, and length composition) extending until the end of 2018 and revision of some data sources (see Vincent, 2019). Additionally

modifications were made to the selectivity functions to improve model fit and stability and the updated length-weight relationship was incorporated. [*S2NewData*]

4. Substitution of the traditional standardized Japanese CPUE with a geostatistical CPUE that combines the offshore and distant water information into an index for each region. [*S3GeoCPUE*]
5. Removal of tag release events (year-quarter and region) with less than 100 tags released and updating reporting rate priors with additional information since the 2016 assessment. [*S4TagGT100*]
6. Inclusion of additional Japanese tagging data, whereby missing release lengths were specified by sampling from the observed release length distribution. [*S5JPTags*]
7. Creation of an 8 region model [*S6_8Region*]
8. Growth was fixed at an estimated value from an assessment model specifically created for the estimation of growth and natural mortality was estimated as a smoothed spline with 5 nodes [*S7GrowthM*]
9. Incorporation of a maturity-at-length ogive using a new MULTIFAN-CL feature [*S8Maturity*]
10. Estimation of the variance parameter (overdispersion τ) of the tagging data to fit the data with more a realistic recognition of tag-return variability [*S9ODEst*]
11. Reduction in the influence of the length composition data by increasing the scalar to 100 [*S10Length100*]

6.2 Sensitivity analyses and structural uncertainty

A vast multitude of models were created in the preparation of the 2019 skipjack stock assessment, but we will restrict the presentation of results to a small set models that are described in further detail below. These axes were used for "one-off" changes from the diagnostic model and several of these sensitivity models were used as axes in the structural uncertainty analyses. The structural uncertainty analysis is conducted by creating a grid of model runs where all possible combinations of the assumptions are explored (see Section 6.3).

The recommendations of the 2019 PAW formed the basis for several of the one-off sensitivity analyses undertaken from the diagnostic model. Each one-off sensitivity model was created by making a single change to the diagnostic model and only one was not included in the structural uncertainty grid. As such, we present the summarized reference points for all models in the structural uncertainty grid for management use. However, we provide the reference point estimates from the one-off sensitivity models in the Appendix (Table A2).

6.2.1 Steepness [*H0.65, H0.95*]

As has been the case in other tuna assessments, we assumed a value of 0.8 for the diagnostic model, 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 (Anon, 2011) and is well established in previous Scientific Committees.

6.2.2 Tag mixing period [*Mix2*]

The tag mixing period is imposed to allow tagged fish time to distribute themselves throughout the region of tagging, although it is somewhat difficult to ascertain how long this period should be. In the diagnostic model, the mixing period was set at one quarter, which means that tags recaptured in the quarter of release are assumed to have mortality rates different from the rest of the population. An alternative model was run assuming a mixing period of two quarters, which means tags recaptured in the quarter of release and the quarter after are not influential in the model.

6.2.3 Scalar of length composition data [*Length50, Length200*]

The difficulties in assigning weighting to the length composition data were discussed in Section 5.5. Two alternative models were considered: a model where the length composition data were given more influence relative to the diagnostic model (*Length50*; corresponds to a maximum effective sample size of 20 fish) and another where those data were given less influence relative to the diagnostic model (*Length200*; corresponds to a maximum effective sample size of 5 fish).

6.2.4 Alternative growth functions [*GrowthLow, GrowthHigh*]

The paucity of reliable modal progression in the length composition data means that it can be difficult to estimate all growth parameters internally in the model. However, it was determined that reasonable growth parameters could be estimated, but were dependent on starting the parameters at reasonable values away from the lower bounds. Additionally, it was discovered that different scalars of the length composition data would influence the estimated growth curve. When the length scalar was above 50 the estimates of growth became unreliable with estimates hitting the lower bound for the length at first age parameter. The estimated growth from a model with a length scalar of 20 was used as the assumed growth rate for the diagnostic model. Two other growth curves estimated with values for the length scalar of 1 and 40 were used in separate one-off sensitivities from the diagnostic model, respectively referred to as *GrowthLow* and *GrowthHigh* (Figure 14).

6.2.5 Alternative Spatial Structure [*5Region*]

A model was created that used the regional structure from the 2016 skipjack tuna stock assessment. Given the difference in the spatial structure, there was an unavoidable difference in the number of fisheries, but assumptions regarding selectivity, catchability, length scalars, etc. were made as consistent as possible between the two models. We present the two spatial structures separately in the structural uncertainty analysis (Section 6.3).

6.2.6 Maturity ogive weighted by regional total biomass [*WtdMaturity*]

A model that used the maturity-at-length ogive with weighted samples from the tropical, subtropical, and temperate regions was created to test the impact of including the non-tropical maturity data (Figure 13; WtdMaturity).

6.3 Structural uncertainty

Stock assessments of pelagic species in the WCPO in recent years have utilized an approach to assess the structural uncertainty in the assessment model by running a "grid" of models that explore the interactions among selected "axes" of uncertainty. The grid contains all combinations of levels of several model quantities, or assumptions, and allows the sensitivity of stock status and management quantities to this uncertainty be determined. 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.

Two separate structural uncertainty grids were run for the 2019 skipjack stock assessment: one for the 8 region model and another for the 5 region model. The grids were constructed from 5 axes of uncertainty with 2–3 levels for each. The values for the diagnostic model are in bold and the levels used in the grid are directly comparable to those presented in Section 6.2 through identical notation. The levels of the grid are:

1. Steepness [0.65 (*H0.65*),**0.8**, and 0.95 (*H0.95*)]
2. Growth [Functions that were estimated using a length scalar of: 1 (*GrowthLow*), **20** (*Diagnostic Growth*), and 40 (*GrowthHigh*)]
3. Length composition scalar [(*50*), **100**, and (*200*)]
4. Assumed tag mixing period [**1** and 2 (*Mix2*)]

This resulted in two grids of 54 models each (Table 5).

7 Results

7.0.1 Consequences of key model developments

The progression of model development from the 2016 to the 2019 diagnostic model is shown in Section 6.1 and the results are displayed in Figures 15 and 16. A summary of the consequences of this progression through the models is as follows:

1. The reference case model for the 2016 assessment refit with the latest version of MULTIFAN-CL (green line; Figure 15) produced a nearly identical result to the 2016 version.
2. The addition of three more years of data, changes to the selectivity curves and length weight-relationship produced results relatively similar estimates of depletion to the previous model but there were differences in the mid-1990s and after 2012 where this model predicted a more depleted status (light purple line; Figure 15). However, the estimated spawning potential for this model was generally smaller than the previous model, particularly at the beginning of the time series.
3. Transitioning to geostatistical CPUE results in a larger spawning potential estimate and less depleted state at the beginning of the time series, but after the 1990s estimates were generally similar until 2010 (red line; Figure 15). After 2010, the geostatistical CPUE model estimated the stock to be more depleted with a smaller spawning potential than the traditional GLM model.
4. The removal of tag release events with fewer than 100 tags release resulted in only minor changes in depletion and spawning potential estimates (blue line; Figure 15).
5. The addition of tags that were missing release lengths from the JPTP resulted in a less depleted stock status and a larger spawning potential compared to the previous step (orange line; Figure 15).
6. The transition to the 8 region model resulted in a much less depleted stock status and a spawning potential that was larger than at the previous step (green line; Figure 16).
7. The estimation of natural mortality as a cubic spline and fixing growth at previously estimated values resulted in depletion estimates similar to the previous model but spawning potential was consistently less (light purple line; Figure 16). These two steps were combined because the estimation of a natural mortality spline did not have a significant effect.
8. The application of the maturity-at-length data to the assessment model resulted in a more depleted state with a lower spawning potential (red line; Figure 16).
9. The penultimate model where overdispersion of the tagging data was estimated resulted in a decrease in the estimated spawning potential in the model and a more depleted status (blue line; Figure 16).

10. In the final step in the progression (diagnostic model), increasing the length composition scalar (i.e., decreasing influence) resulted in estimates of depletion that were less depleted than the previous step prior to the mid 2000s and in the last year of the assessment (orange line; Figure 16). The estimated spawning potential for this step was larger than the previous step for most of the time series.

7.1 Fit of the diagnostic model to data sources

7.1.1 Catch data

A high penalty was applied to the catch data for all fisheries except the longline fishery because these fisheries were assigned arbitrary values. As a result, the model fit to the catch data is nearly exact (Figure 17).

7.1.2 Standardized CPUE

There was substantial temporal variability in the standardized CPUE indices used in the assessment, but despite this the model-predicted CPUE fit the indices relatively well, particularly the long term trends in observed abundance (Figures 18 and 19). The model largely captured the substantial seasonal variability observed for the pole-and-line fishery in Regions 1, 2, 3, and to some extent 4 (P-ALL-1, P-ALL-2, P-ALL-3, and P-ALL-4). There was a general trend among the temperate and equatorial pole-and-line fisheries (P-ALL-3, P-ALL-4, P-ALL-7, and P-ALL-8) that the increase in seasonal variability in CPUE during the 1990s was poorly fit for the high observations but well fit for the lower observations. This corresponds with a tendency for the effort deviations in Regions 7 and 8 being positive (Figure 20 P-ALL-7 and P-ALL-8).

The model poorly predicted the variability in the CPUE for the Indonesia and Philippine purse seine fishery (S-ID.PH-5), but estimated the general trend well (Figure 18). The CPUE for the associated purse seine fishery in Region 6 was generally well fit by the model.

The CPUE of fisheries not receiving standardized CPUE indices have far less influence in estimating temporal trends in total biomass. This is a consequence of the lower effort deviation penalties applied and time-varying catchability in these fisheries. Temporal trends in effort deviations can be indicative of model misspecification of temporal changes in catchability. Effort deviation estimates were centered around zero and did not display any temporal trends (Figure 21).

7.1.3 Size composition data

The length composition data were reasonably fit by the model as shown by the comparison of the observed and predicted numbers-at-length (Figures 22 to 25). There is some overestimation by the

model for lengths larger than about 65 cm for the pole-and-line fisheries in Regions 6, 7, and 8. The pole-and-line fishery in Region 5, which was not grouped with the other pole-and-line fisheries, overestimated the frequency of fish of length of about 35 cm but underestimated the frequency of fish around 50 cm. On the other hand, the estimates for the longline fisheries are well fit by the model Figure 23). For the purse seine fisheries there is some tendency to overestimate the smallest and largest lengths in some regions, but this is likely due to the assumption that selectivity is shared among the regions. Compared to the 2016 assessment, the model no longer estimates large numbers of fish at very small lengths for any of the fisheries. This is likely a result of the change in the growth curve used in the diagnostic model. Similarly, the predicted length composition from Indonesia (Z-ID-5) and Vietnam (Z-VN-5) in this assessment no longer overestimate the largest length frequencies due to these fisheries now being assigned their own selectivity functions.

This agreement between the observed and model-predicted numbers at length is reflected in the comparison of temporal variation in observed and model-predicted median fish lengths (Figures 26 and 27). Some lack of fit to the pole-and-line fisheries is shown for Regions 2 and 3 (P-ALL-2 and P-ALL-3), where the model predicts a fairly constant median length, but the time-series of observed lengths shows a gradual increase particularly since the mid-2000s. Similarly, the observed length composition from P-ALL-4 shows a marked increase in the last few years, which could potentially be accounted for by using selectivity time blocks within the model. The tendency to overestimate the length of the median fish for the purse seine fisheries in Region 6 (SU-ALL-6) is evident in Figure 27.

7.1.4 Tagging data

On the aggregated scale, the model appears to fit the tagging data very well (Figure 28), but on the tag program scale there are some differences in the quality of fit between them (Figure 29). The model generally predicts the number of tags recaptured by the JPTP to be larger than the observed values, whereas the other programs are fit very well. The difference in the fit is partially due to there being many fewer tag returns in the JPTP compared to the other programs and thus less weight within the model.

The numbers of observed tags returned by year of recapture compared to the model predicted tag returns show relatively good agreement between the two, though some of the years with high recaptures are underestimated by the model (Figure 30). Predicted tag returns by grouped fisheries (i.e., associated and unassociated purse seine grouped) fit the observed data well (Figures 31 and 32). The predicted proportion of tags returned by region of release and recapture in each quarter matched the observed tag recovery data well (Figure 33).

7.2 Model parameter estimates (diagnostic model)

7.2.1 Catchability

Purse seine fisheries where catchability was permitted to vary over time generally showed an increasing trend, but purse seine fisheries in the temperate and subtropical region predicted an increase followed by a decrease (Figure 34). The increase in catchability in the equatorial purse seine fisheries depended on the region and are between a two to six fold increase over the time series. Catchability for the purse seine fishery in Region 1 (S-ALL-1) shows an increasing trend until the mid-1980s and then a decreasing trend to very low levels at the end of the time series. The purse seine fishery in Region 2 shows an increase in catchability until about 2010 and then a sudden decrease in catchability. Catchability for the purse seine fishery in Region 3 was estimated to vary largely until the late 1980s and then gradually declined. The pole-and-line fishery in Region 6 was stable for the first decade of the time series and then declined until the mid-1990s. There was then a gradual increase with some sudden drops in catchability. This potentially reflects changes in the composition of the fleet (with different catchability of different flags) and the pronounced decline in effort for this fishery over time.

7.2.2 Selectivity

A range of selectivity patterns are shown by the different fisheries in the model and can be largely classified by fishing gear. The age-specific selectivity coefficients are displayed in Figure 35 and length-specific selectivity curves are displayed in Figure 36. Figure 36 displays the selectivity-at-age coefficients plotted against the mean length-at-age, where the dots represent the age. The pole-and-line and purse seine fisheries select mostly young fish that are three to eight quarters old, whereas the longline fisheries catch the largest sized fish. The Philippines miscellaneous fishery in Region 5 (Z-PH-5) selected for young and very old fish, which reflects the high variability in the length composition data.

7.2.3 Movement

To interpret the movement among regions for the diagnostic model, it is beneficial to examine the empirical movements of tagged fish (Figure 12) and the estimated seasonal movement parameters (Figure 37). In addition, Figure 38 portrays the proportion of total biomass in each region by origin (which region the fish recruited into) of fish based on the average recruitment within each region, the estimated movement rates and natural mortality. The left panel of Figure 38 assumes no fishing mortality is occurring while the right panel plot assumes the fishing mortality rate by age and season averaged over the time period 2008-2017. From these plots, it is evident that there is significant movement from Regions 1, 2, and 3 into Region 4 but only minimal movement out of this region.

Additionally there is substantial movement of other sources into Regions 7 and 8, but most other regions are primarily composed of fish originating from that region. There is movement of fish from Region 7 into Region 8 during all quarters of the year but there is relatively little movement out of Region 8.

The model appears to estimate movement consistent with the expected seasonal dynamics of skipjack in most regions. However, the estimated lack of movement out of Regions 4 and 8 is surprising. This may contribute in part to the large spawning potential estimated for these regions. The relatively few tag recaptures from releases in Region 4 may account for some of this lack of estimated movement. Additionally, the environmentally driven movements into and out of Region 8 cannot be captured by the constant movement assumed in the model.

This model did not assume age-dependent movement because the increased number of parameters did not result in a large enough improvement in the fit to the data to warrant the increased complexity. This differs from the 2016 skipjack tuna stock assessment that estimated age dependent movement.

7.2.4 Natural mortality

Natural mortality estimated in this stock assessment differs from the last assessment due to the use of the cubic spline function. For the diagnostic model, natural mortality was estimated to be high for the first few age classes and then decrease until age 6 quarters by about half, before moderately increasing with increasing age (Figure 39). The natural-mortality-at-age as estimated by all models in the structural uncertainty grids (both 8 region and 5 region) are presented in Figure 40, which are colored according to the assumed mixing rate of the tag data. It is apparent that by increasing the mixing period to 2 quarters much of the information in the tagging data related to natural mortality is lost, particularly for the oldest ages.

7.2.5 Maturity-at-Age

Maturity-at-age as calculated by the 2019 skipjack tuna assessment model differs from the values assumed in the 2016 assessment (Figure 41). The previous assessment assumed that 100% of fish were mature by age 3 quarters, where this assessment estimates that only about 50% of this fish are mature at this age-class. Greater than 90% of the population becomes mature after about 6 quarters of age. Maturity-at-age in the structural uncertainty grid varied according to the growth function assumed for the specific model. The GrowthHigh and diagnostic growth had nearly identical estimates of maturity-at-age, whereas the low growth resulted in a maturity-at-age that was marginally less at each age (Figure 41). The maturity-at-age from the one-off sensitivity model was estimated to be between the diagnostic growth and the estimates from the GrowthLow model at each age class.

7.2.6 Tag Reporting Rates

The estimated tag reporting rates by fishery recapture group (see groupings in Table 2) are displayed in Figure 42. The most important groups for scaling the population size are the purse seine fisheries in Regions 5–8 and account for the overwhelming majority of recaptures from the SPC tagging programs. Conversely, the pole-and-line/purse seine fisheries in Regions 1, 2, 3, and 4 recapture most of the JPTP tag releases. The reporting rates for the equatorial purse seine fisheries, with the exception of the Indonesian-Philippines purse seine fishery in Region 5 for the PTTP and the distant water purse seine in Region 5 for the JPTP, are below the upper bound set at 0.9. The estimate of reporting rate for the grouped pole-and-line, longline, and Japanese purse seine for the SSAP and the JPTP reside at the upper bound. Conversely, a large number of reporting rate groups for Region 5 fisheries are estimated at or close to the lower bound for the JPTP and SSAP. This is a consequence of no usable tag recaptures for these fisheries being available for these programs.

7.2.7 Growth

The growth function in the diagnostic model was fixed at values estimated from a stock assessment model where the length scalar was equal to 20. This estimated growth curve started from a mean length of 23 cm at age-class one. Mean length-at-age increased quickly until about age-class 8 quarters, after which growth slowed down until reaching a length of 84 cm in the oldest fish in the model (age-class 16). The standard deviation of length-at-age was constant across the ages (Figure 14).

As previously noted, two additional growth curves were estimated from two assessment models where the length scalar was specified at 1 or 40. The resulting growth curves had moderately lower and higher mean lengths in the oldest age groups of 78 and 85 respectively. The standard deviation of length-at-age was constant across ages and was similar to the model used as the diagnostic model.

7.3 Stock assessment results

7.3.1 Recruitment

The estimated distribution of recruitment across regions must be interpreted with caution to a degree, as MULTIFAN-CL has the ability to use a combination of movement and regional recruitment to distribute the population in a way that maximizes the total objective function. The estimated recruitment over the time series shows the general trend of being relatively stable in the first decade of the assessment period and then gradually increasing throughout the rest of the time series (Figures 43 to 45). The correlation between annual recruitment with catch and juvenile F are significant ($p < 0.05$; Figure 46). The majority of recruitment occurs in Regions 5, 7, and 8, with only minor amounts of recruitment occurring in the other regions. Recruitment in Region 8 increased

particularly in 2015, which is likely driven by the large catch in this region that year(Figure 45). Recruitment in all regions is seasonal, particularly in Regions 1–4 ((Figure 43).

The estimated SRR for the diagnostic model is presented in Figure 47, which shows the general increase in recruitment and decrease in spawning potential over time.

7.3.2 Biomass

With the exception of 2012–2015, the trend in total biomass estimated in the 2016 skipjack tuna assessment are fairly consistent with the trends estimated by the diagnostic model (Figures 45, 48 and 49). The 2016 assessment model predicted an increase in total biomass in the last few years of the assessed time series, but the estimates from the 2019 assessment show only a moderate increase in total biomass. Spawning potential is estimated to be relatively stable for the majority of the time series, but nearly a one-third decline in spawning potential occurred in 2009 due primarily to decreases in Regions 4 and 8 (Figure 48). A decrease in the total biomass in recent years was predicted for Region 5.

The diagnostic model estimates total biomass to be larger than estimates from the 2016 skipjack tuna stock assessment, but estimates spawning potential to be smaller (Figure 45). The increase in total biomass can largely be attributed to the transition to the 8 region model where large biomass is predicted in Region 4, which is larger than the estimated total biomass in Region 1 of the 5 region 2016 stock assessment model. The decrease in spawning potential is attributed to the different definition of the adult population, i.e. the change in the maturity schedule.

7.3.3 Fishing mortality

Average fishing mortality rates for juvenile and adult age-classes increase significantly throughout the time series (Figure 50). A large increase in fishing mortality in 2008 and 2009 was estimated by the model, and fishing mortality increases throughout the entire assessment period, unlike the estimate from the 2016 stock assessment.

The temporal trends in fishing-mortality-at-age vary by region, but the overall fishing-mortality-at-age increases over time (Figure 51). Fishing mortality in Regions 1 and 3 is estimated to be very high seasonally during the first decade of the assessment, but then decreases and remains at the same level until the end of the time series. The fishing mortality in Regions 1–4 are highly seasonal which is consistent with the CPUE for these regions (Figure 18). The fishing mortality for the equatorial regions generally increases throughout the time series, particularly in the last two decades of the assessment. Fishing mortality on the youngest age-class is only significant in Region 5 where the miscellaneous fisheries operate.

Changes in fishing mortality-at-age and the population age structure are shown for decadal time

intervals in Figure 52. Since the 1980s, the increase of fishing mortality to the current levels is due to the increase of catches of both juvenile and adult fish by both associated purse seine sets and the mixed gear fisheries in Indonesia, Philippines, and Vietnam. Fishing mortality on ages 3–6 quarters also increased through time consistent with the increased fishing mortality from the purse seine fishery (Figure 51).

7.4 Multi-model inference - stepwise model development, sensitivity analyses and structural uncertainty

7.4.1 One-off changes from the structural uncertainty analysis

Comparisons of the spawning potential and depletion trajectories for the diagnostic model and one-off sensitivity models are provided in Figures 53 and 54. The key reference points are compared for the structural uncertainty grid in Tables 6 and 7 and for the one-off sensitivity models in the Appendix Table A2. In conjunction with discussions from the PAW, the results from these one-off sensitivities informed the selection of the grid axes and levels therein. The one-off sensitivity models that form the basis of the axes of uncertainty will be discussed later in the context of the structural uncertainty grid (Section 7.4.2), however the results for the maturity one-off sensitivity are presented below.

Maturity ogive weighted by regional total biomass [*WtdMaturity*]

The model with the maturity ogive derived by weighting the estimated proportion in each region was nearly identical to the diagnostic model in terms of depletion estimation. The estimate of spawning potential for this model was minutely lower than the diagnostic model (Figure 54). Reference point estimates for this model were identical for most depletion-based reference points and only slightly different for MSY-based reference points.

7.4.2 Structural uncertainty analysis

Results of the structural uncertainty analysis are summarized in several forms: box and violin plots of $F_{recent}/F_{\text{MSY}}$ and $SB_{recent}/SB_{F=0}$ for the different levels of each of the four axes of uncertainty for the two grids separately (Figures 55 to 58), Majuro plots showing the estimates of $F_{recent}/F_{\text{MSY}}$ and $SB_{recent}/SB_{F=0}$ (and $SB_{latest}/SB_{F=0}$ for comparison) across all models in the two grids (Figures 59 to 61), plots of the distribution of dynamic depletion in each region for the 8 region and 5 region grid (Figures 62 and 63), and averages and quantiles across the 54 models in each grid for all of the reference points and other quantities of interest (Tables 6 and 7).

Many of the results of the structural uncertainty analysis are consistent with results of previous assessments of tuna stocks in the WCPO that used the same uncertainty axes, namely steepness and tag mixing period. The growth axis of uncertainty was not previously tested for skipjack but results are consistent with other tuna species in the WCPO where such axes have been investigated. The general features of the two uncertainty analyses are as follows:

- Each grid contains 54 models that display a moderate range of estimates of stock status and reference points, which suggest that the stock is more depleted than the estimate from the 2016 assessment.
- The most influential axis was the tag mixing period. The models with a mixing period of 1 quarter estimated $F_{recent}/F_{\text{MSY}}$ well above the estimate for a mixing period of 2 quarters (Figures 55 and 56).
- The second most influential axis of uncertainty in the grid was the SRR steepness. The $F_{recent}/F_{\text{MSY}}$ was most influenced by the change in steepness parameter, but there was also some differences in the $SB_{recent}/SB_{F=0}$ reference point for this axis.
- Growth was the next most influential axis in the grid. The difference between the diagnostic growth and high growth was limited but the low growth models showed a lower $F_{recent}/F_{\text{MSY}}$ and higher $SB_{recent}/SB_{F=0}$.
- The length composition scalar had minimal impact on the structural uncertainty grid and the range of reference point estimates nearly overlaps entirely among the three levels.
- Both of the structural uncertainty grids estimate $SB_{recent}/SB_{F=0}$ to be below the target reference point, though the 8 region model grid is marginally more optimistic compared to the 5 region model grid.
- Overfishing is not estimated to be occurring for any of the models in either of the structural uncertainty grids, nor is the stock estimated to be in an overfished state.
- Local spawning biomass depletion in the region around PNG and the Solomon Islands approached 0.2 in 2010, but has since increased to near 0.5 for both the 8 region and 5 region models (Region 6 and Region 5, respectively).
- Local depletion is most pronounced in the region covering Southeast Asian areas, being close to 0.2 in the most recent period. (Region 5 and Region 4, respectively).
- The least depleted but seasonally variable regions for the 8 region and 5 region models were the northern regions of the model (Regions 1–4 and Region 1, respectively).

7.4.3 Analyses of stock status

There are several ancillary analyses related to stock status that are typically undertaken on the diagnostic model (dynamic Majuro analyses, fisheries impacts analyses etc.). The shift toward relying more on multi-model inference in recent years 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 and 7) and the dynamic depletion plots of the structural uncertainty grid time series plots of fisheries depletion for all models in the grid (Figures 64 and 65). We also present the fished and unfished spawning biomass trajectories for the diagnostic model (Figure 66) and the dynamic Majuro and Kobe plots for the 8 region diagnostic model and 5 region equivalent.

Fishery impact

It is 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 biomass (Figure 67). The early impacts on the population were primarily driven by pole-and-line fishing, but the impact of that gear has generally declined to be relatively insignificant in Regions 6, 7, and 8. The impact of the pole-and-line fishery on Regions 1, 2, 3, and 4 has remained relatively constant over the entire time series. Equatorial purse seine fishing is estimated to have had the most significant impact on spawning biomass withing Regions 6, 7, and 8. The impact of the associated purse seine fishery has decreased in recent years in Regions 6 and 7 but has increased in Region 8. The miscellaneous fisheries were estimated to have a significant impact on spawning biomass in Region 5.

Yield analysis

The yield analyses conducted in this assessment incorporates the spawner recruitment relationship (Figure 47) into the equilibrium biomass and yield computations. Importantly, in the diagnostic 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.

The ratio of SB_{MSY} to SB_0 was estimated to be between 0.121–0.219 for the 8 region models (Table 6) and between 0.123–0.218 for the 5 region models (Table 7; median = 0.175 and 0.176, respectively).

A plot of the yield distribution under different values of fishing effort relative to the current effort are shown in Figure 68 for models representing different axes of the structural uncertainty grid. For the diagnostic model, it is estimated that MSY would be achieved by approximately doubling fishing

mortality, although the resulting increase in yield would be relatively small. The right-hand arm of the yield curve displays a gradual decline in yield with increasing fishing mortality. The different example models shown display a similar pattern over the scale of fishing mortality although the absolute value of the yield curve differs significantly. The model with a mixing period of 2 quarters estimates that MSY would occur at 2.32 times the current fishing mortality and would result in a 17% increase in yield.

The yield analysis also enables the calculation of the MSY level that could be achieved assuming the different patterns of age-specific fishing mortality observed through the history of the fishery (Figure 69). The MSY level for the diagnostic model was estimated to be stable for the entire time series of the assessment.

Dynamic Majuro and Kobe plots and comparisons with Limit and Target Reference Points

The section summarizing the structural uncertainty grid (Section 7.4.2) 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 and Kobe plots, the methods of which are presented in Section 5.7.4. The large number of model runs in the structural uncertainty grid precludes presenting this process for all runs, however the dynamic Majuro and Dynamic Kobe plots for the diagnostic and *5Region* models are presented in Figure 70.

At the beginning of the assessment period the stock was almost at unexploited levels. Both the dynamic Majuro and Kobe plots show the steady increase in depletion of the stock. The dynamic Majuro plot indicates that the stock went below the interim target reference point in the late 2000s and is currently still below that level ($50\%SB_{F=0}$). The models do not estimate the stock to be undergoing overfishing and the stock is not overfished.

8 Discussion and conclusions

The 2019 skipjack tuna stock assessment generally estimated the stock status to be below the interim skipjack target reference point ($50\%SB_{F=0}$) but above the limit reference point for all models.

8.1 Changes to the previous assessment

The addition of three more years of data (tagging, catch, effort, length compositions) and numerous other model changes resulted in an estimated stock status more depleted compared to the 2016 stock assessment. The previously estimated large increase in recruitment during the time period 2012 to 2015 was not present in this assessment. Additionally, this assessment estimated fishing

mortality to have increased since 2012, which is the opposite of the trend estimated from the 2016 assessment. This difference is likely due to the former assessment estimating a decrease in purse seine catchability for that time period. The more depleted stock status can also be partly attributed to the new maturity function and estimation of tag overdispersion in the assessment. Numerous changes to the previous assessment were made with differing effects on the estimated stock status.

Maturity-at-age

Maturity-at-length data became available since the last assessment and were incorporated into the assessment using recent developments in MULTIFAN-CL. The maturity-at-length data were incorporated into the assessment model and were converted into a maturity-at-age through estimated growth curves. This change in maturity has minimal effect on model parameter estimates, but changes the definition of the adult population. The SRR will differ as a result in changes to the spawning potential, which will impact the unfished biomass calculations and have a large impact on management reference points (both MSY and depletion-based).

Growth

The estimation of growth within the skipjack stock assessment has been problematic for the past few assessments (Rice et al., 2014; McKechnie et al., 2016). However, estimates of growth conducted using otoliths do not correspond well with the observed length frequency and result in unreasonable model estimates (Tanabe et al., 2003; McKechnie et al., 2016). The growth curves for the skipjack tuna assessment since 2011 estimated the mean length of the youngest age in the model to reside on the lower bound of 10 cm. This was avoided in this assessment by starting this parameter at a value more consistent with the smallest sizes observed in the catch length frequency. However, estimation of this and other growth parameters became problematic when the length composition scalar was increased (i.e., less influence of the length data). The estimated growth curve was found to be dependent on the value of the length scalar and thus were incorporated into the structural uncertainty grid. The growth curve is an important component of any length based stock assessment and can have a large impact on the assessed stock status.

Length Scalar

An additional change from the 2016 stock assessment was the decision to decrease the influence of the length composition data. This decision resulted from the observed conflict between the size and tagging data in the total biomass and overdispersion likelihood profiles (Figures A1 and A3). Therefore, it was believed that there was model mis-specification regarding selectivity due to the

model assuming no time-series or random variability. Initial experiments to incorporate these time-varying selectivities weakened the impact of the size data on the biomass scaling similar to simply reducing the influence of the length data (increasing the scalar). The approach of using time-varying selectivities allowed for a better fit to the length composition data but resulted in many more parameters to be estimated compared to the simple down-weighting of the length data. The size sampling design in the equatorial regions is dominated by length measurement of fish caught in large schools in the purse seine fisheries, whereas the northern regions are dominated by measurement from the pole-and-line fisheries (Figure 9). Given this sampling design it was believed that samples may be more corrected that purely random samples of the population. Therefore, we opted to more strongly down-weight the size data, noting the need to further investigate alternative selectivity specifications to address model mis-specification, and consider self-scaling methods that can possibly estimate effective sample size (see Section 8.2.1 for additional details).

Tag overdispersion

The estimation of the overdispersion of the tagging data was included in the structural uncertainty analysis of the 2016 stock assessment (McKechnie et al., 2016); the current assessment chose to use the approach of estimating this overdispersion parameter for all models in the grid. The reference case model from the previous assessment used a Poisson distribution, which assumes the variance is equal to the mean. This is an unlikely assumption due to the schooling behavior of skipjack tuna and the variability in the observed data. The variability in tag returns is better accounted for by the estimation of the overdispersion parameter (τ) of the tagging data. Additionally, the estimation of the overdispersion parameter was remarkably stable in all the models tested, as is shown by the rapid change in the likelihood profile for the model when fixed at different values (Figure A3). Finally, the estimate of the overdispersion parameter was influenced not only by the tagging data but the other data components because the parameter is based on the model's ability to fit the observed tagging data. Therefore, we allowed the overdispersion parameter to respond to all of the different combinations of model settings used in the grid by estimating it for each model.

Regional structure

The largest change from the 2016 skipjack tuna stock assessment was the transition to the 8 region model. In both the 8 region and 5 region models, the median estimate of $SB_{recent}/SB_{F=0}$ was below the skipjack tuna target reference point ($50\%SB_{F=0}$). The 5 region model used in the 2016 skipjack tuna assessment resulted in an 80% uncertainty interval for $SB_{recent}/SB_{F=0}$ of 0.317 – 0.504, whereas the 8 region model was more optimistically estimated at 0.363 – 0.521. This difference in estimated stock status was primarily due to an higher biomass estimated by the 8 region model that can be attributed to Region 4 of that spatial structure (Figures 1a and 71). The diagnostic

model estimated that a large proportion of fish from Regions 1 –3 move to Region 4 (Figure 37) and that the exploitation rate in the latter area is low. Thus the model predicts that this area has a very large adult population. The estimate of biomass in this region is driven strongly by tagging data from the pole-and-line and longline fisheries in this region (Figures A4 and A5), whereas the CPUE in this region did not have as large of an impact in the likelihood profiles for these fisheries. However, the estimated biomass in the region relative to the other regions is consistent with the relative biomass estimates for Regions 1-4, 7, and 8 from the geostatistical CPUE model and SEAPODYM (Figure 72; Kinoshita et al., 2019; Senina et al., 2016).

Tag release lengths

This assessment included tags from the JPTP that were missing length at release, which were excluded from the 2016 skipjack tuna stock assessment. The simple method of sampling from observed recent measurements relies on the assumption that the length at release has been consistent through time. Yet, preliminary tests comparing the different methods for filling in the lengths at release were similar. The inclusion of these data provides valuable information on the scale of the population within regions and the movement among them, which is reflected in the changes to the estimated depletion and biomass.

Improvements with minor impact

Many of the improvements made to the assessment had relatively little impact on the estimated stock status overall:

- removal of spurious historical catch data for the pole-and-line fishery in Region 4 of the 5 region model and the purse seine in Region 1 of the 5 region model,
- addition of missing catch data to the Indonesian miscellaneous gear fishery in Region 4 of the 5 region model (see Vincent (2019) for more details),
- replacement of Japanese pole-and-line length composition data with the most up to date data,
- incorporation of an updated length-weight relationship,
- changes in the number of nodes of the cubic splines for selectivity of various fisheries,
- use of logistic selectivity curves for the longline fisheries,
- estimation of a cubic spline with 5 nodes for natural mortality,
- update of reporting rate priors, and
- removal of tag release events with less than 100 tags released.

8.2 Sources of uncertainty

8.2.1 Data weighting and conflict among data components

The existence of data conflicts in integrated stock assessment models is commonplace. While further investigation of this data conflict is warranted, it is clear from the likelihood profile and the structural uncertainty grid that the tagging data are an important source of information for the stock assessment. This conclusion was also made in the 2016 stock assessment (McKechnie et al., 2016). The likelihood profile for both the 5 region and 8 region model suggest that the tagging data would fit the data best with a total biomass of around 5 million metric tons (Figures A1 and A2). This suggests that the data have a consistent influence on the model despite the difference in regional structure. Conversely, the length and CPUE data suggest a lower biomass for the 5 region model compared to the 8 region model. This difference in the best fit by the CPUE data is surprising given that both the 5 region and 8 region models are calculated from the same geostatistical model with only a difference in the regional grouping of clusters. CPUE data are not fit in the objective function of MULTIFAN-CL like other data sources with an observed and model predicted component. Instead the data are fit based on the estimated effort deviations from the effort calculated from the CPUE and the time varying penalties assigned to each observation. Therefore, the difference in the fit to the CPUE between the two regional structures is likely due to differences in the penalties and requires additional investigation.

The length composition scalars had less influence in the structural uncertainty grid than previous years. This was a consequence of the larger scalar value of the length composition data than previous assessments. The decision to reduce the influence of the length composition in the assessment model arose due to the conflict between this data source and the tagging data. The preferred method for dealing with conflict from size composition data is to allow flexibility in the selectivity functions. The selectivity splines used for most fisheries in this assessment are very flexible and can take a wide variety of shapes. A preliminary investigation was conducted on a model with selectivity blocked into decades for all fisheries except the longline. A likelihood profile for this blocked selectivity model was similar to that of the diagnostic model (not shown). Attempts to use the Dirichlet multinomial distribution and the self scaling multinomial were unsuccessful, but developments to the latter method are ongoing and further investigation of methods to resolve data conflict are recommended.

One of the difficulties in fitting integrated models is the need to assign relative weightings to the different data components. It is increasingly recognized that data components should be actively down-weighted in some situations, e.g. down-weighting length compositions to allow CPUE data to provide the most information on trends in abundance (Francis, 2011). How to weight the tagging data is less clear and accepted recommendations are currently lacking. For skipjack tuna the tagging data provides most of the information on scaling biomass and parameters such as natural mortality and movement. The approach of estimating the overdispersion of the tagging data within the

assessment model is believed to be the best approach currently available.

8.2.2 Robustness of CPUE indices

As noted in previous assessments, the JP pole-and-line fisheries provide the standardized CPUE indices in all regions except Regions 5 and 6, but this fishery now makes up less than 10% of the total WCPO skipjack catch and the area covered by the fishery in 2018 is very limited (Kinoshita et al., 2019). The index produced for the 8 region model was possible primarily as a result of the geostatistical method, as the traditional method had difficulty estimating a time series trend for a few regions. If these fleets continue to reduce, and contract their effort to restricted areas of the assessment regions, there will be further problems with their utility in indexing relative abundance. It is possible, and increasing likely, that this data source will not support standardization analyses if data becomes too sparse. Investigation of alternative sources for a valid index of the relative size of the stock over time is imperative.

8.3 Recommendations for further work

8.3.1 WCPFC-specific recommendations

Given the currently agreed upon skipjack tuna interim target reference point (TRP; $50\%SB_{F=0}$), the management recommendations from this stock assessment will differ from the previous stock assessments. Yet the recommendations for future work will remain relatively unchanged.

The tagging data is a critical component of the skipjack stock assessment and it is unlikely that these models can provide robust estimates for recent time-periods without it. It is therefore recommended that regular large-scale tagging cruises and complementary tag recovery work continue to be undertaken in a way that provides the best possible data for stock assessment purposes. In particular, the tagging of skipjack tuna in the eastern equatorial region and between 10°N and 30°N is needed to more accurately estimate the population size in this region. The difficulties in obtaining tag returns from the 2017 tag release highlights the importance of continued cooperation and support from all CCMs to facilitate the timely return of recaptured tags to the tagging agency (SPC or Japan) in order to ensure the accuracy of stock assessments in the region.

Another important component of the skipjack tuna stock assessment is the standardized CPUE from the Japanese pole-and-line fishery. A stock assessment model cannot operate properly without an index of abundance to inform the assessment model of trends in biomass. Therefore, it is imperative that alternative sources for creation of a standardized index are explored in order for stock assessments to be conducted to provide management advice.

The finalized data from Indonesia and Vietnam, that came from their annual catch estimate meetings, were not available until June 28th 2019. This gave very limited time to conduct the assessment with

the full dataset. We recommend that the annual catch estimate meetings for these countries be held earlier in the year, particularly when skipjack tuna is assessed, due to the inclusion of the most recent year within the assessment model. Ideally these would occur in time for the data provision deadline of April 30^{th}, but finalized data should be available no later than June 1^{st}.

8.3.2 Biological studies and data investigations

Estimation of growth functions

Difficulty in estimating growth within the skipjack tuna stock assessment has been noted in previous assessment for the species. Growth parameters have been shown to have a significant influence on stock status and management quantities in this and other tuna assessments (Harley et al., 2014; McKechnie et al., 2015, 2016). For skipjack, the main opportunity for additional data on growth is from tagging-based length-increment data. Incorporation of these data source into the MULTIFAN-CL model is currently a high priority and has been so for a number of years.

Standardized CPUE analyses

Investigation of standardizing the purse-seine fishery data to create an alternative CPUE index is currently underway. However, there are many potential issues that the dynamics of the purse seine fishery and biology of the tuna that must be overcome for this analysis to provide a reliable index of abundance. As the name implies, a CPUE index is based on the catch divided by the effort, but the assumption within a stock assessment model is that the catch per effort is linearly proportional to biomass or number of fish. However, the catch within a year in the purse seine fishery in nearly directly proportional to the number of sets that year, which indicates that this assumption is likely violated. A potential avenue to investigate in order to obtain a standardized index from the purse seine fishery is to measure catch relative to the time spent searching. However, the form of searching for schools of tuna has changed over time as technologies have improved, e.g., helicopters, FADs, and most recently drones. In order to standardize in this manner, information through the historical period of the fishery would need to be provided from the fishing industry to the scientist conducting the analysis. Therefore, we strongly advocate for the cooperation between the purse seine fishing industry and SPC scientists. Given the uncertainty in the ability to continue using the pole-and-line fishery as an index of abundance in future assessments, the investigation of these alternative methods is a worthwhile and needed endeavor.

An alternative to the fishery dependent CPUE commonly used in assessments in the WCPO is fishery independent indices of abundance. The application of a fishery independent survey for highly migratory species such as tuna are often immediately disregarded due to the cost involved to conduct a survey over such a large area. Despite this cost the application of some fishery independent

surveys, such as aerial surveys (Bauer et al., 2015) and acoustic surveys (Uranga et al., 2017; Melvin et al., 2018), have been used for Atlantic bluefin tuna (*Thunnus thynnus*). If designed appropriately, a fishery independent survey could be used to inform the abundance of all tuna and commercially important species in the WCPO. To appropriately conduct such a massive survey would require significant amounts of funding and cooperation among all CCMs. Given the importance of the skipjack fishery and the potential for application to other species, a thorough investigation of the feasibility of such a survey is warranted.

Tagging data examination

Additional investigation regarding the tagging data are recommended given current uncertainties and the importance of this data source in the assessment. The largest uncertainty in the grid was the mixing period assumed for the tagging data. Investigation of individual-based models can provide some insight into the appropriate period of mixing that should be assumed (Scutt Phillips et al., 2018). These investigations are ongoing at SPC, but initial results suggest that the mixing period may be specific to the spatial and environmental conditions at release. Continued investigation of these individual based models and their application to the assumed mixing period for individual tag release events is a proposed area for future analysis.

Due to the quarterly dynamics assumed in MULTIFAN-CL, there is the potential that the mixing period is unevenly applied across tag releases within a quarter. Under the current methods used in the assessment, it is possible that tags could be released a day before the end of the quarter and recaptured a few days later in the next. Assuming a one quarter mixing period, these tag returns (assuming reported with usable recapture information) would influence the model fit. Alternatively, a tag released on the first day of the quarter and recaptured 85 days later (i.e., in the same quarter) would be excluded under the current methods. Therefore, the effective mixing period applied to the tag release events is variable depending on the date of release. Further investigation regarding this issue is needed, such as externally removing tag recaptures before a certain period at liberty. However, if the filtering of the tagging data were to be conducted external to MULTIFAN-CL, then a correction for the reporting rate is required, which would likely entail making numerous assumptions. Therefore, careful investigation of this issue is needed.

There remain some questions regarding the influence of the length composition in the estimation of this parameter in the model. An alternative formulation of the likelihood function for the tagging data is currently being created for use in the assessment model. A thorough investigation of this proposed likelihood function would be an area of research that would be beneficial for all stock assessments in the WCPO.

This stock assessment included additional tag releases from the Japanese Tagging Program by randomly sampling from the measured release lengths. Future improvement in the methods to

include this historical data could be conducted by ensuring that the length at release is less than the length at recapture, when available. If an estimate of growth is available, the back calculation of length given the time at liberty could be conducted.

Time varying movement

Movement of skipjack between the western and central Pacific Ocean are believed to by driven by environmental events such as El Niño. Currently, MULTIFAN-CL does not have capabilities to estimate or assign time-varying movement. Development of this feature within MULTIFAN-CL could be beneficial to the assessment and management strategy evaluations of skipjack.

8.4 Main assessment conclusions

The general conclusions of this assessment are as follows:

- Total biomass and spawning potential remained relatively stable, with fluctuations, until the mid 2000s, after which it declined. Estimated recruitment shows an increasing trend from 1980 to the recent period.
- Average fishing mortality rates for juvenile and adult age-classes increase throughout the period of the assessment.
- The 8 region model structure provided slightly more optimistic estimates of stock status when compared to the 5 region model structure. In both cases, the stock was assessed to be above the adopted LRP, and fished at rates below F_{MSY}, with 100% probability. We conclude the skipjack stock is not overfished, nor subject to overfishing.
- Overall median depletion over the recent period (2015-2018; $SB_{recent}/SB_{F=0}$) was 0.44 (80 percentile range 0.36-0.52) for the 8 region model, and 0.40 (80 percentile range 0.30-0.50) for the 5 region model.
- Results from both regional structures indicate a stock status currently on average below the interim TRP for skipjack.
  - For the 8 region grid, 47 of the 54 models (85%) estimated $SB_{recent}/SB_{F=0}$ to be less than the TRP ($50\%SB_{F=0}$).
  - For the 5 region grid, 48 of the 54 models (87%) estimated $SB_{recent}/SB_{F=0}$ to be less than the TRP ($50\%SB_{F=0}$).
- Recent median fishing mortality (2014-2017; $F_{recent}/F_{\text{MSY}}$) was 0.44 (80 percentile range 0.34-0.61) for the 8 region model, and 0.48 (80 percentile range 0.35-0.66) for the 5 region model.

We therefore conclude that skipjack tuna is not currently overfished, nor undergoing overfishing.

Acknowledgments

We thank the various fisheries agencies, in particular NRIFSF, and especially Hidetada Kiyofuji, Yoshi Aoki, and Junji Kinoshita for the provision of the catch, effort, tagging, and length composition data used in this analysis. Special thanks must go to Nicholas Ducharme-Barth, Tiffany Cunningham, and Keith Bigelow who provided standardized CPUE indices for the analyses. We also thank participants at the preparatory stock assessment workshop (Noumea, April 2016; Pilling and Brouwer, 2019) for their contributions to the assessment. We would like to thank the SPC carrot-munching sheep moppet for the comic relief, making of plots, and the fly outbreak that provided aggression relief.

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.

Anon (2011). Report of the 2011 ISSF stock assessment workshop. Issf-technical-report-2011-02, Rome, Italy.

Bauer, R., Sylvain, B., Brisset, B., and Fromentin, J.-M. (2015). Aerial surveys to monitor bluefin tuna abundance and track efficiency of management measures. *Marine Ecology Progress Series*, 534:221–234.

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.

Bigelow, K., Garvilles, E., Bayate, D., and Cecilio, A. (2019). Relative abundance of skipjack for the purse seine fishery operating in the Philippines Moro Gulf (region 12) and high seas pocket # 1. Technical Report WCPFC-SC15-2019/SA-IP-08, Pohnpei, Federated States of Micronesia.

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

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

Davies, N., Fournier, D., Takeuchi, Y., Bouyè, F., and Hampton, J. (2018). Developments in the MULTIFAN-CL software 2018-2019. Technical Report WCPFC-SC-2019/SA-IP-02, Pohnpei, Federated States of Micronesia.

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.).

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

Gaertner, D., gado de Molina, A., Ariz, J., Pianet, R., and Hallier, J. P. (2008). Variability of the growth parameters of the skipjack tuna (*katsuwonus pelamis*) among areas in the eastern atlantic: analysis from tagging data within a meta-analysis approach. *Aquatic Living Resources*, 21:349–356.

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. and Williams, P. (2016). Annual estimates of purse seine catches by species based on alternative data sources. WCPFC-SC12-2016/ST-IP-03, Bali, Indonesia, 3–11 August 2016.

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.

Hoyle, S., Kleiber, P., Davies, N., Harley, S., and Hampton, J. (2011). Stock assessment of skipjack tuna in the western and central pacific ocean. WCPFC-SC7-2011/SA-WP-04, 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. Technical Report 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.

Kinoshita, J., Aoki, Y., Ducharme-Barth, N., and Kiyofuji, H. (2019). Standardized catch per unit effort (cpue) of skipjack tuna of the japanese pole-and-line fisheries in the wcpo from 1972 to 2018. Technical Report WCPFC-SC15-2019/SA-WP-10, Pohnpei, Federated States of Micronesia.

Kiyofuji, H., Aoki, Y., Kinoshita, J., Ohashi, S., and Fujioka, K. (2019a). A conceptual model of skipjack tuna in the western and central Pacific Ocean (WCPO) for the spatial structure configuration. Technical Report WCPFC-SC15-2019/SA-WP-11, Pohnpei, Federated States of Micronesia.

Kiyofuji, H., Ohashi, S., Kinoshita, J., and Aoki, Y. (2019b). Overview of historical skipjack length and weight data collected by the japanese pole-and-line fisheries both of commercial and research vessel (r/v) from 1953 to 2017. Technical Report WCPFC-SC15-2019/SA-IP-12, Pohnpei, Federated States of Micronesia.

Kleiber, P., Fournier, D., Hampton, J., Davies, N., Bouye, F., and Hoyle, S. (2019). *MULTIFAN-CL User's Guide*. http://www.multifan-cl.org/.

Lawson, T. (2011). Purse-seine length frequencies corrected for selectivity bias in grab samples collected by observers. Technical Report WCPFC-SC7-2011/ST-IP-02, Pohnpei, Federated States of Micronesia, 9–17 August 2011.

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.

Lehodey, P., Bertignac, M., Hampton, J., Lewis, A., and Picaut, J. (1997). El Nino Southern Oscillation and tuna in the western Pacific. *Nature*, 389:715–718.

Leroy, B. (2000). Preliminary results on skipjack (*Katsuwonus pelamis*) growth. SKJ-1, 13th Meeting of the Standing Committee on Tuna and Billfish. Noumea, New Caledonia.

Maunder (2001). Growth of skipjack tuna (katsuwonus pelamis) in the easter pacific ocean, as eestimate from tagging data. Technical Report Bulletin of the Inter-American Tropical Tuna Commision, vol. 22.

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

McKechnie, S., Hampton, J., Pilling, G. M., and Davies, N. (2016). 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., Pilling, G., and Hampton, J. (2017). 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.

Melvin, G. D., Munden, J., and Finley, M. (2018). Development of fishery independent index of abundance for Atlantic bluefin tuna on the Gulf of St Lawrence. *Cellect. Vol. Sci. Pap. ICCAT*, 74(6):2570–2585.

Ohashi, S., Aoki, Y., Tanaka, F., Aoki, A., and Kiyofuji, H. (2019). Reproductive traits of female skipjack tuna *katsuwonus pelamis* in the western central Pacific Ocean (WCPO). Technical Report WCPFC-SC15-2019/SA-WP-10, Pohnpei, Federated States of Micronesia.

Peatman, T. (2019). Analysis of tag seeding data and reporting rates. Technical Report WCPFC-SC15-2019/SA-IP-06, Pohnpei, Federated States of Micronesia.

Pilling, G., Scott, R., Davies, N., and Hampton, J. (2016). Approaches used to undertake management projections of WCPO tuna stocks based upon MULTIFAN-CL stock assessments. WCPFC-SC12-2016/MI-IP-04, Bali, Indonesia, 3–11 August 2016.

Pilling, G. M. and Brouwer, S. (2019). Report from the spc pre-assessment workshop, noumea, april 2019. Technical Report WCPFC-SC14-2018/SA-IP-01, Pohnpei, Federated States of Micronesia.

Rice, J., Harley, S., Davies, N., and Hampton, J. (2014). Stock assessment of skipjack tuna in the Western and Central Pacific Ocean. WCPFC-SC10-2014/SA-WP-05, Majuro, Republic of the Marshall Islands, 6–14 August 2014.

Scott, R., Davies, N., Pilling, G. M., and Hampton, J. (2016). Retrospective forecasting of the 2014 WCPO bigeye tuna stock assessment. WCPFC-SC12-2016/SA-WP-02, Bali, Indonesia, 3–11 August 2016.

Scutt Phillips, J., Sen Gupta, A., Senina, I., Sebille, E., Lange, M., Lehodey, P., J., H., and Nichol, S. (2018). An individual-based model of skipjack tuna Katsuwonus pelamis movement in the tropical pacific ocean. *Progress in Oceanography*, 164:63–74.

Senina, I., Lehodey, P., Calmettes, B., Nichol, S. J., Caillot, S., J., H., and P., W. (2016). Predicting skipjack tuna dynamics and effects of climate change using seapodym with fishing and tagging data. WCPFC-SC12-2016/EB WP01.

Sibert, J. R., Hampton, J., Fournier, D., and Bills, P. J. (1999). An advection-diffusion-reaction model for the estimation of fish movement parameters from tagging data, with application to skipjack tuna (*Katsuwonus pelamis*). *Canadian Journal of Fisheries and Aquatic Sciences*, 56:925–938.

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

Tanabe, T., Kayama, S., and Ogura, M. (2003). Precise age determination of young to adult skipjack tuna (*Katsuwonus pelamis*) with validation of otolith daily increment. SCTB16/SKJ-08, 16th Meeting of the Standing Committee on Tuna and Billfish, Mooloolaba, Queensland, Australia, 9–16 August.

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.

Uranga, J., Arrizabalaga, H., Boyra, G., Hernandez, M. C., ni, N. G., Arregui, I., Fernandes, J. A., Yurramendi, Y., and Santiago, J. (2017). Detecting the presence-absence of bluefin tuna by automated analysis of medium-range sonars on fishing vessels. *PLoS ONE*, 12(2):e0171382.

Vidal, T., Pilling, G., Tremblay-Boyer, L., and Usu, T. (2019). Standardized cpue for skipjack tuna katsuwonus pelamis from the papua new guinea archipelagic purse seine fishery. Technical Report WCPFC-SC15-2019/SA-IP-05, Pohnpei, Federated States of Micronesia.

Vincent, M. T. (2019). Summary of fisheries structures for the 2019 assessment of skipjack tuna in the western and central pacific ocean. Technical Report WCPFC-SC15-2019/SA-IP-09, Pohnpei, Federated States of Micronesia.

Vincent, M. T., Aoki, Y., Kiyofuji, H., Hampton, J., and Pilling, G. M. (2019). Background analyses for the 2019 stock assessment of skipjack tuna. Technical Report WCPFC-SC15-2019/SA-IP-04, Pohnpei, Federated States of Micronesia.

Wild, A. and Hampton, J. (1994). A review of the biology and fisheries for skipjack tuna, *Katsuwonus pelamis*, in the Pacific Ocean. FAO Fisheries Technical Paper (FAO) 336(2):1–51.

9 Tables

Table 1: Definition of fisheries for the MULTIFAN-CL skipjack analysis for the diagnostic case (8 region) model (see 2016 assessment for the 5 region). Gears: P = pole and line; S = purse seine unspecified set type; SA = purse seine associated set type; SU = purse seine unassociated set type; LL = longline; Dom = the range of artisanal gear types operating in the domestic fisheries of Philippines, Indonesia, and Vietnam. Flag/fleets: PH = Philippines; ID = Indonesia; VN = Vietnam; DW = distant water; ALL = all nationalities.

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

Table 2: Summary of the groupings of fisheries within the diagnostic (8 region) assessment model for estimation of selectivity, catchability , tag recaptures, and tag reporting rates. In the table Sel is selectivity, Nodes is the number of nodes in the selectivity spline, SeasCat is seasonal catchability, TimVarCat is time varying catchability, TimVarCatCV is the CV on the time varying catchability random walk, EffPen is the effort deviation penalties grouping, EffPenCV is the CV on the effort deviations, Recaptures are the grouping of fisheries for tag recaptures, and Reporting is the sharing of the reporting rate parameters among fisheries within a tagging program . Note that effort is missing for all L and 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	Sel Group	Sel Type	SeasCat	TimVarCat	TimVarCatCV	EffPen	EffPenCV	Recaptures	Reporting
F1 P-ALL-1	1	1	Nodes 4	Y	N	NA	time-variant	0.20	1	1
F2 S-ALL-1	1	2	Nodes 4	Y	Y	0.1	scaled	0.41	2	1
F3 L-ALL-1	1	3	Logistic	N	N	NA	constant	0.22	3	1
F4 P-ALL-2	2	1	Nodes 4	Y	N	NA	time-variant	0.20	4	1
F5 S-ALL-2	2	2	Nodes 4	Y	Y	0.1	scaled	0.41	5	1
F6 L-ALL-2	2	3	Logistic	N	N	NA	constant	0.22	6	1
F7 P-ALL-3	3	1	Nodes 4	Y	N	NA	time-variant	0.20	7	1
F8 S-ALL-3	3	2	Nodes 4	Y	Y	0.1	scaled	0.41	8	1
F9 L-ALL-3	3	3	Logistic	N	N	NA	constant	0.22	9	1
F10 Z-PH-5	5	4	Nodes 6	N	N	NA	constant	0.22	10	2
F11 Z-ID-5	5	5	Nodes 5	N	N	NA	constant	0.22	11	3
F12 S-ID.PH-5	5	6	Nodes 5	N	N	NA	time-variant	0.20	12	4
F13 P-ALL-5	5	7	Nodes 6	Y	Y	0.1	scaled	0.41	13	1
F14 SA-DW-5	5	8	Nodes 4	Y	Y	0.1	scaled	0.41	14	5
F15 SU-DW-5	5	9	Nodes 4	Y	Y	0.1	scaled	0.41	14	5
F16 Z-VN-5	5	10	Nodes 5	N	N	NA	constant	0.22	15	6
F17 L-ALL-5	5	11	Logistic	N	N	NA	constant	0.22	16	1
F18 P-ALL-6	6	12	Nodes 4	Y	Y	0.1	scaled	0.41	17	1
F19 SA-ALL-6	6	8	Nodes 4	Y	Y	0.1	time-variant	0.20	18	7
F20 SU-ALL-6	6	9	Nodes 4	Y	Y	0.1	scaled	0.41	18	7

Table2– Continued from previous page

Fishery	Region	Sel Group	Sel Type	SeasCat	TimVarCat	TimVarCatCV	EffPen	EffPenCV	Recaptures	Reporting
F21 L-ALL-6	6	11	Logistic	N	N	NA	constant	0.22	19	1
F22 P-ALL-4	4	1	Nodes 4	Y	N	NA	time-variant	0.20	20	1
F23 L-ALL-4	4	3	Logistic	N	N	NA	constant	0.22	21	1
F24 P-ALL-7	7	1	Nodes 4	Y	N	NA	time-variant	0.20	22	1
F25 SA-ALL-7	7	8	Nodes 4	Y	Y	0.1	scaled	0.41	23	8
F26 SU-ALL-7	7	9	Nodes 4	Y	Y	0.1	scaled	0.41	23	8
F27 L-ALL-7	7	11	Logistic	N	N	NA	constant	0.22	24	1
F28 P-ALL-8	8	1	Nodes 4	Y	N	NA	time-variant	0.20	25	1
F29 SA-ALL-8	8	8	Nodes 4	Y	Y	0.1	scaled	0.41	26	9
F30 SU-ALL-8	8	9	Nodes 4	Y	Y	0.1	scaled	0.41	26	9
F31 L-ALL-8	8	13	Logistic	N	N	NA	constant	0.22	27	1

Table 3: Summary of the tagging file used in the 2019 diagnostic case model, showing the raw number of usable releases, the corrected effective number of releases, the correction ratio, and the raw and effective recapture rates by tagging program

Programme	Raw	Effective	Recaptures	Correction	Raw.rate	Eff.rate
JPTP	127353	78321	5524	0.61	0.04	0.07
PTTP	271698	136158	35432	0.50	0.13	0.26
RTTP	90667	54808	10899	0.60	0.12	0.20
SSAP	78671	60525	4237	0.77	0.05	0.07
Total	568389	329812	56092	0.58	0.10	0.17

Table 4: Description of symbols used in the yield and stock status analyses. For the purpose of this assessment, "recent" for F is the average over the period 2014–2017 and for SB is the average over the period 2015–2018 and "latest" is 2018

Symbol	Description
C_{latest}	Catch in the last year of the assessment (2018)
F_{recent}	Average fishing mortality-at-age for a recent period (2014–2017)
$Y_{F_{recent}}$	Equilibrium yield at average fishing mortality for a recent period (2014–2017)
f_{mult}	Fishing mortality multiplier at maximum sustainable yield (MSY)
F_{MSY}	Fishing mortality-at-age producing the maximum sustainable yield (MSY)
MSY	Equilibrium yield at F_{MSY}
$F_{recent}/F_{\text{MSY}}$	Average fishing mortality-at-age for a recent period (2014–2017) relative to F_{MSY}
SB_{latest}	Spawning biomass in the latest time period (2018)
SB_{recent}	Spawning biomass for a recent period (2015–2018)
$SB_{F=0}$	Average spawning biomass predicted in the absence of fishing for the period 2008–2017
SB_{MSY}	Spawning biomass that will produce the maximum sustainable yield (MSY)
$SB_{MSY}/SB_{F=0}$	Spawning biomass that produces maximum sustainable yield (MSY) relative to the average spawning biomass predicted to occur in the absence of fishing for the period 2008–2017
$SB_{latest}/SB_{F=0}$	Spawning biomass in the latest time period (2018) relative to the average spawning biomass predicted to occur in the absence of fishing for the period 2008–2017
$SB_{latest}/SB_{\text{MSY}}$	Spawning biomass in the latest time period (2018) relative to that which will produce the maximum sustainable yield (MSY)
$SB_{recent}/SB_{F=0}$	Spawning biomass for a recent period (2015–2018) relative to the average spawning biomass predicted to occur in the absence of fishing for the period 2008–2017
$SB_{recent}/SB_{\text{MSY}}$	Spawning biomass for a recent period (2015–2018) relative to the spawning biomass that produces maximum sustainable yield (MSY)
$20\%SB_{F=0}$	WCPFC adopted limit reference point – 20% of spawning biomass in the absence of fishing average over years $t-10$ to $t-1$ (2008–2017)

Table 5: Description of the structural sensitivity grid used to characterize uncertainty in the assessment, where the options for the diagnostic case are indicated in bold.

Axis	Levels	Option 1	Option 2	Option 3
Region Structure	2	**8Region**	5 Region	
Steepness	3	0.65	**0.80**	0.95
Mixing period	2	**1 quarter**	2 quarters	
Growth	3	**Default level**	Low Growth	High Growth
Length composition sample sizes divided by:	3	50	**100**	200

Table 6: Summary of reference points over the 54 models with the 8 region model in the structural uncertainty grid.

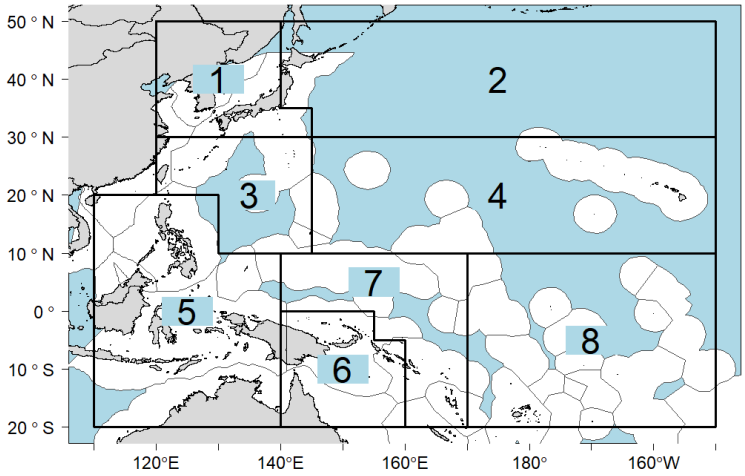
	Mean	Median	Min	10	90	Max
C_{latest}	1755216	1755663	1749846	1753470	1757047	1757083
$Y_{F_{recent}}$	1879015	1857400	1679600	1710240	2043520	2135200
f_{mult}	2.286	2.258	1.472	1.641	2.919	3.705
F_{MSY}	0.223	0.222	0.180	0.189	0.264	0.270
MSY	2297481	2288400	1953600	1993800	2761880	2825600
$F_{recent}/F_{\text{MSY}}$	0.461	0.443	0.270	0.343	0.610	0.679
$SB_{F=0}$	6215233	6243661	5247096	5548226	6867055	7349558
SB_{MSY}	1100633	1063500	631900	711460	1544800	1688000
$SB_{MSY}/SB_{F=0}$	0.175	0.176	0.117	0.130	0.223	0.230
$SB_{latest}/SB_{F=0}$	0.413	0.412	0.325	0.344	0.481	0.525
$SB_{latest}/SB_{\text{MSY}}$	2.471	2.376	1.551	1.665	3.441	3.925
$SB_{recent}/SB_{F=0}$	0.440	0.438	0.336	0.363	0.521	0.551
$SB_{recent}/SB_{\text{MSY}}$	2.631	2.581	1.601	1.777	3.664	4.139

Table 7: Summary of reference points over the 54 models with the 5 region model in the structural uncertainty grid.

	Mean	Median	Min	10	90	Max
C_{latest}	1749645	1750173	1746602	1747800	1750975	1751070
$Y_{F_{recent}}$	1739689	1704400	1518400	1576040	1913320	1994000
f_{mult}	2.228	2.096	1.385	1.519	2.896	3.687
F_{MSY}	0.222	0.222	0.176	0.187	0.264	0.267
MSY	2089941	2068400	1716800	1779480	2445280	2617600
$F_{recent}/F_{\text{MSY}}$	0.477	0.477	0.271	0.345	0.660	0.722
$SB_{F=0}$	5819526	5784069	4639059	4972127	6724511	7551201
SB_{MSY}	993772	956850	554000	630520	1343300	1701000
$SB_{MSY}/SB_{F=0}$	0.168	0.165	0.112	0.127	0.214	0.225
$SB_{latest}/SB_{F=0}$	0.372	0.361	0.268	0.289	0.457	0.501
$SB_{latest}/SB_{\text{MSY}}$	2.318	2.234	1.337	1.448	3.203	3.885
$SB_{recent}/SB_{F=0}$	0.410	0.401	0.299	0.317	0.504	0.541
$SB_{recent}/SB_{\text{MSY}}$	2.553	2.488	1.483	1.592	3.505	4.193

10 Figures

(a) 8 Region



(b) 5 Region

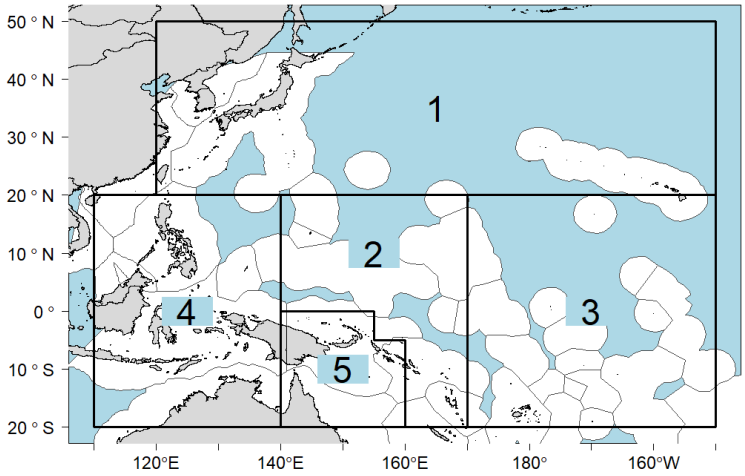


Figure 1: The geographical area covered by the stock assessment and the boundaries for the 8 region (top) and 5 region (bottom) assessment models.

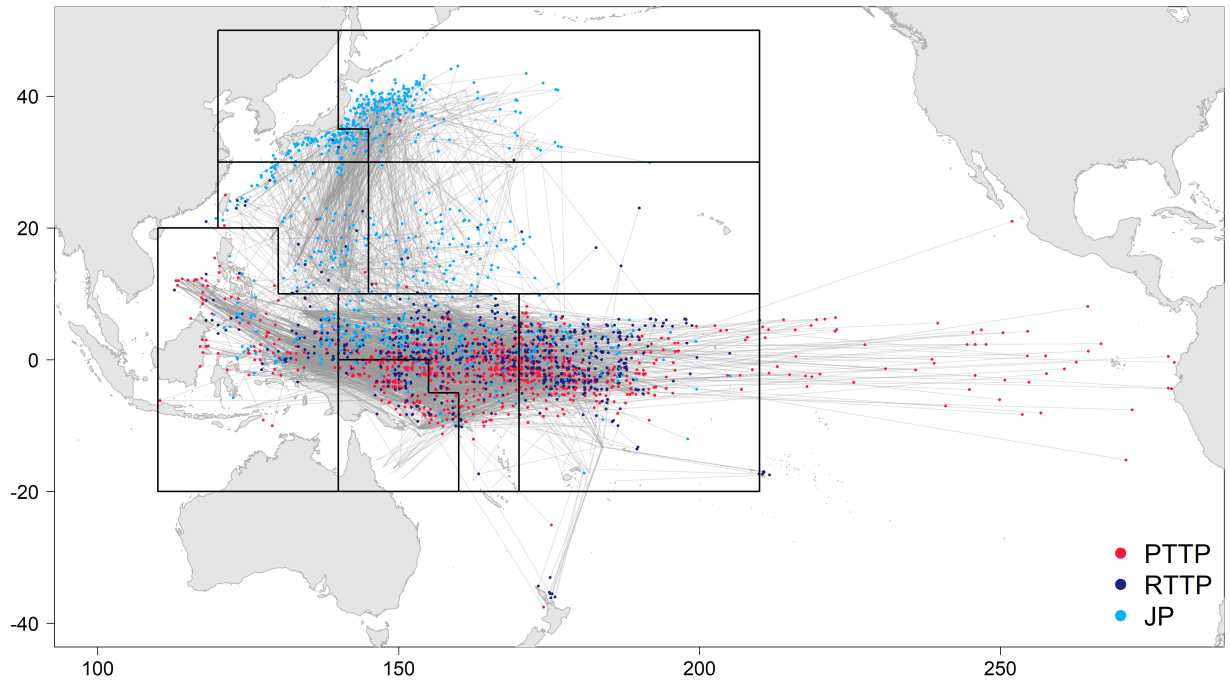


Figure 2: Plot of tag recaptures greater than 1,000 nautical miles from the point of release by the program of release for those tags that were released within the assessment region.

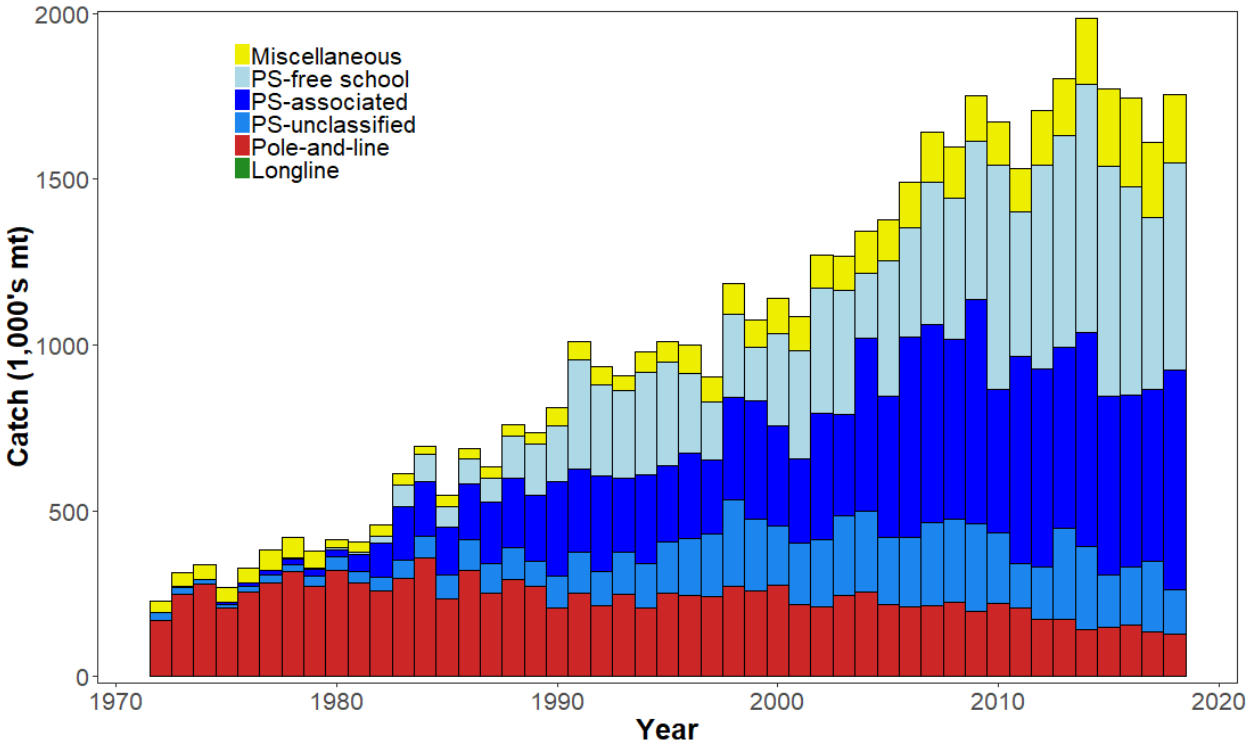


Figure 3: Time series of total annual catch (1000's mt) by fishing gear from the diagnostic model over the full assessment period.

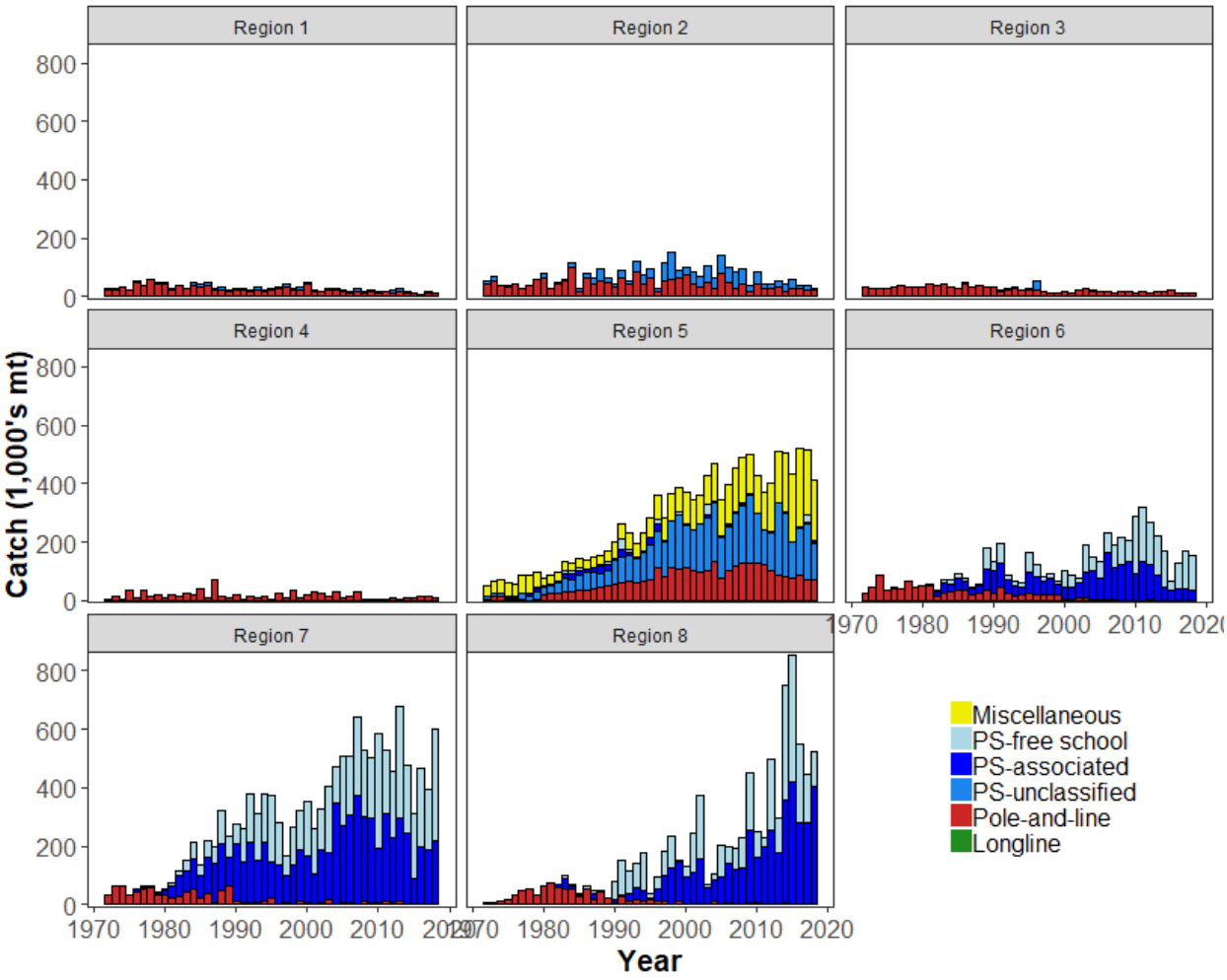


Figure 4: Time series of total annual catch (1000's mt) by fishing gear and assessment region from the diagnostic model over the full assessment period.

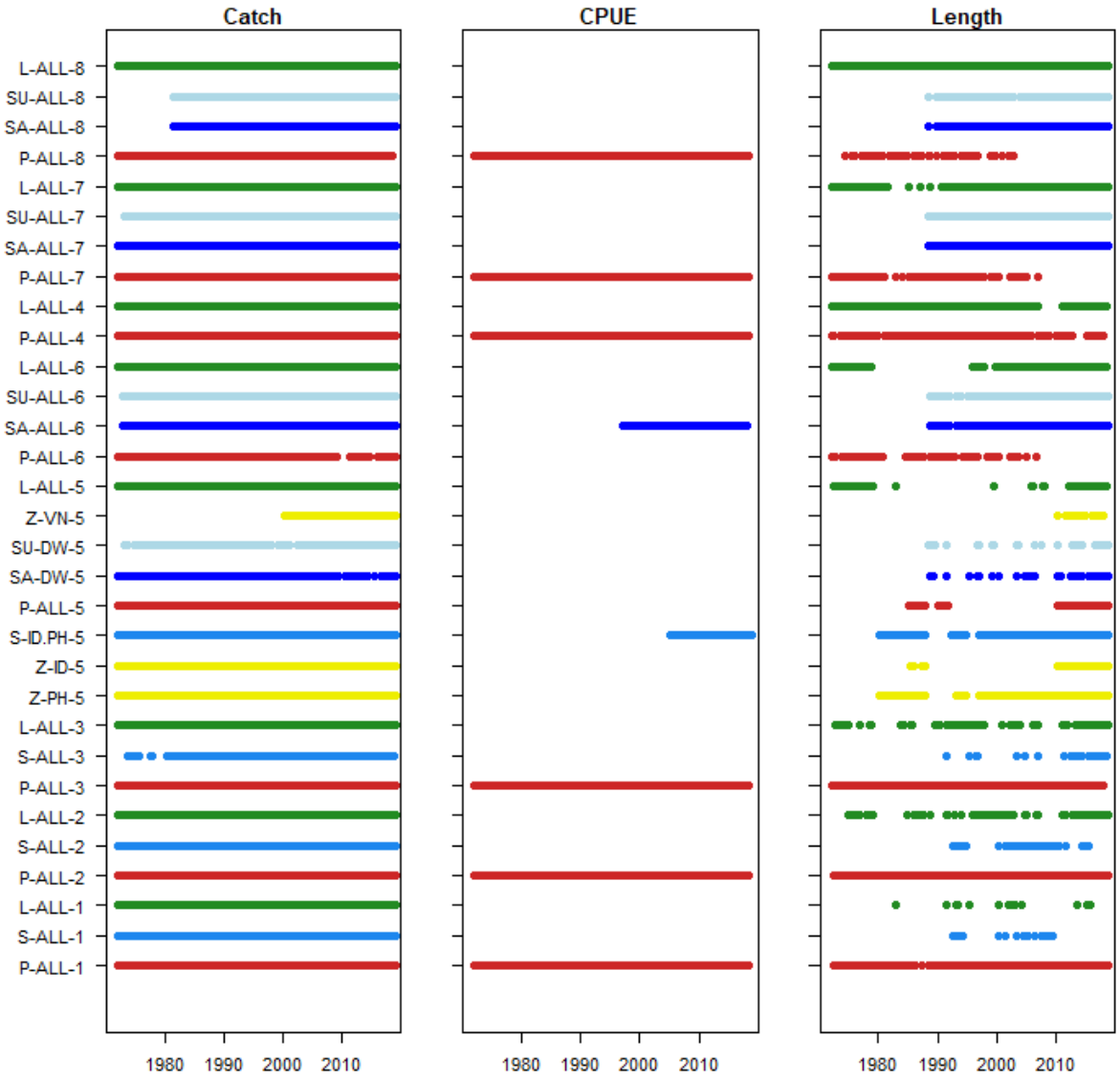


Figure 5: Presence of catch, standardised CPUE, and length frequency data by year and fishery for the diagnostic model. The different colors refer to longline (green), pole-and-line (red), purse seine (blue) and miscellaneous (yellow).

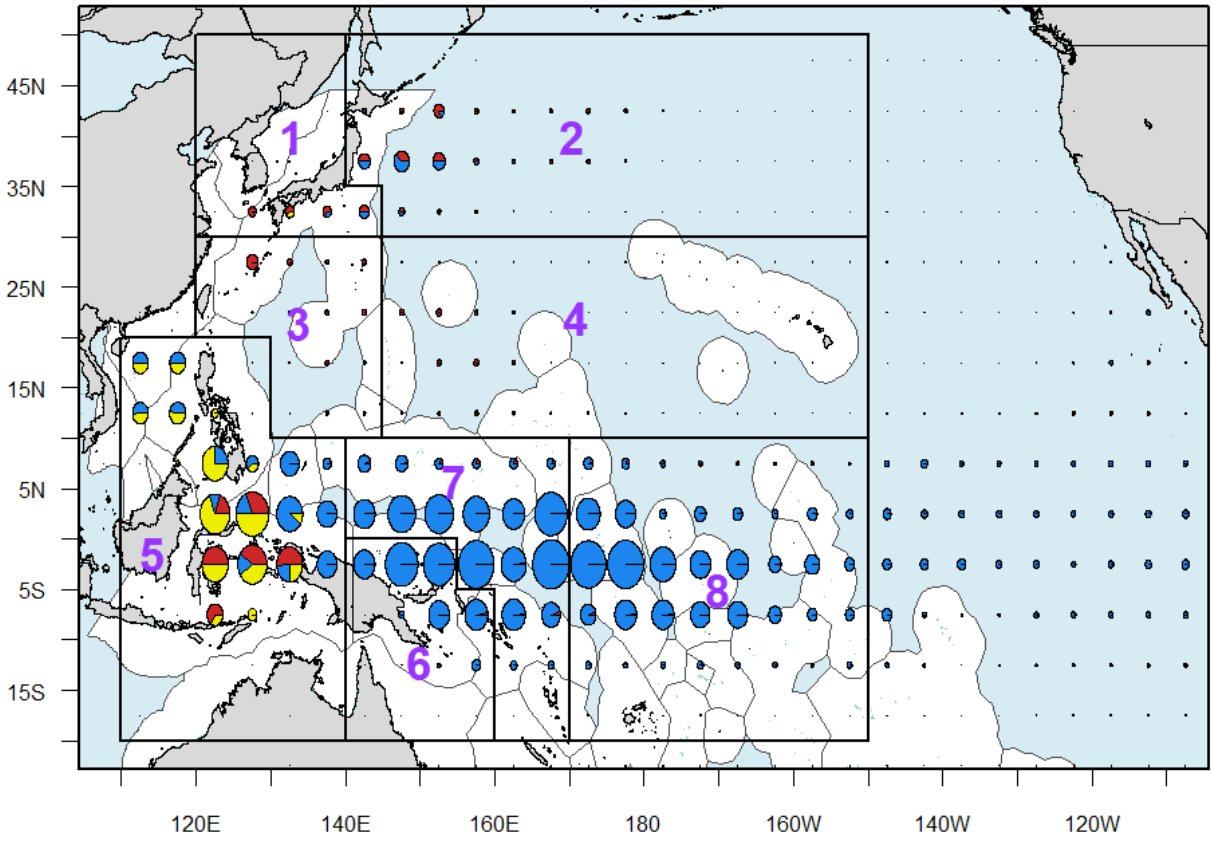


Figure 6: Distribution and magnitude of skipjack catches for the most recent decade of the stock assessment (2009-2018) by 5° square and fishing gear: longline (green), pole-and-line (red), purse seine (blue) and miscellaneous (yellow), for the entire Pacific Ocean. Overlayed are the regional boundaries for the stock assessment.

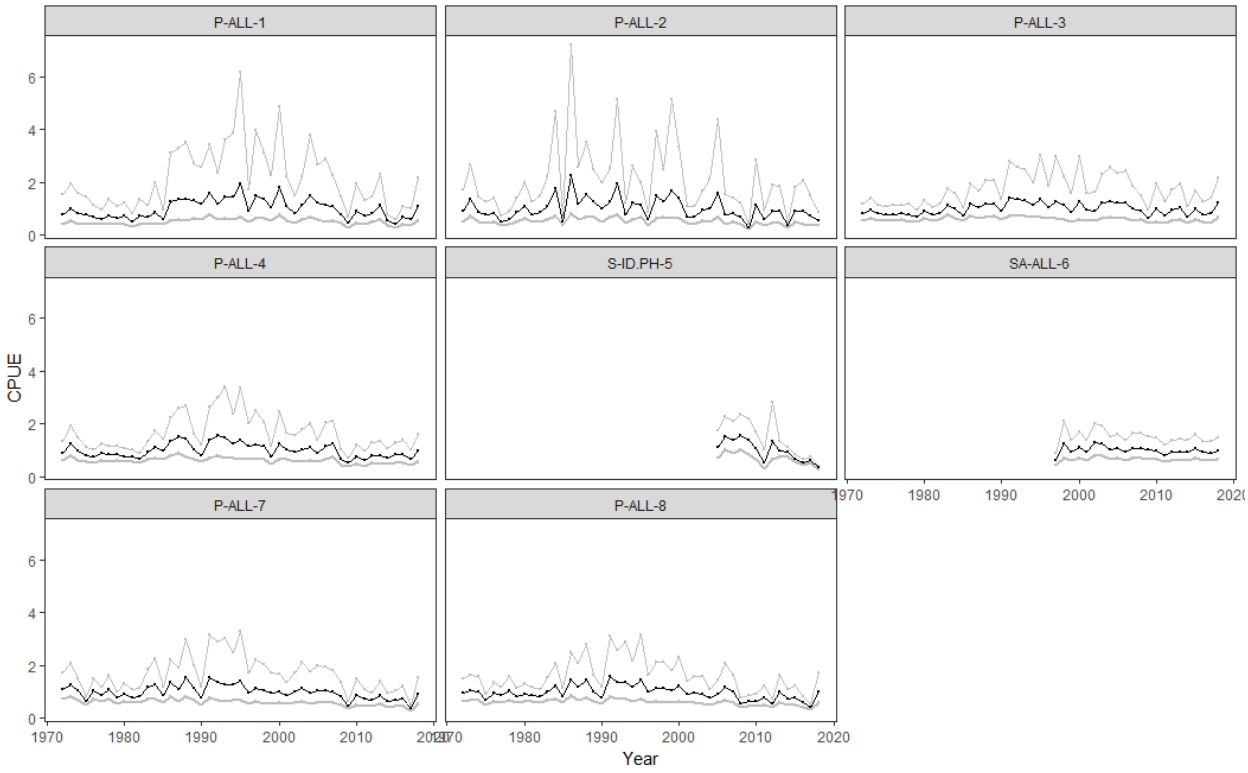


Figure 7: Standardized catch-per-unit-effort (CPUE) indices for the pole-and-line fisheries from the geostatistical model in units of mt/day for regions 1, 2, 3, 4, 7, and 8 (P-ALL-1, P-ALL-2, P-ALL-3, P-ALL-4, P-ALL-7, P-ALL-8), the ID/PH domestic purse seine fishery in units of mt/day for region 5 (S-ID.PH-5), and the associated purse seine fishery in units of mt/set for region 6 (S-ASS-ALL-6), used in the diagnostic model. See Bigelow et al. (2019), Kinoshita et al. (2019) and Vidal et al. (2019) for further details of the estimation of these CPUE indices. The light gray lines represent the 95% confidence intervals from which the effort deviation penalties used in the diagnostic model are derived.

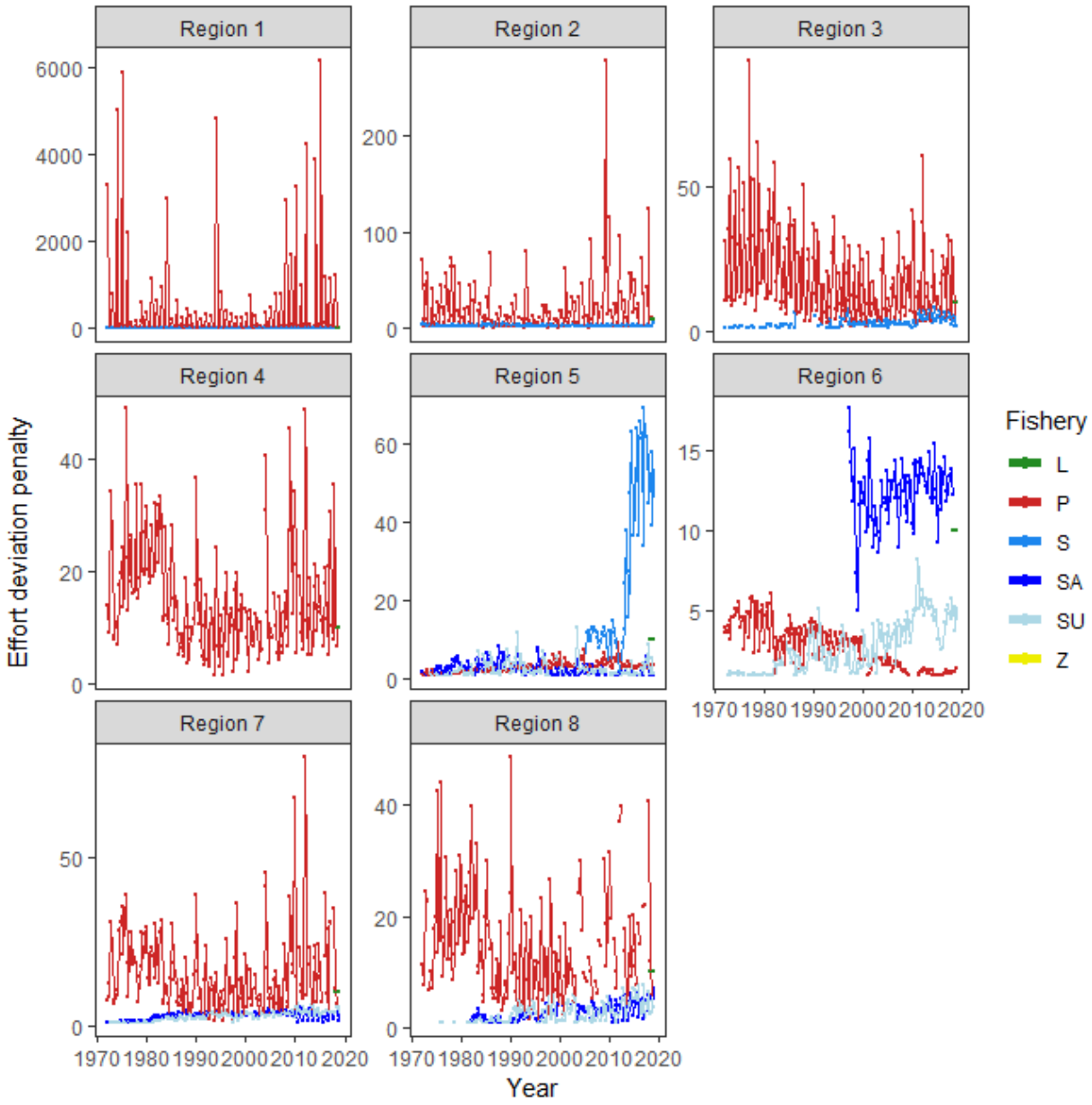


Figure 8: Plot of the effort deviation penalties applied to each fishery, by region, where a higher penalty gives more weight to the CPUE for that observation. longline (LL) fisheries and miscellaneous (Z) fisheries do not have effort in the model and thus no penalties.

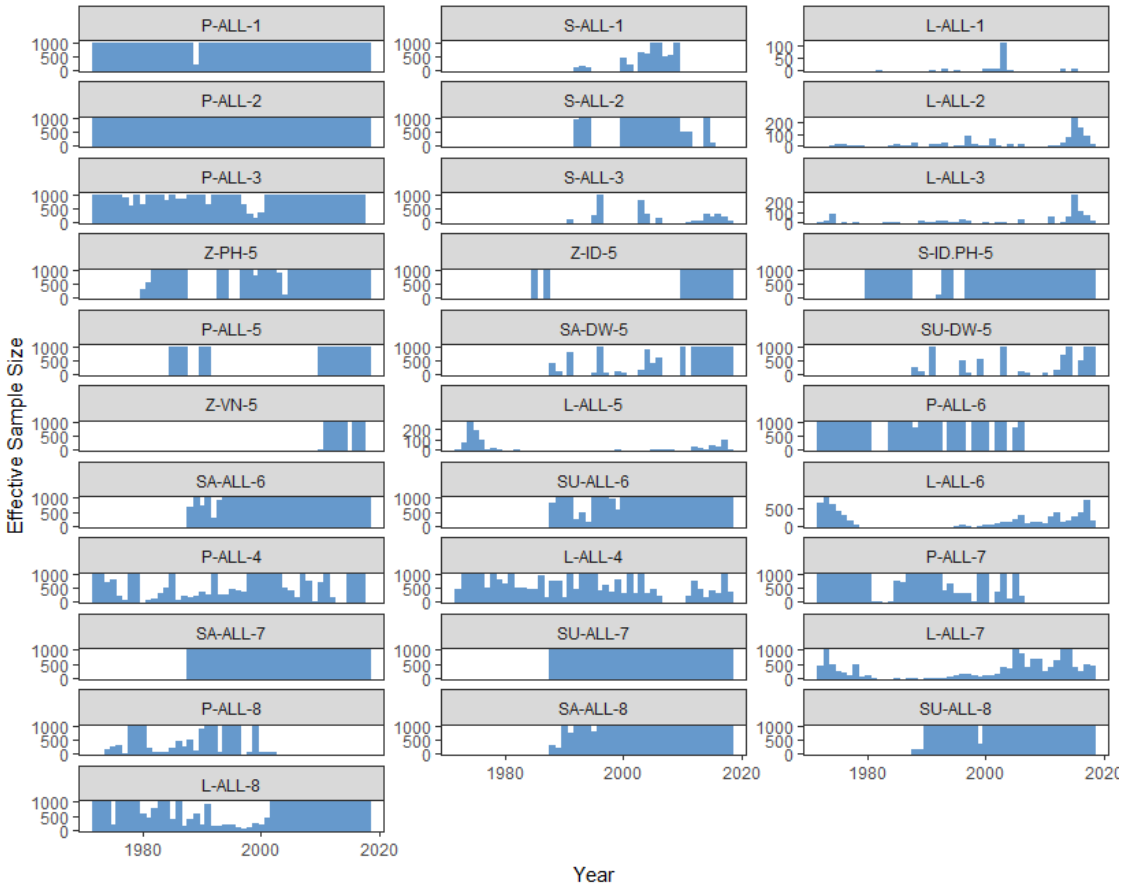


Figure 9: Number of length frequency samples available for each fishery for the diagnostic model capped at a maximum sample size of 1,000. These samples are further down-weighted in the model according to defined settings presented in Section 5.5.

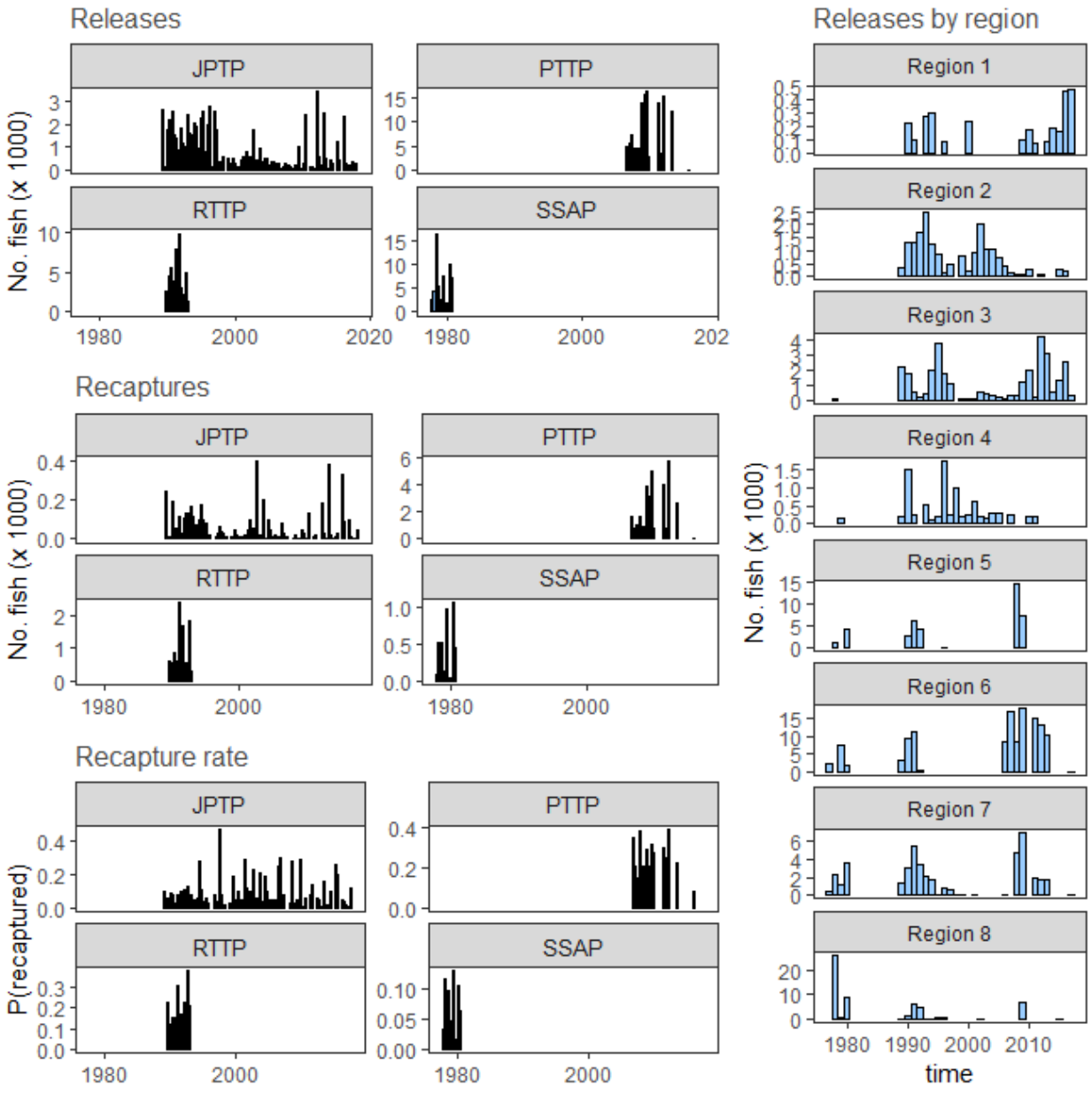


Figure 10: Summary of the number of releases, recaptures, and recapture rate of tags, by tagging program and region.

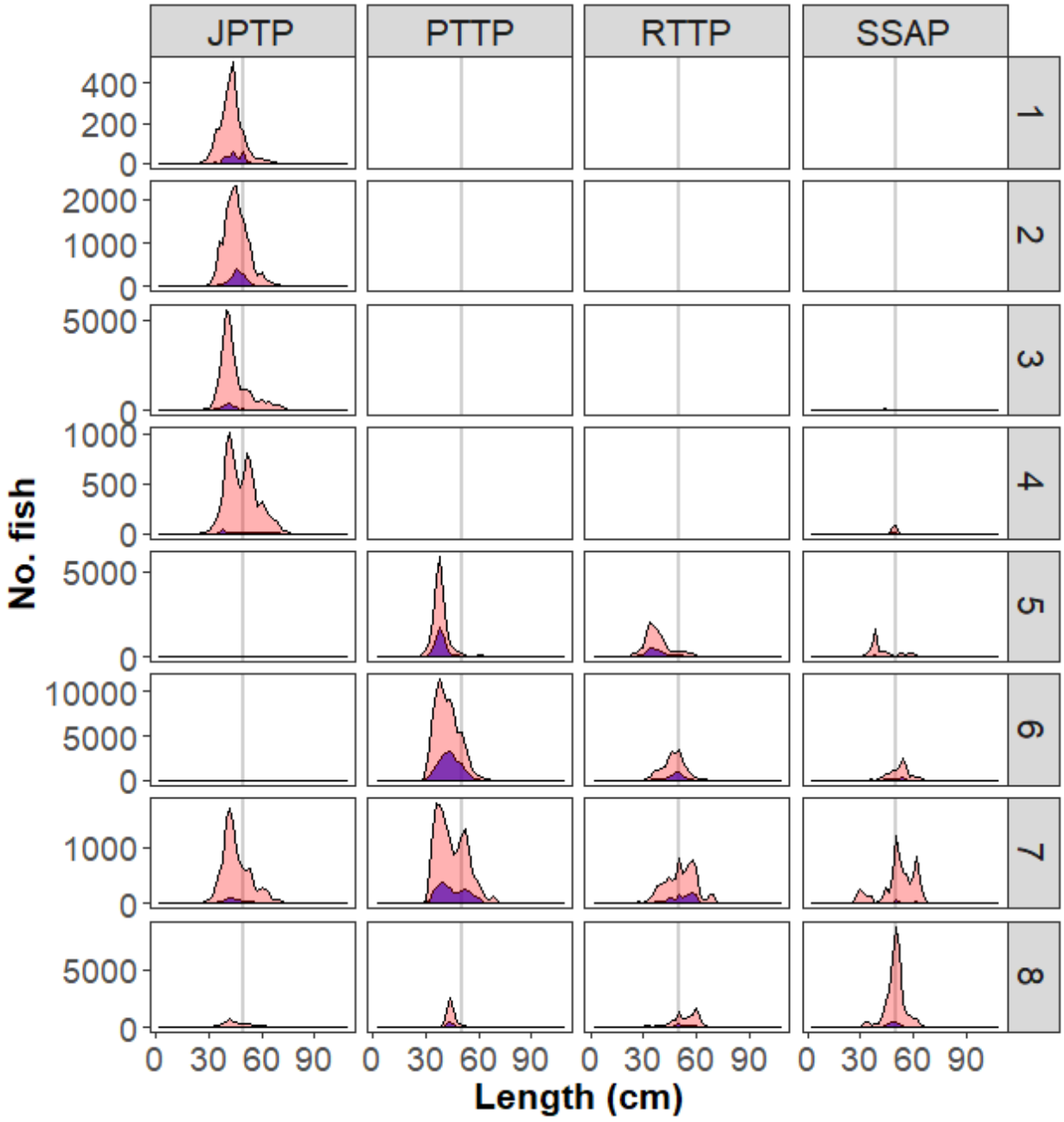


Figure 11: Summary of the length composition of released (pink) and recaptured (purple) fish for the different tagging programs (columns) and regions (rows), where the length of 50 cm is denoted by the vertical line.

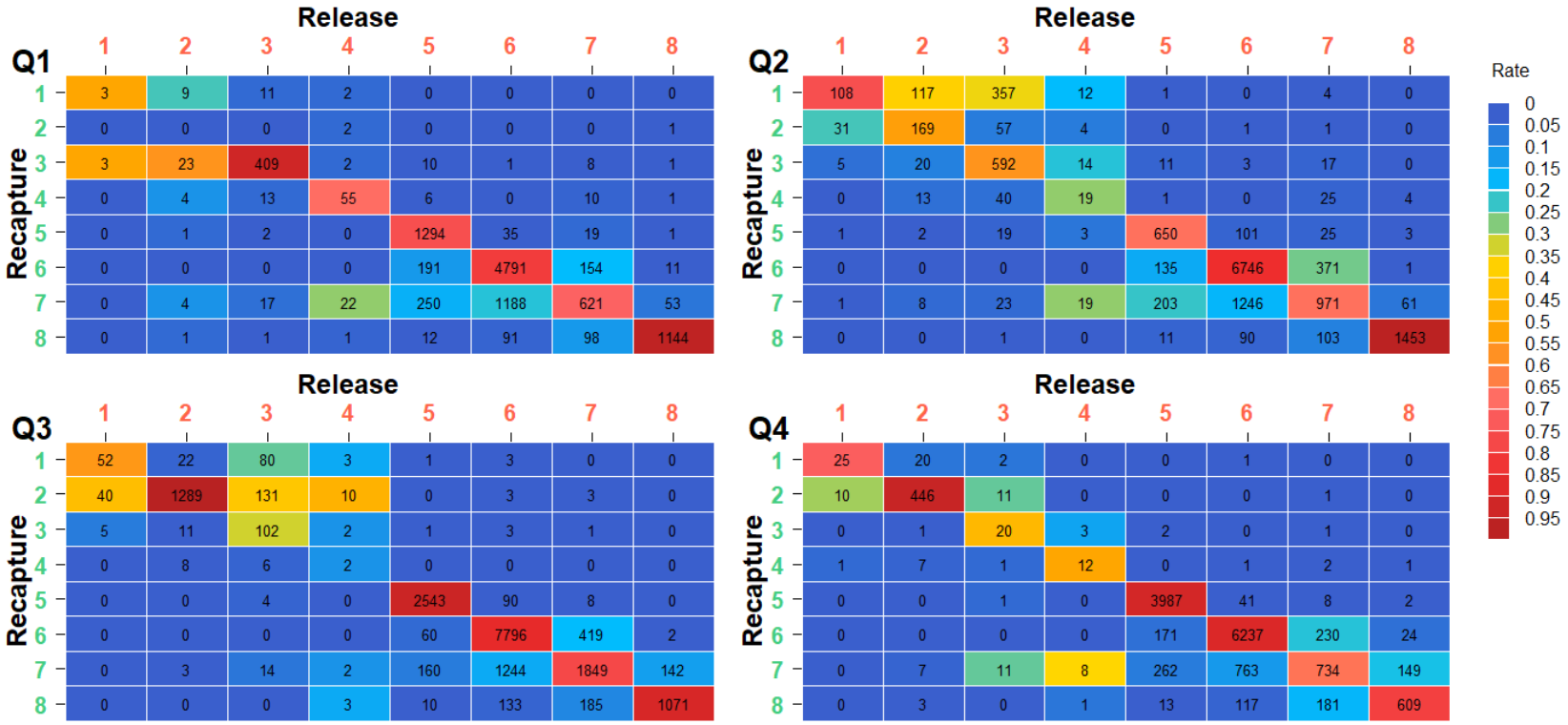


Figure 12: Summary of the number of tags recaptured by region of release (columns), region of recapture (rows), and quarter of recapture (panel). The shade of the tile corresponds to the proportion of tags released from that region that are recovered in the corresponding region.

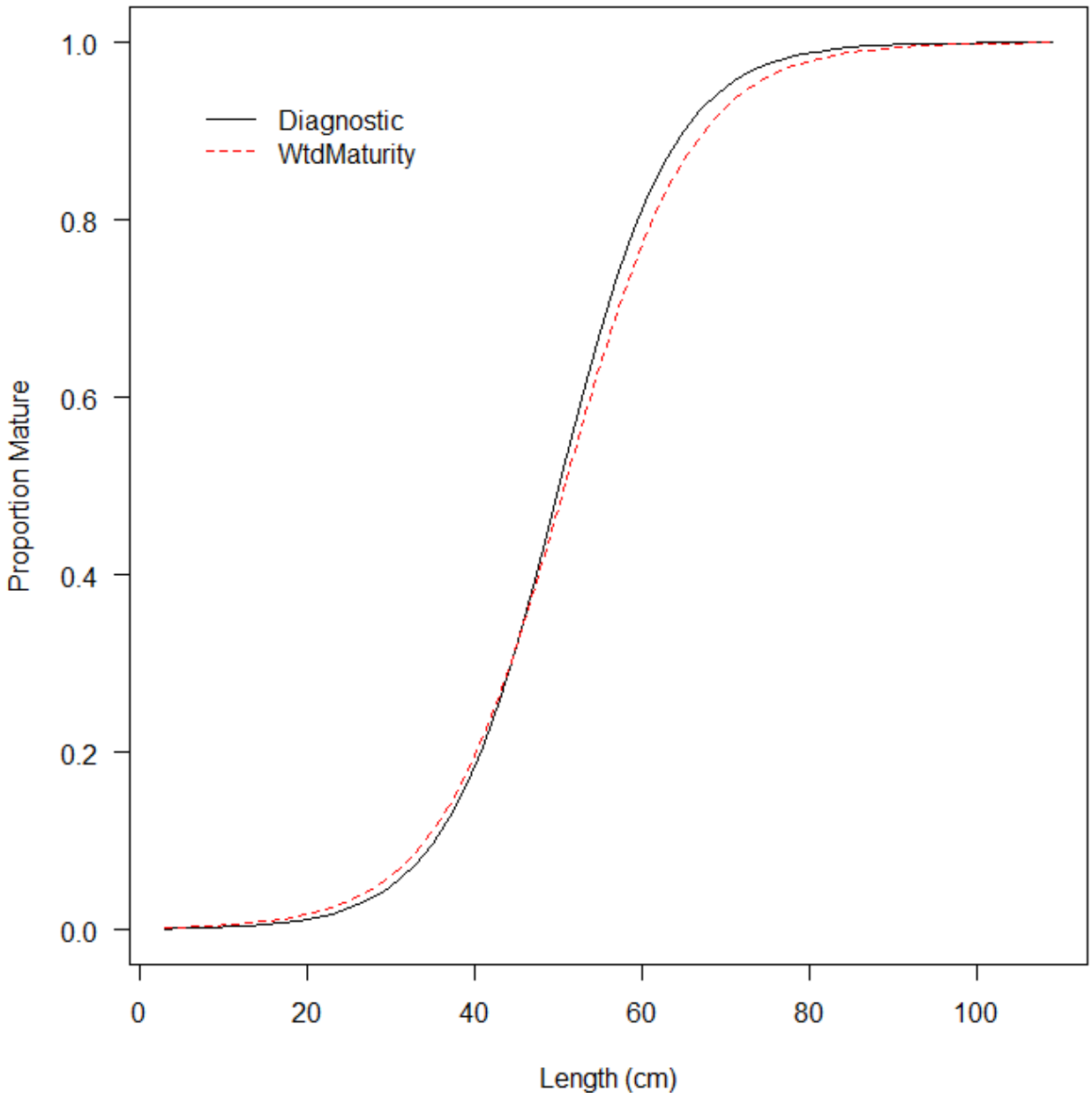


Figure 13: Maturity-at-length assumed for all models in the structural uncertainty grid (solid black line) and the one-off sensitivity model (*WtdMaturity*; dashed red line).

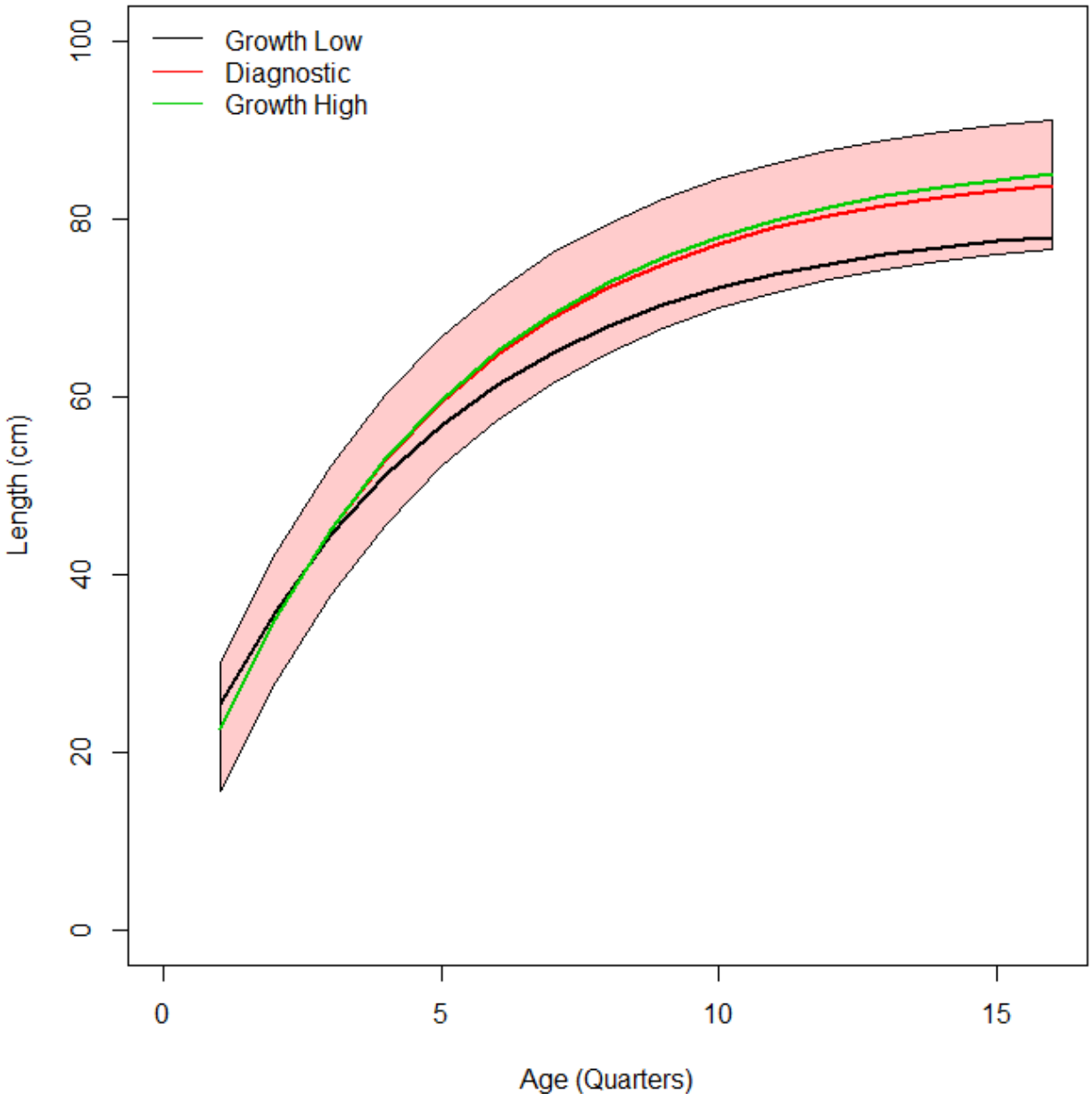


Figure 14: Growth curves estimated by different scaling factors of the length composition data that were fixed within the structural uncertainty grid. The red shaded region is two times the standard deviation at length for the diagnostic growth curve.

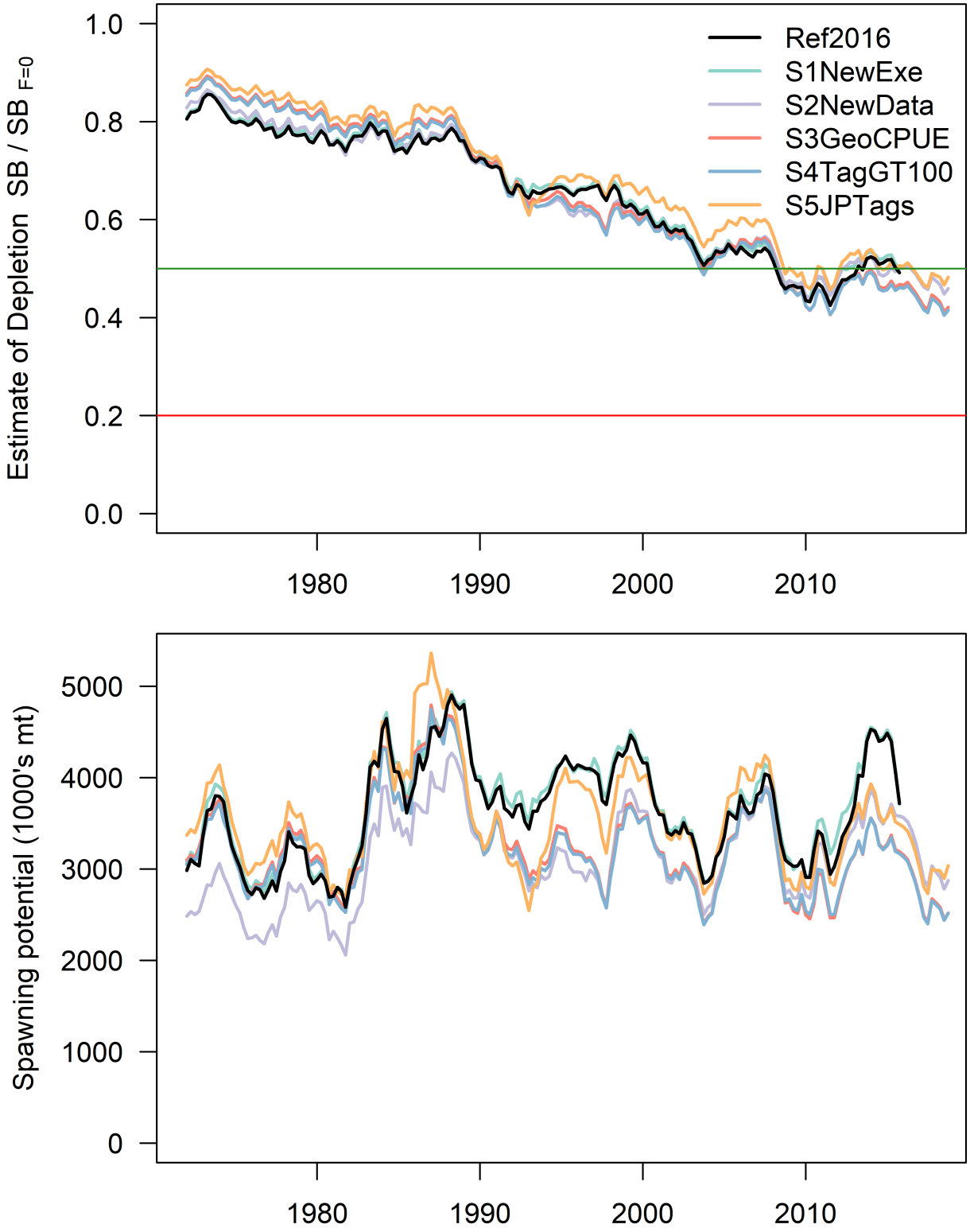


Figure 15: Estimated dynamic depletion of spawning potential ($SB_t/SB_{t,F=0}$) (top panel) and the estimated spawning potential are shown for the first set of stepwise models in the progression from the 2016 reference case to the 2019 diagnostic case.

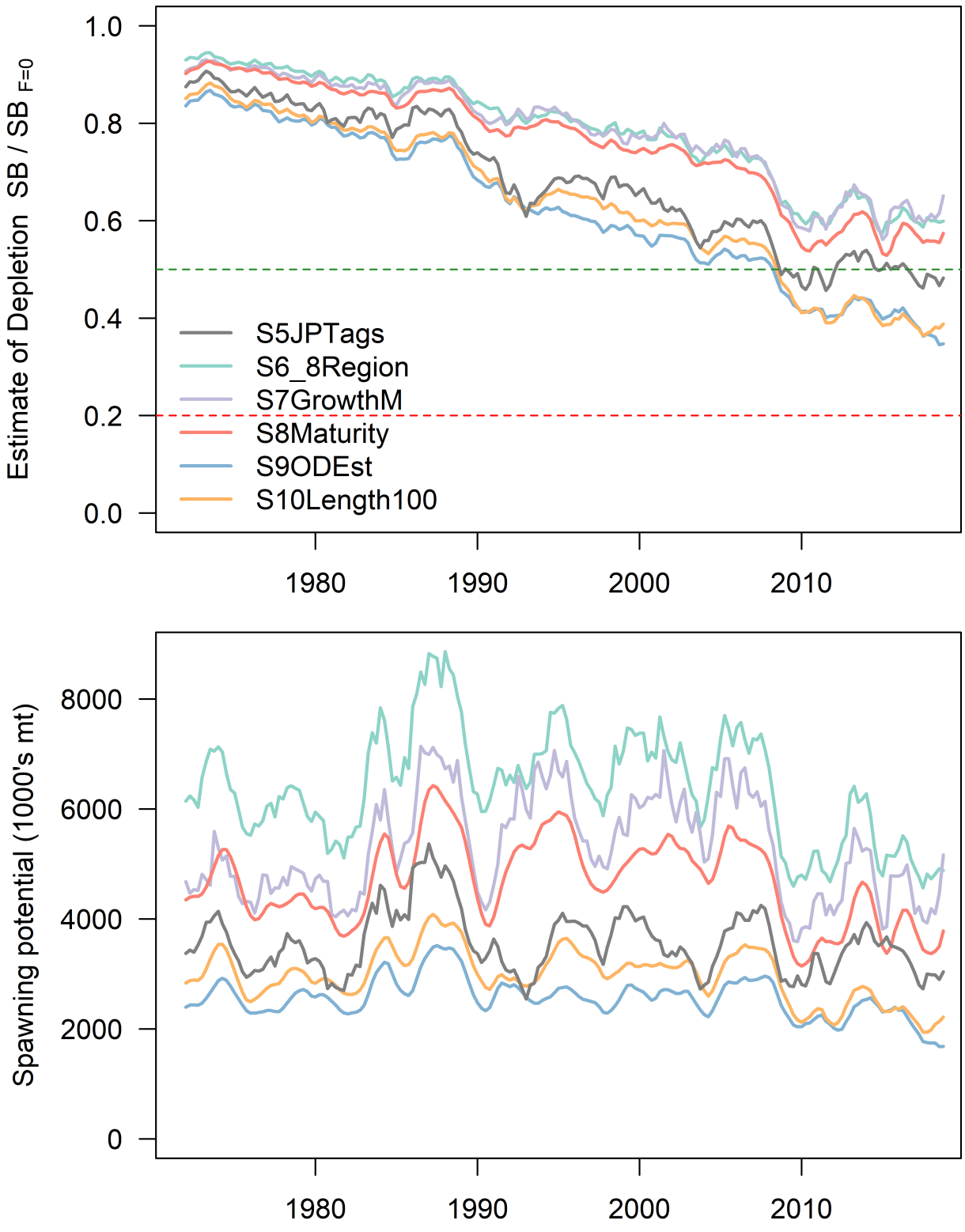


Figure 16: Estimated dynamic depletion of spawning potential ($SB_t/SB_{t,F=0}$) (top panel) and the estimated spawning potential are shown for the second set of stepwise models in the progression from the 2016 reference case to the 2019 diagnostic case (*Length100*), where the gray line is the last step in the Figure 16.

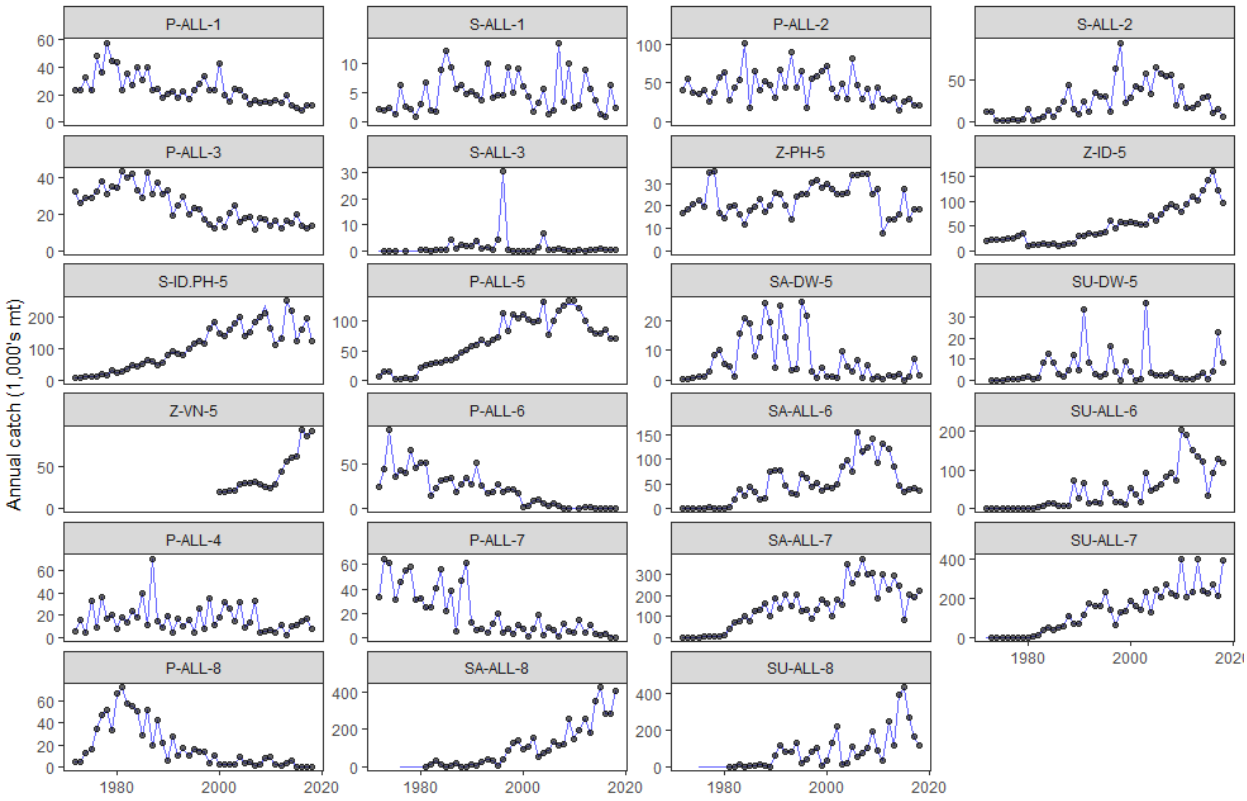


Figure 17: Observed (black points) and model-predicted (blue lines) catch in weight for the fisheries in the diagnostic model excluding the longline fisheries because they were given small nominal catches.

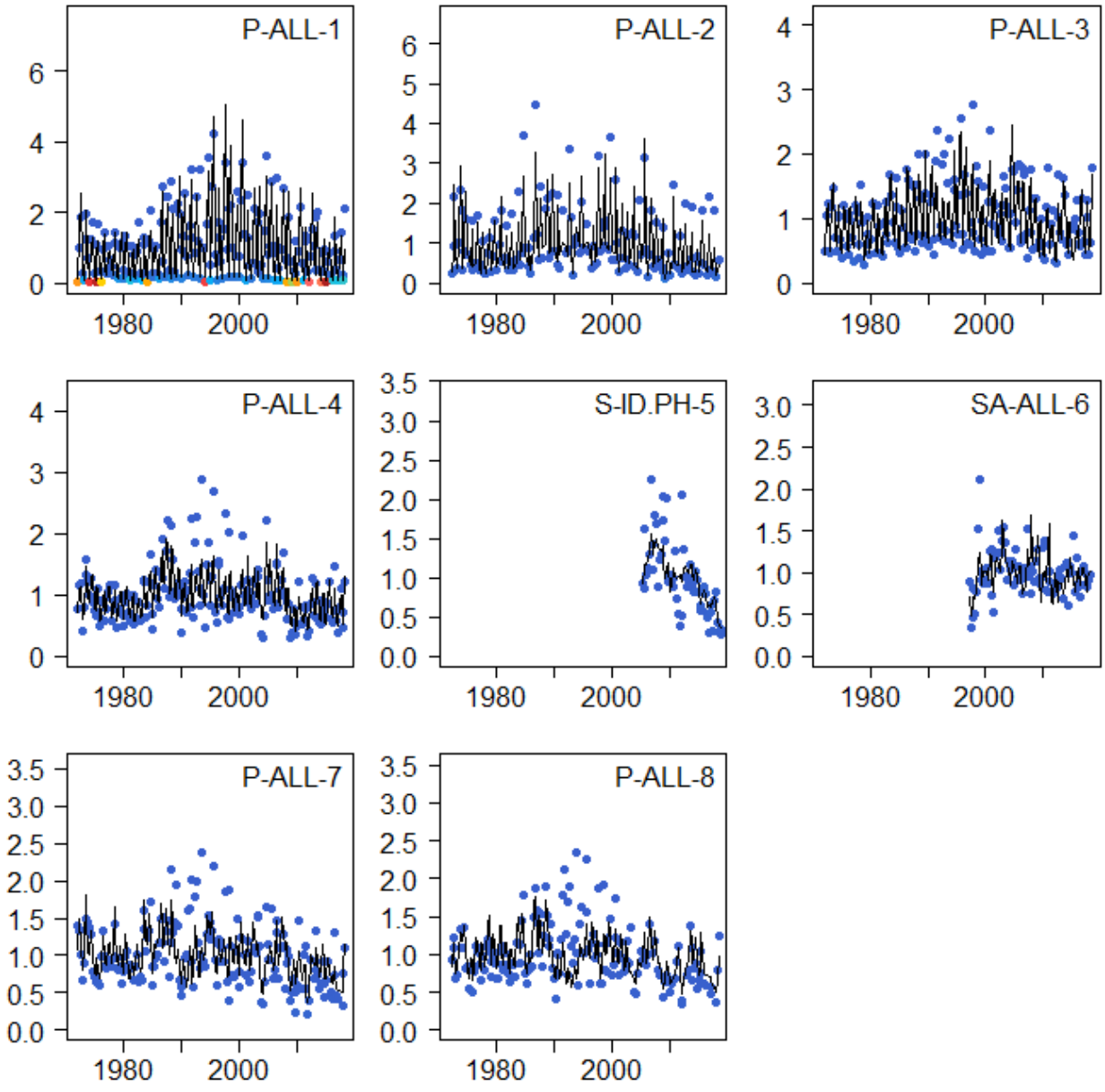


Figure 18: Observed (colored points) and model-predicted (black lines) CPUE for the eight fisheries which received standardised CPUE indices in the diagnostic model, where the color of the point indicates the penalty applied to the observed CPUE and brighter colors indicate a larger penalty (see Figure 8).

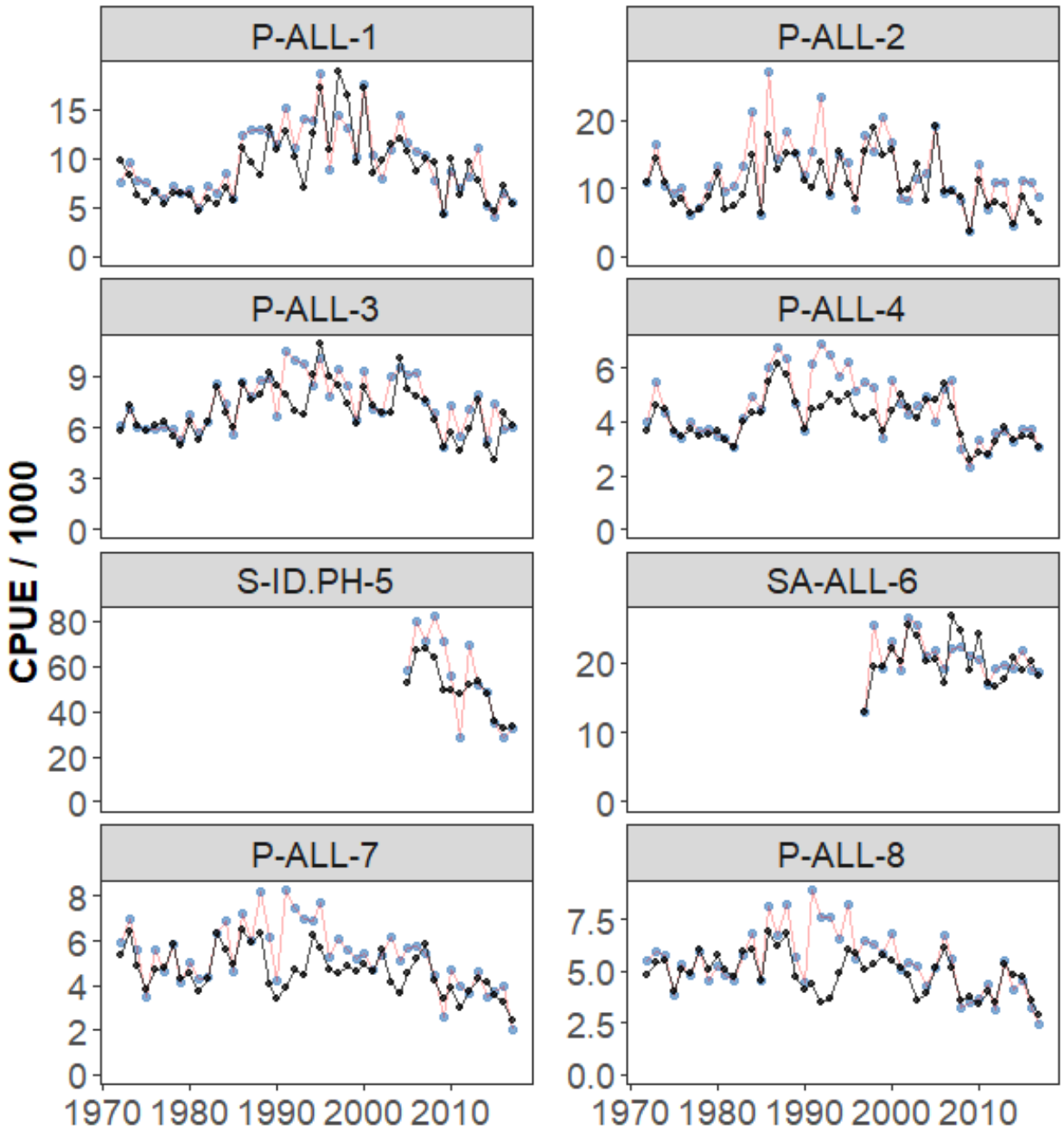


Figure 19: Observed (blue points and red lines) and model-predicted (black points and lines) annual mean CPUE for the eight fisheries which received standardised CPUE indices in the diagnostic model.

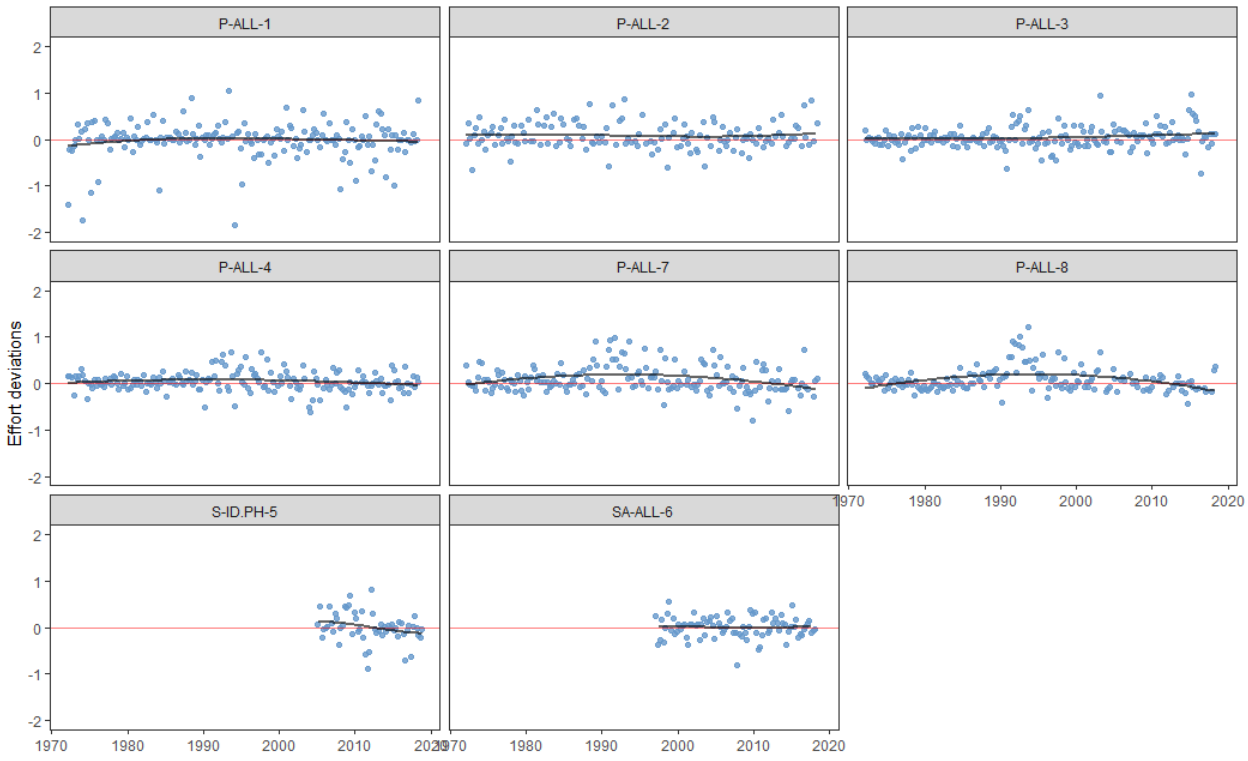


Figure 20: Effort deviations by time period for each of the fisheries receiving standardised CPUE indices in the diagnostic model. The line represents a Lowess smoothed fit to the effort deviations.

Figure 21: Effort deviations by time period for each of the fisheries that did not receive standardised CPUE indices in the diagnostic model. The line represents a Lowess smoothed fit to the effort deviations.

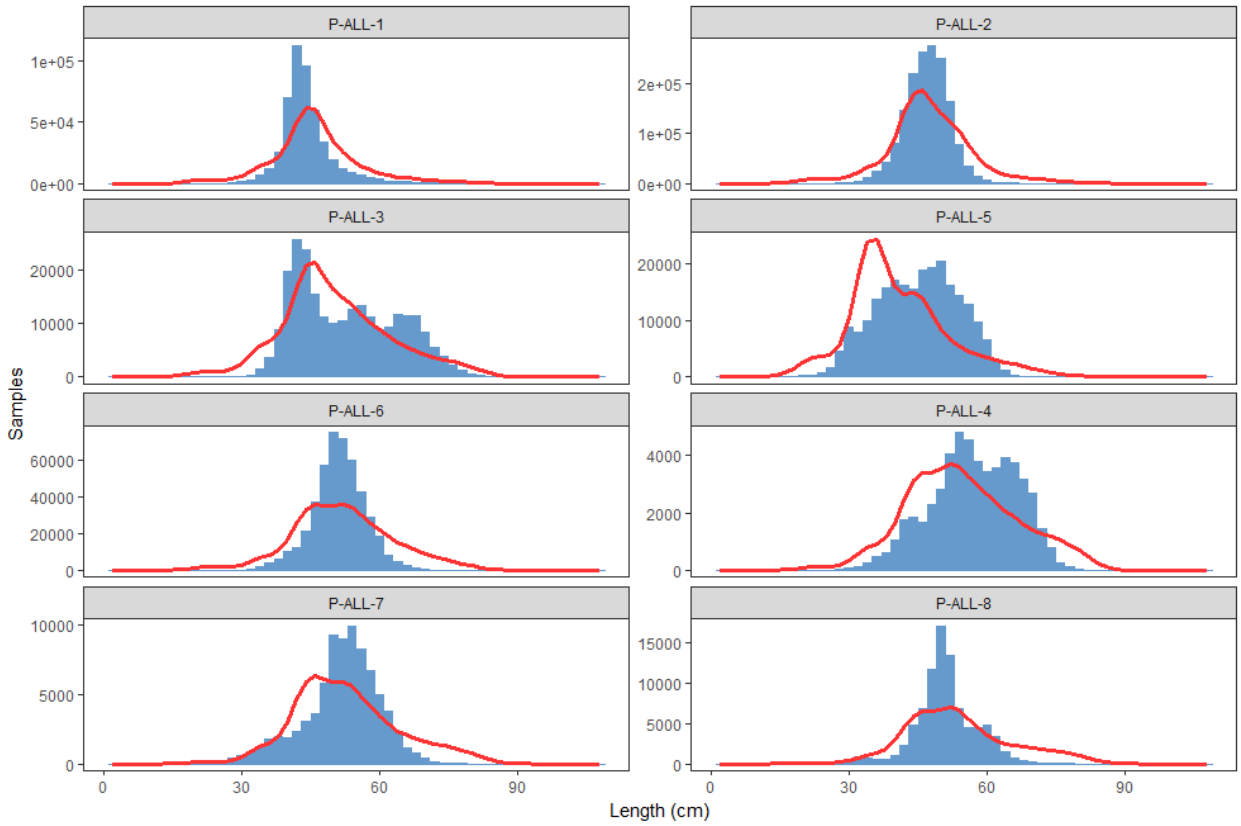


Figure 22: Composite (all time periods combined) observed (blue histograms) and predicted (red line) catch-at-length for pole-and-line fisheries the diagnostic model.

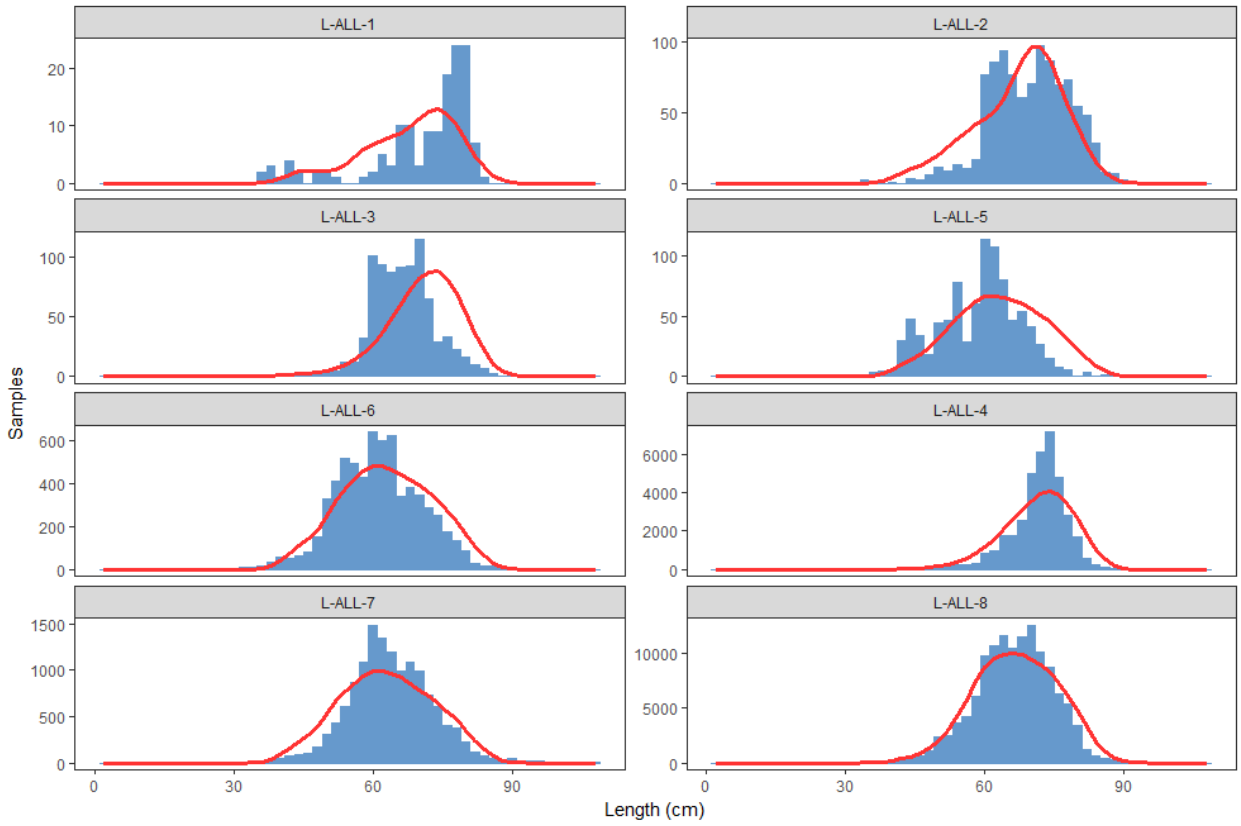


Figure 23: Composite (all time periods combined) observed (blue histograms) and predicted (red line) catch-at-length for longline fisheries for the diagnostic model.

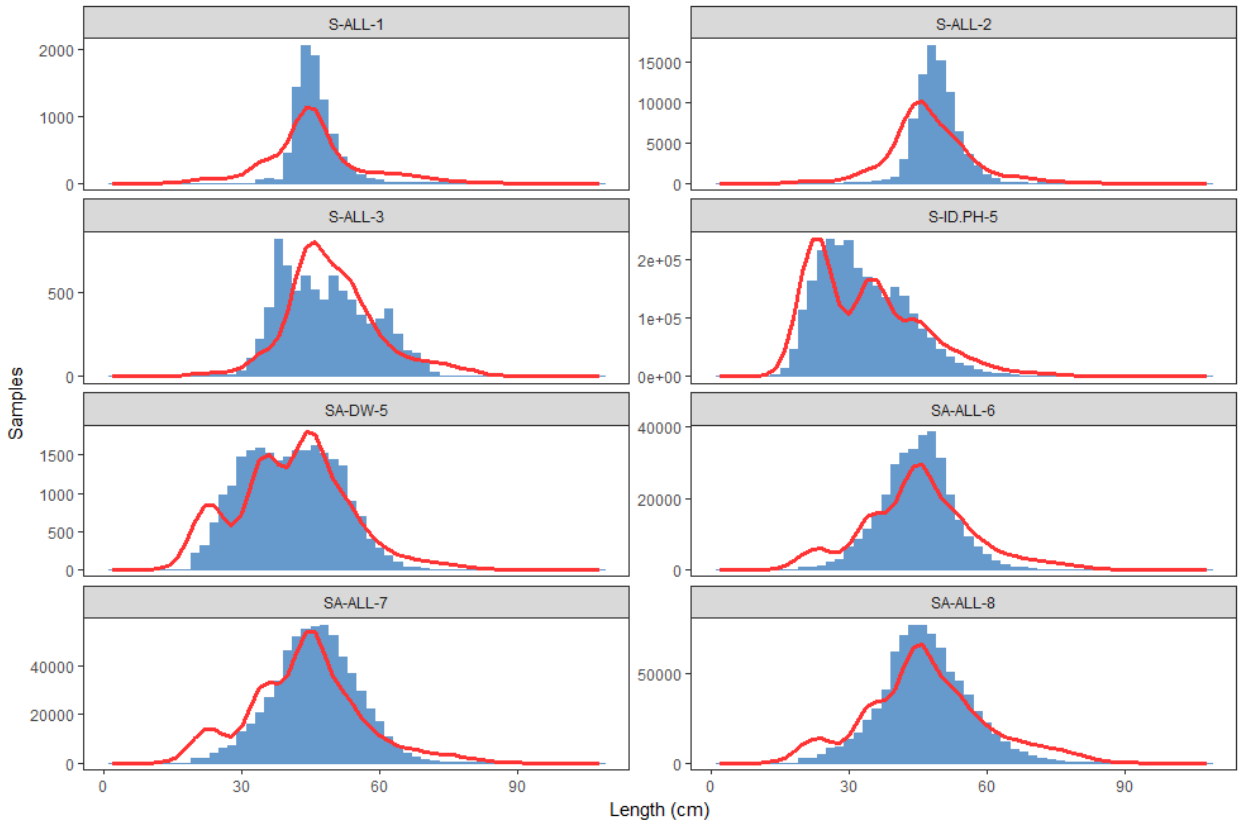


Figure 24: Composite (all time periods combined) observed (blue histograms) and predicted (red line) catch-at-length for associated or combined purse seine fisheries for the diagnostic model.

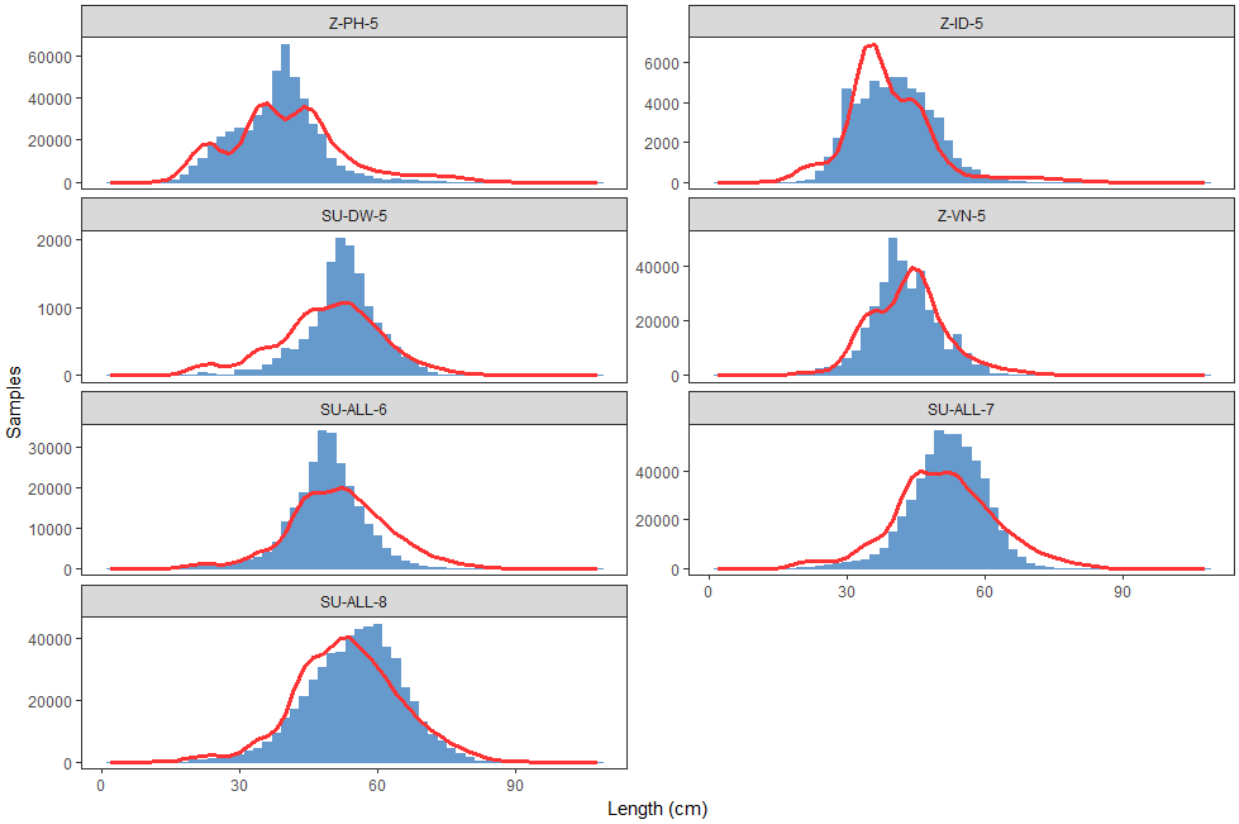


Figure 25: Composite (all time periods combined) observed (blue histograms) and predicted (red line) catch-at-length for unassociated purse seine and miscellaneous fisheries for the diagnostic model.

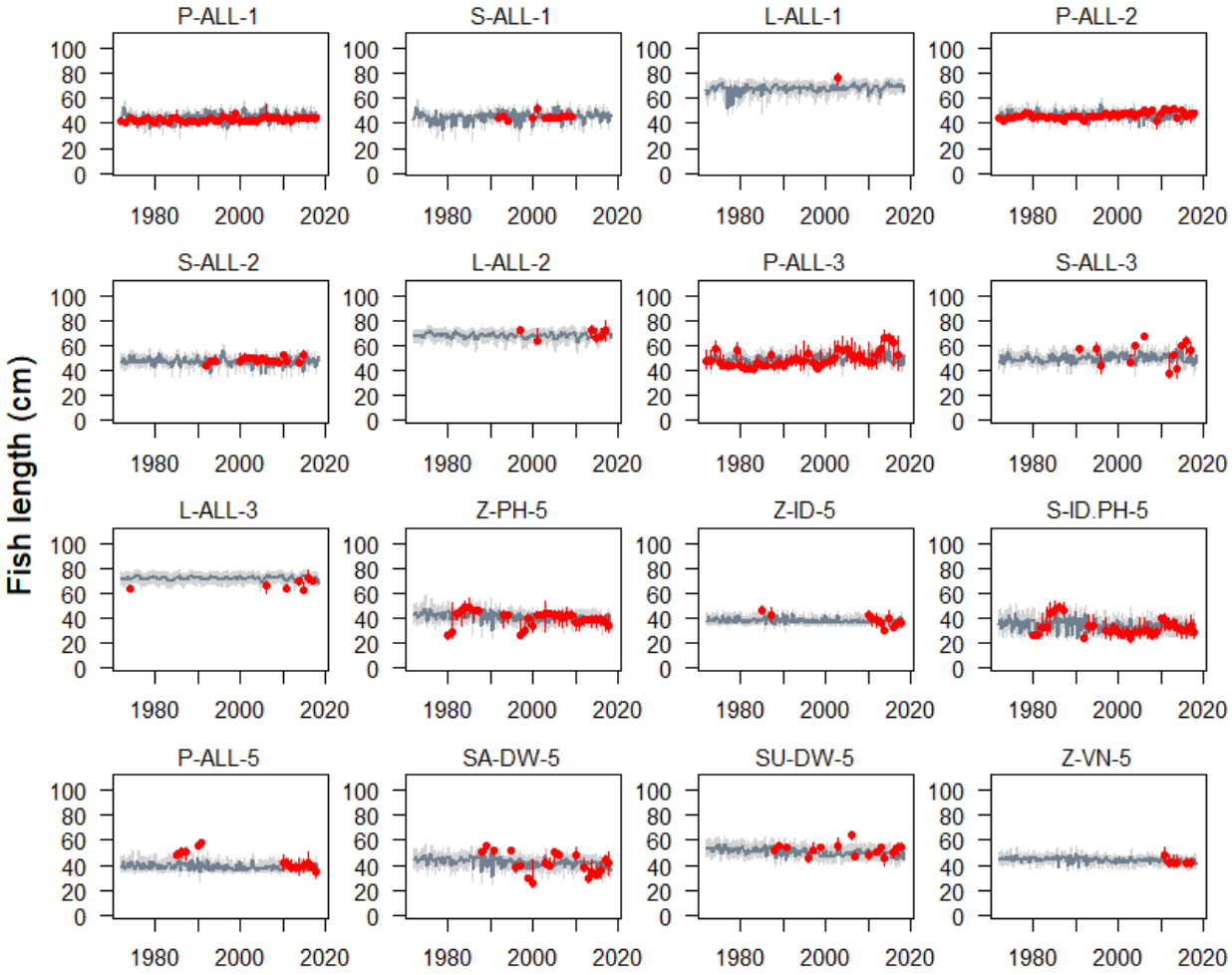


Figure 26: A comparison of the observed (red points) and predicted (gray line) median fish length (FL, cm) for the first set of fisheries with samples for the diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples with a minimum of 30 fish per quarter are plotted.

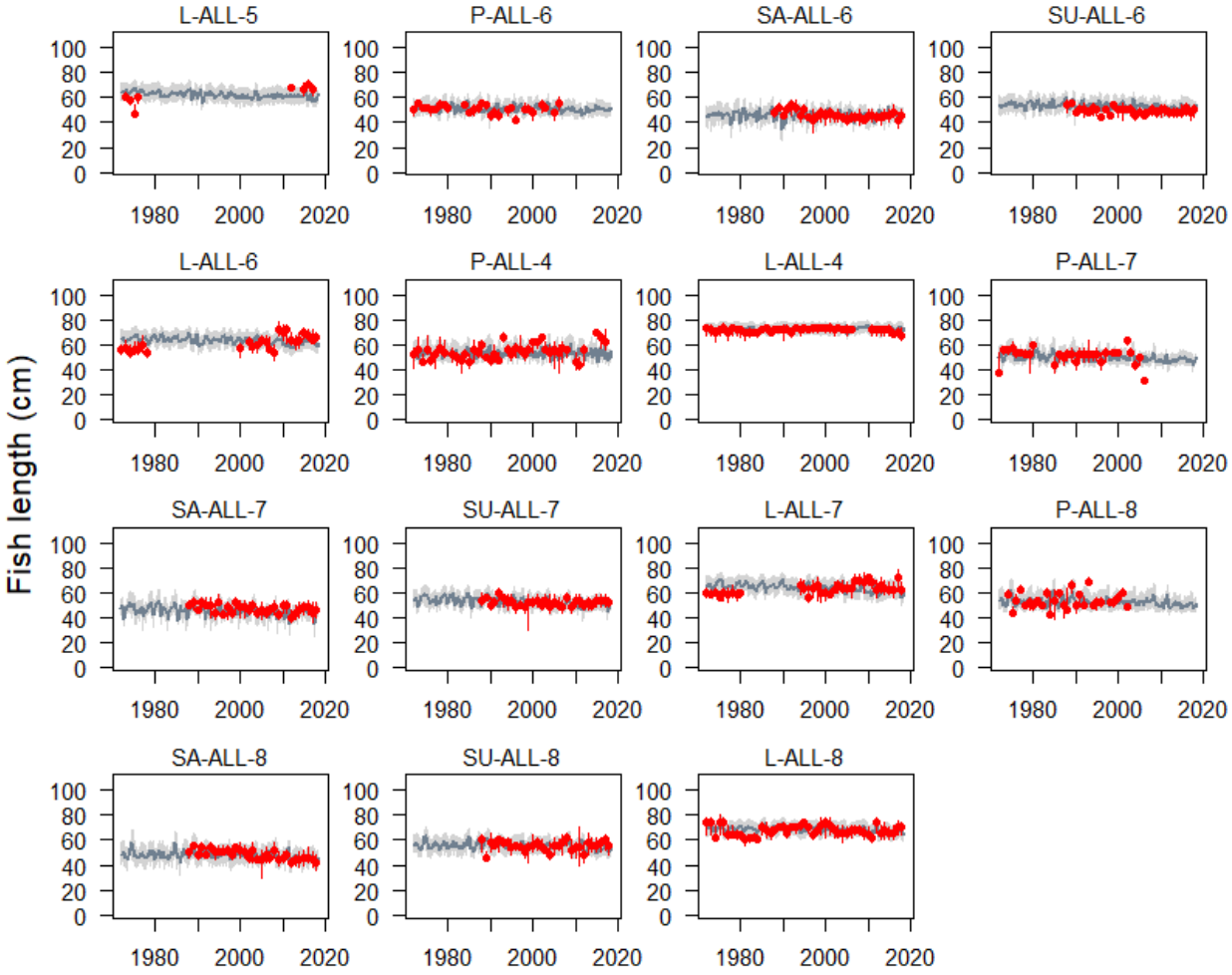


Figure 27: A comparison of the observed (red points) and predicted (gray line) median fish length (FL, cm) for the second set of fisheries with samples for the diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples with a minimum of 30 fish per quarter are plotted.

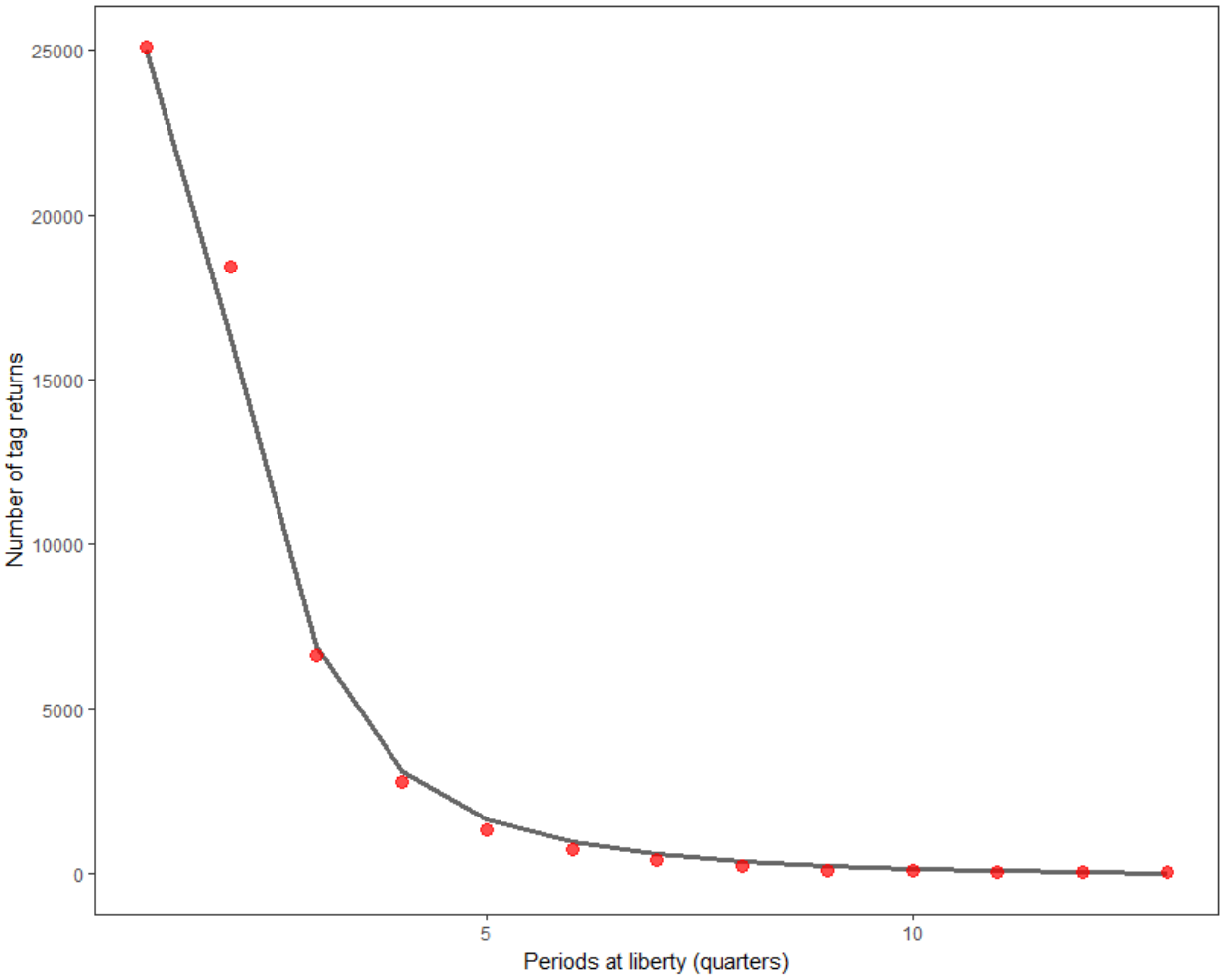


Figure 28: Observed (red points) and model-predicted (black line) tag attrition across all tag release events for the diagnostic model. The exact correspondence between observed and predicted values for the first quarter-at-liberty is a direct result of the tag mixing assumption.

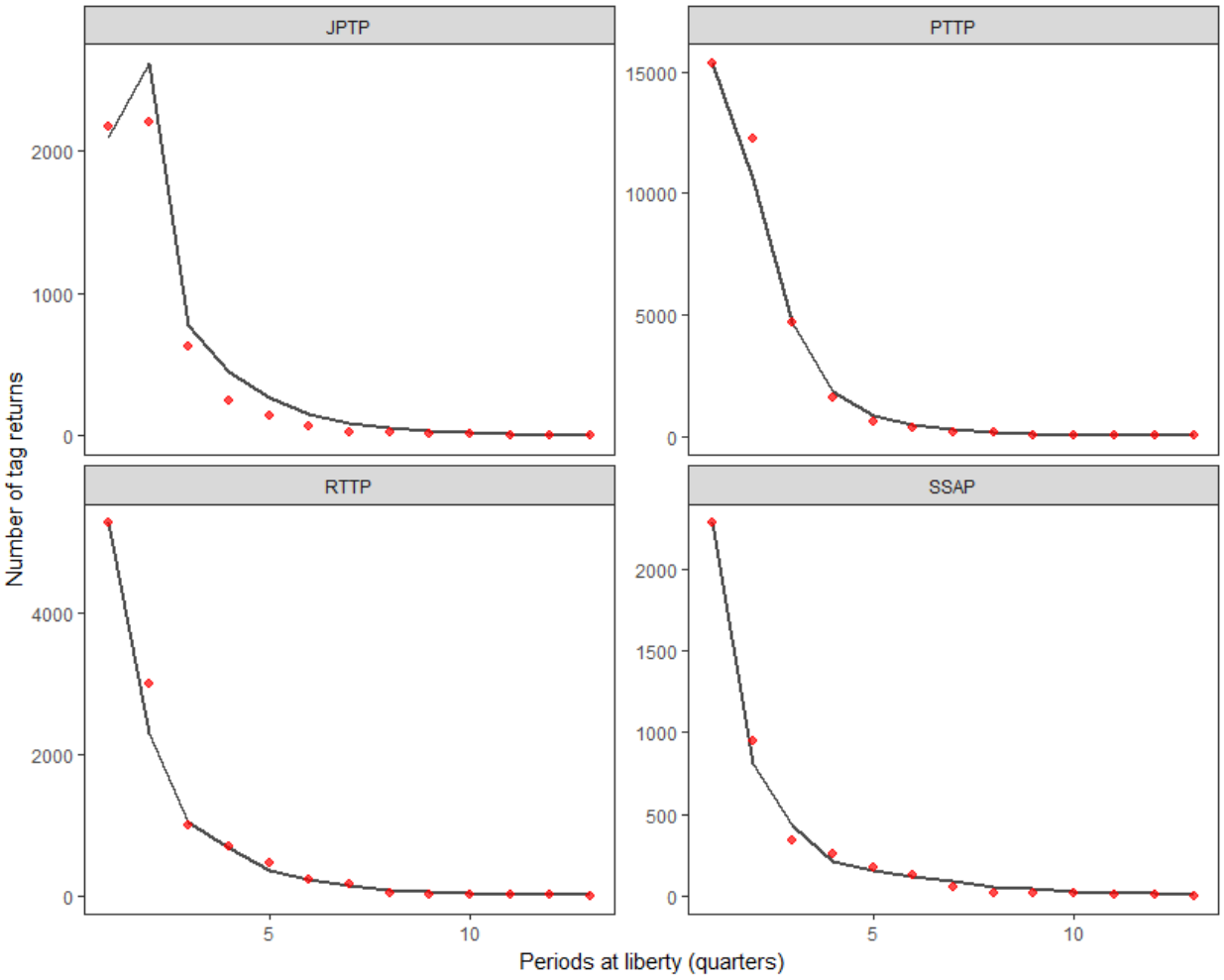


Figure 29: Observed (red points) and model-predicted (black line) tag attrition by tagging program for the diagnostic model. The exact correspondence between observed and predicted values for the first quarter-at-liberty is a direct result of the mixing assumption.

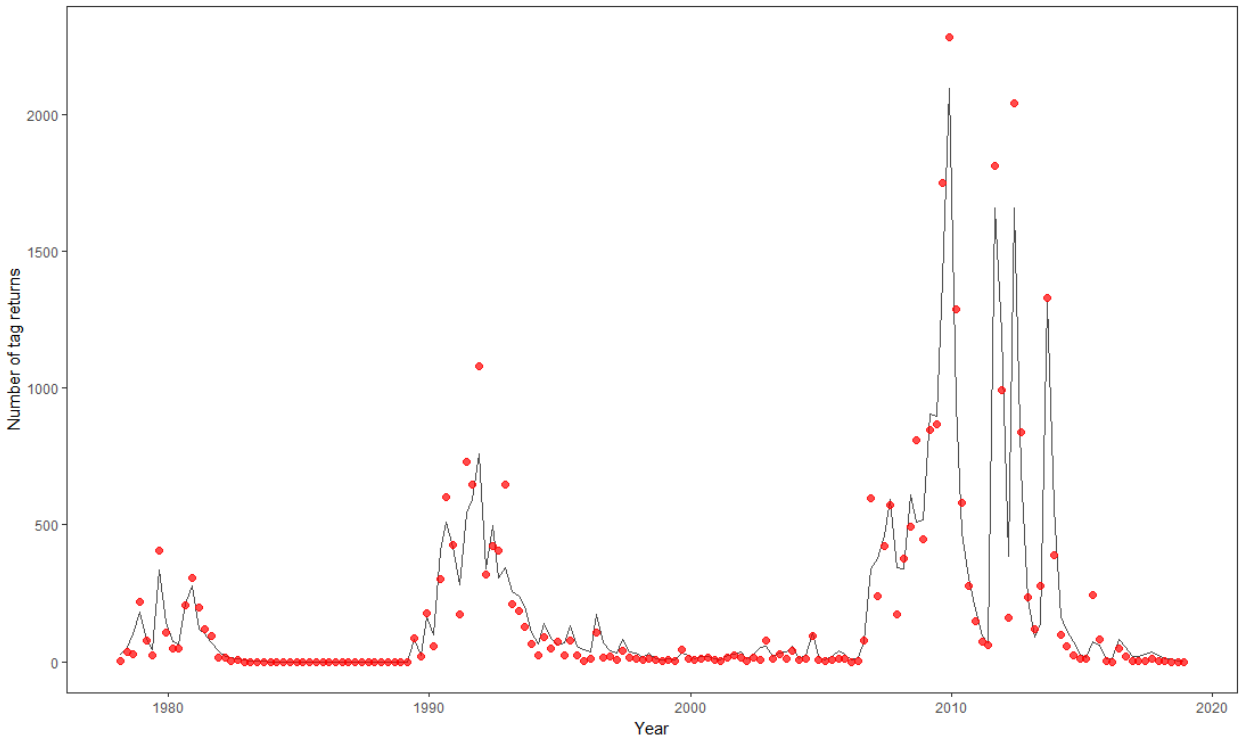


Figure 30: Observed (red points) and model-predicted (black line) tag returns over time, with returns in the mixing period removed, for the diagnostic model across all tag release events with all tag recapture groupings aggregated.

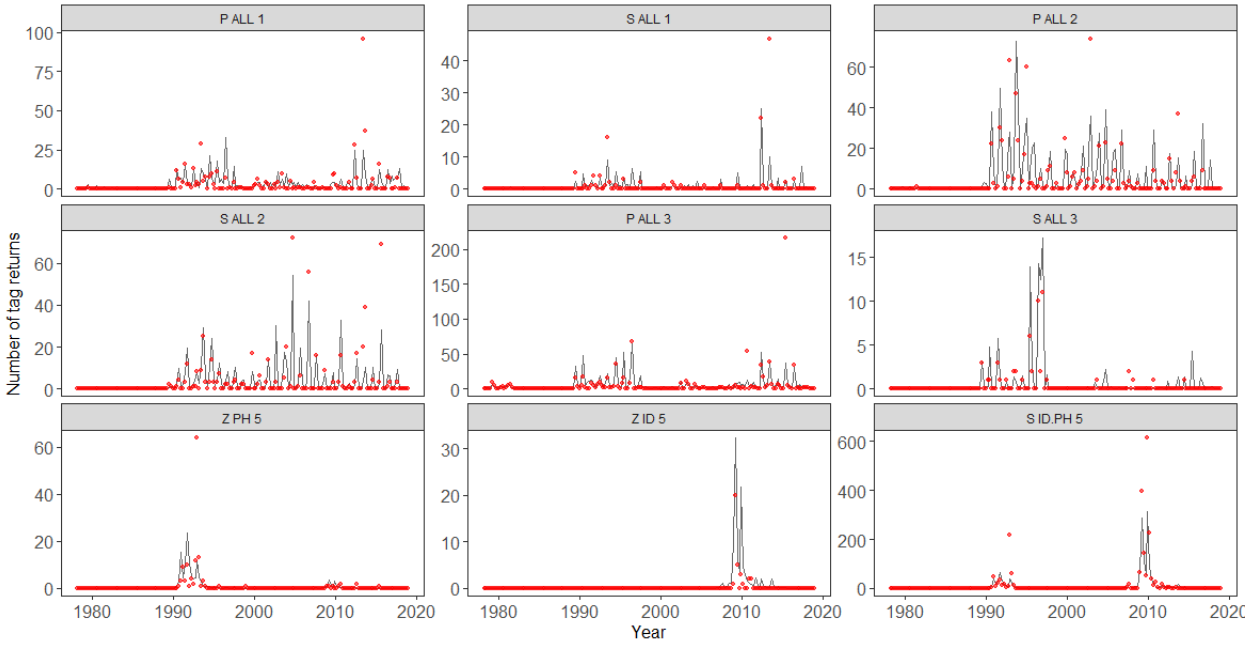


Figure 31: Observed (red points) and model-predicted (black line) tag returns, with returns in the mixing period removed, in the diagnostic model across all tag release events for the first set of fisheries. Groups with extremely low numbers of recaptures are uninformative and not shown.

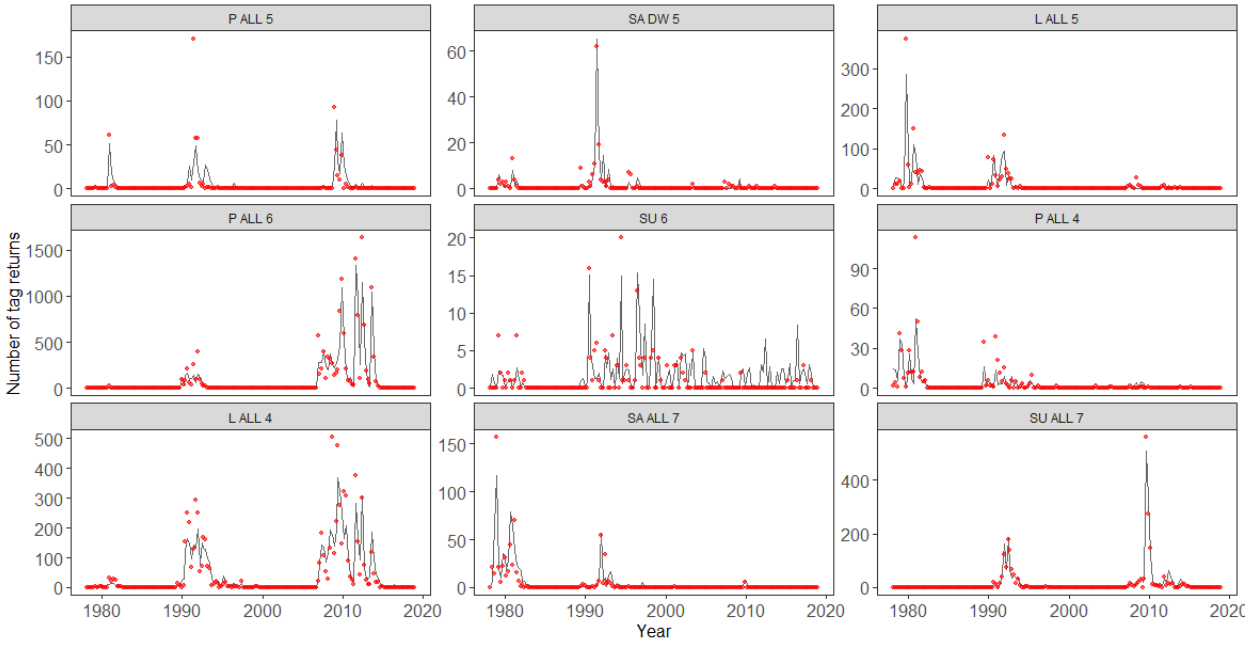


Figure 32: Observed (red points) and model-predicted (black line) tag returns, with returns in the mixing period removed, in the diagnostic model across all tag release events for the second set of fisheries. Groups with extremely low numbers of recaptures are uninformative and not shown.

(a) Observed tag returns

(b) Model predicted tag returns

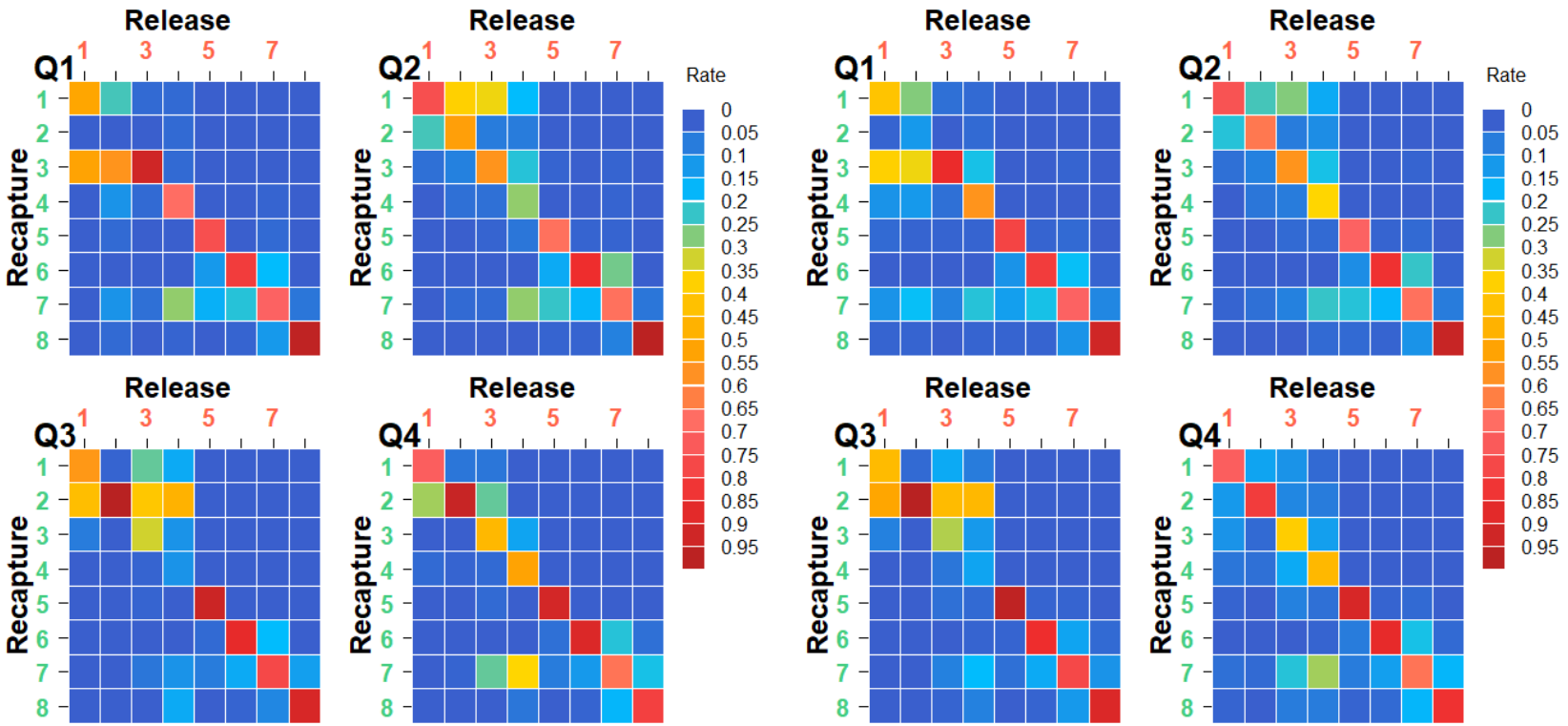


Figure 33: Comparison of the observed (33a) and predicted (33b) proportion of tags returned by region of release (columns), region of recapture (rows), and quarter of recapture (panel) where the color of the tile indicates the proportion of tags returned from the region of release.

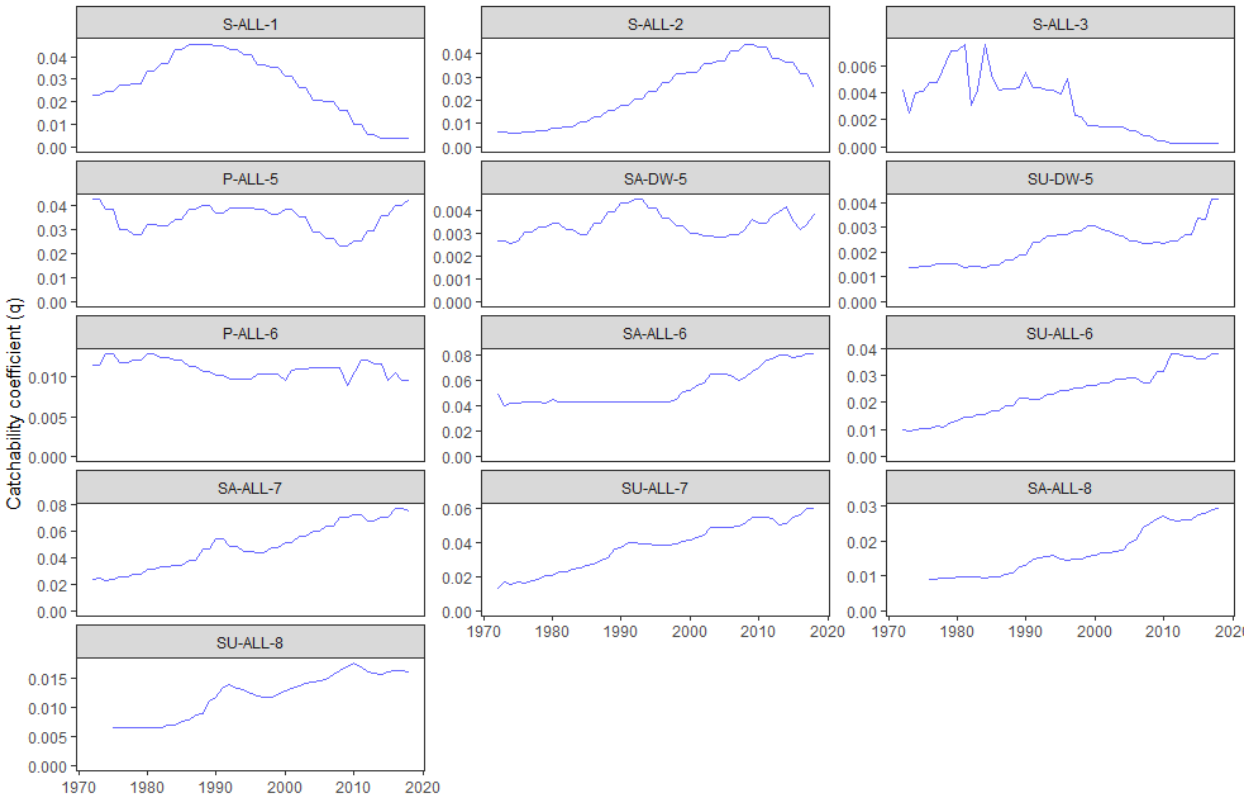


Figure 34: Estimated time series of catchability (including seasonal variability) for those fisheries assumed to have random walk in these parameters.

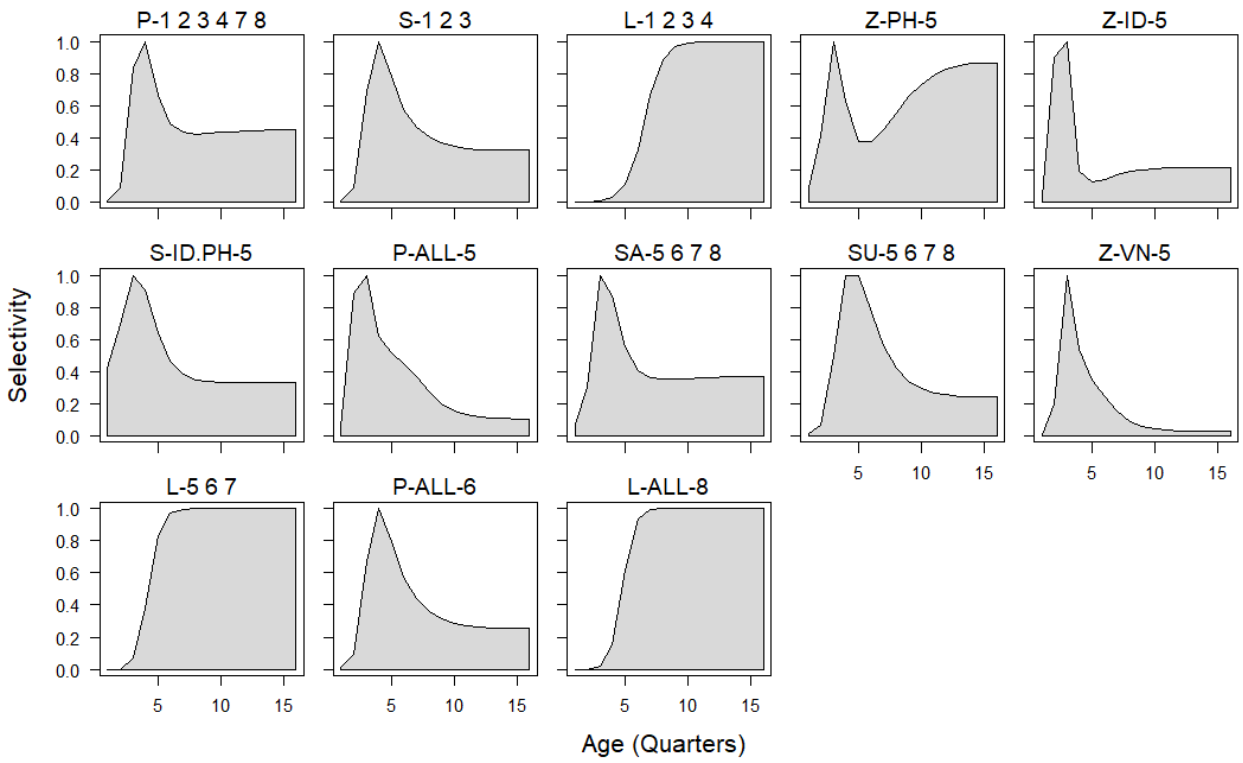


Figure 35: Selectivity-at-age in number of quarters by groups of fisheries with shared selectivities.

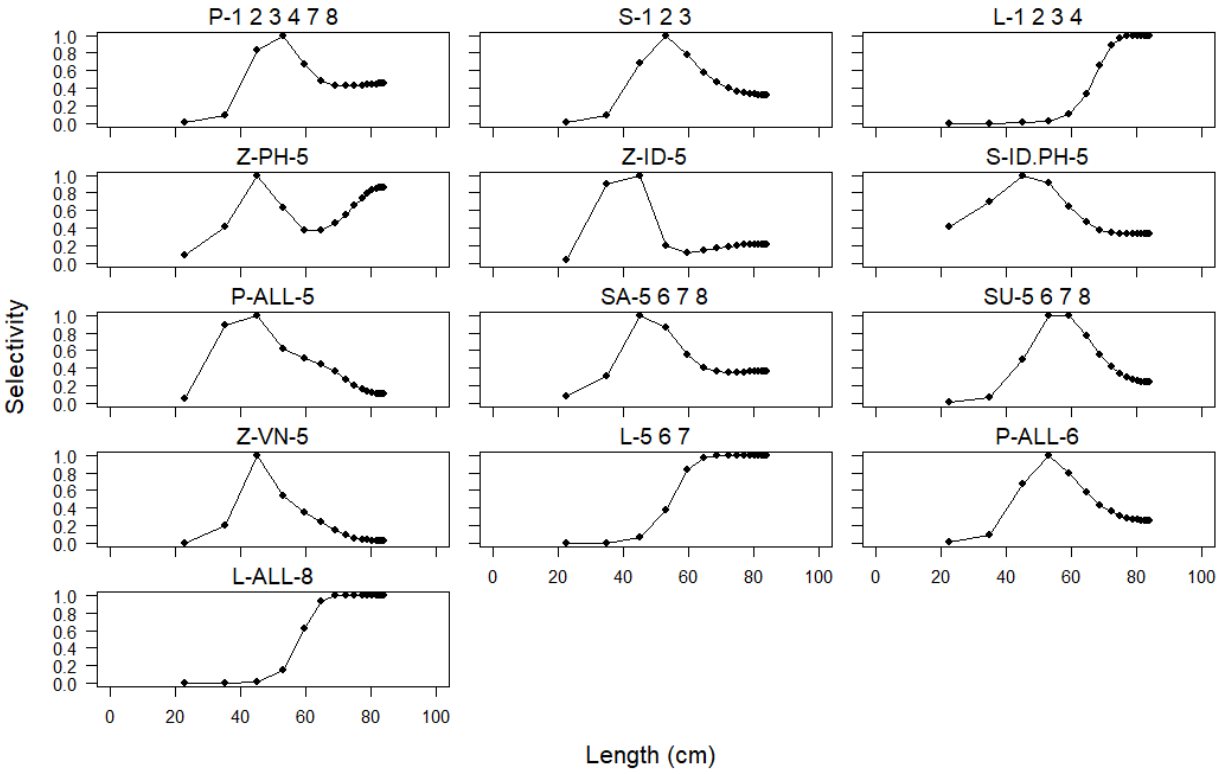


Figure 36: Length-specific selectivity coefficients by groups of fisheries with shared curves, where the points indicate the mean length at age.

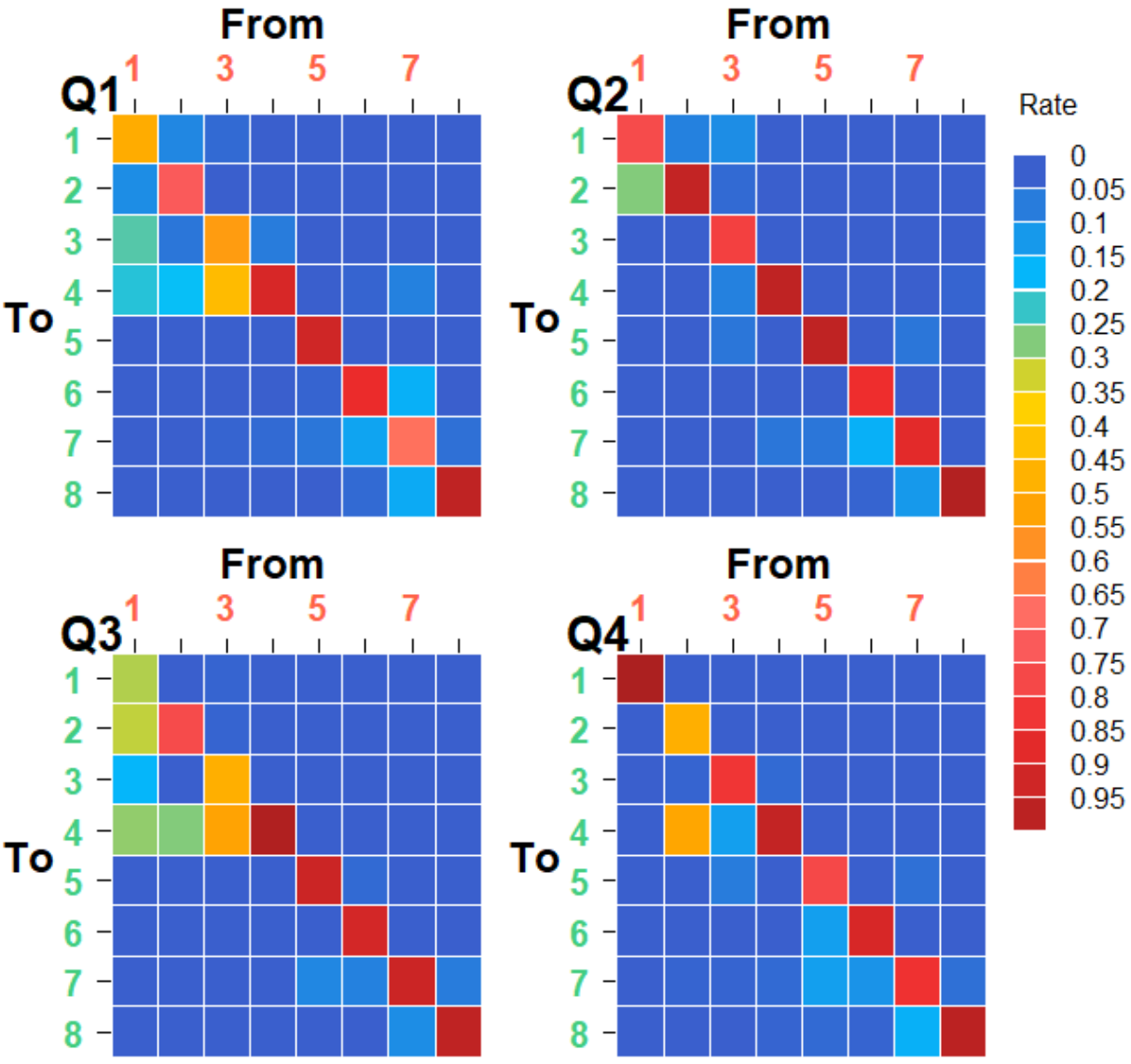
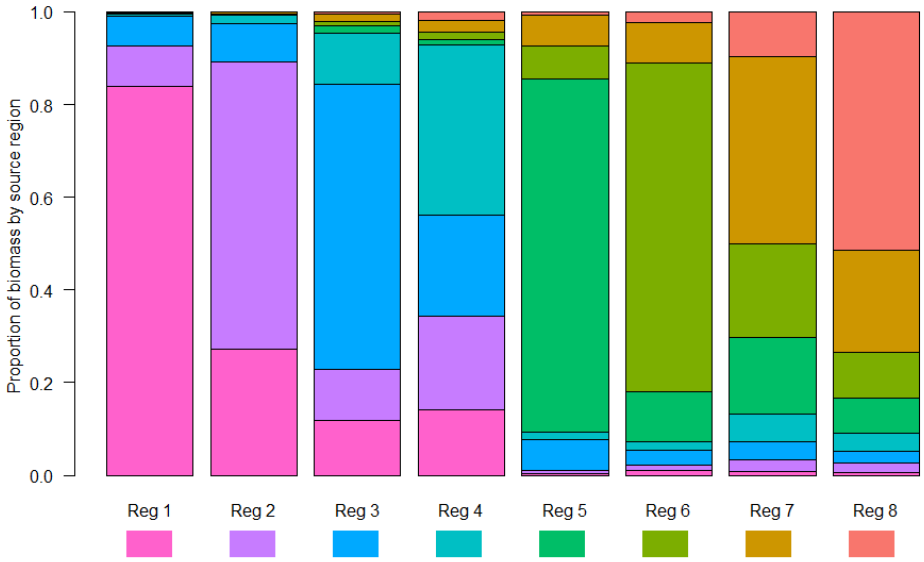


Figure 37: Estimated movement coefficients by quarter for the diagnostic case model. The red numbers (horizontal axis) indicate the source model region; the green numbers (vertical axis) indicate the receiving regions. The color of the tile shows the magnitude of the movement rate (proportion of individuals from region x moving to region y in that quarter), with each column adding up to 1.

(a) No F



(b) F average 2008–2017

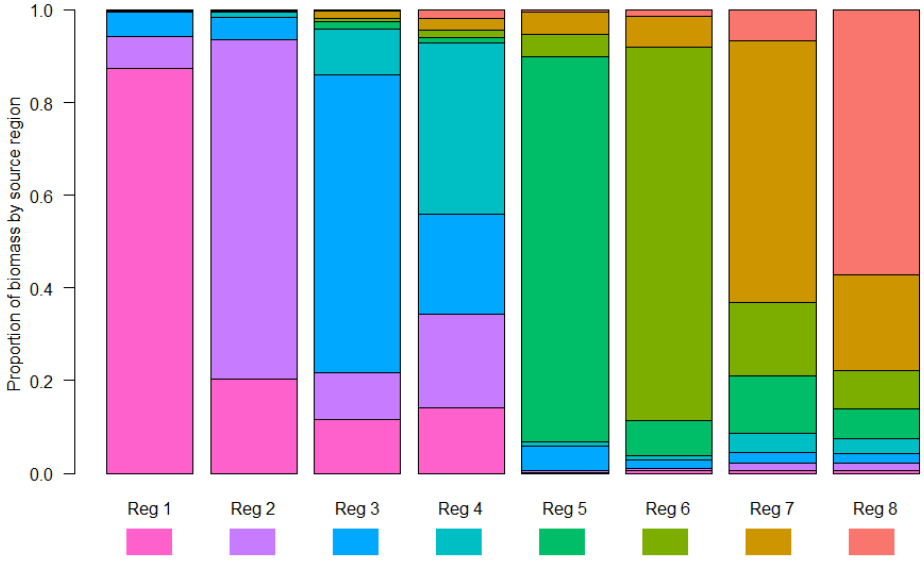


Figure 38: The expected percent contribution of total biomass (by weight) from each region based on the long-term average distribution of recruitment between regions, estimated movement parameters, and natural mortality. The columns represent the average total biomass within a region after mixing and the colors are the region of recruitment presented below the corresponding label on the x-axis. In panel 38a fishing mortality is not included in the calculations, whereas panel 38b includes the regional fishing mortality at age averaged across 2008–2017.

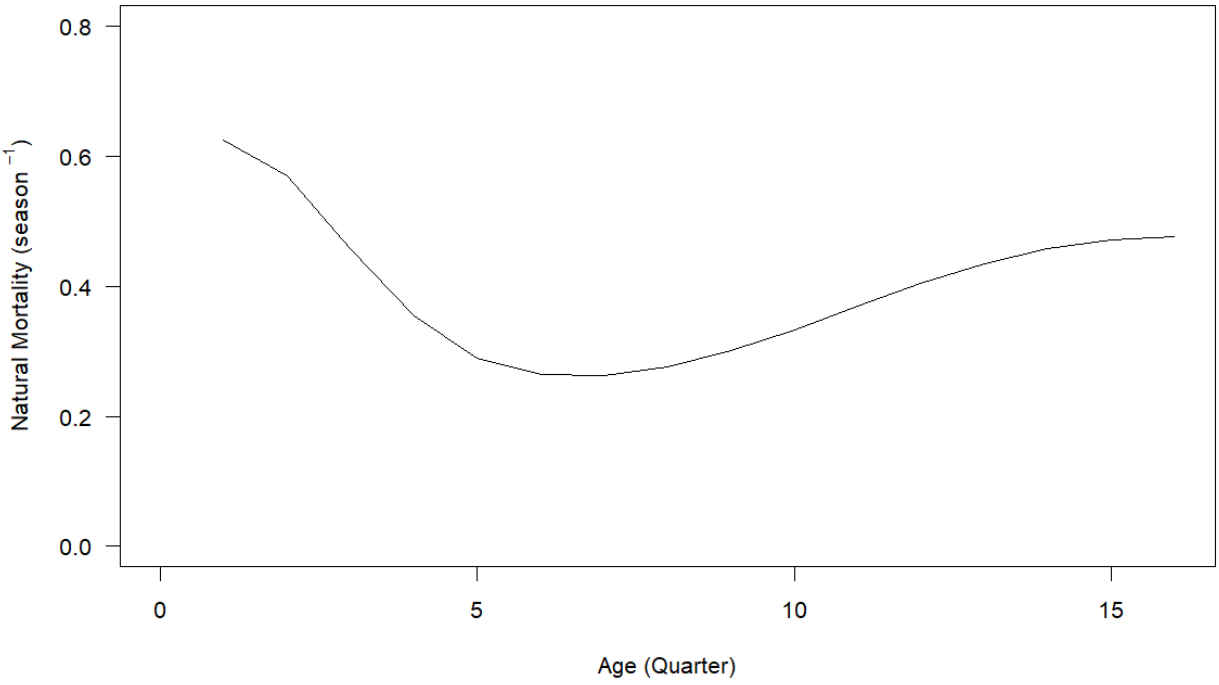


Figure 39: Quarterly natural mortality-at-age as estimated by the diagnostic model.

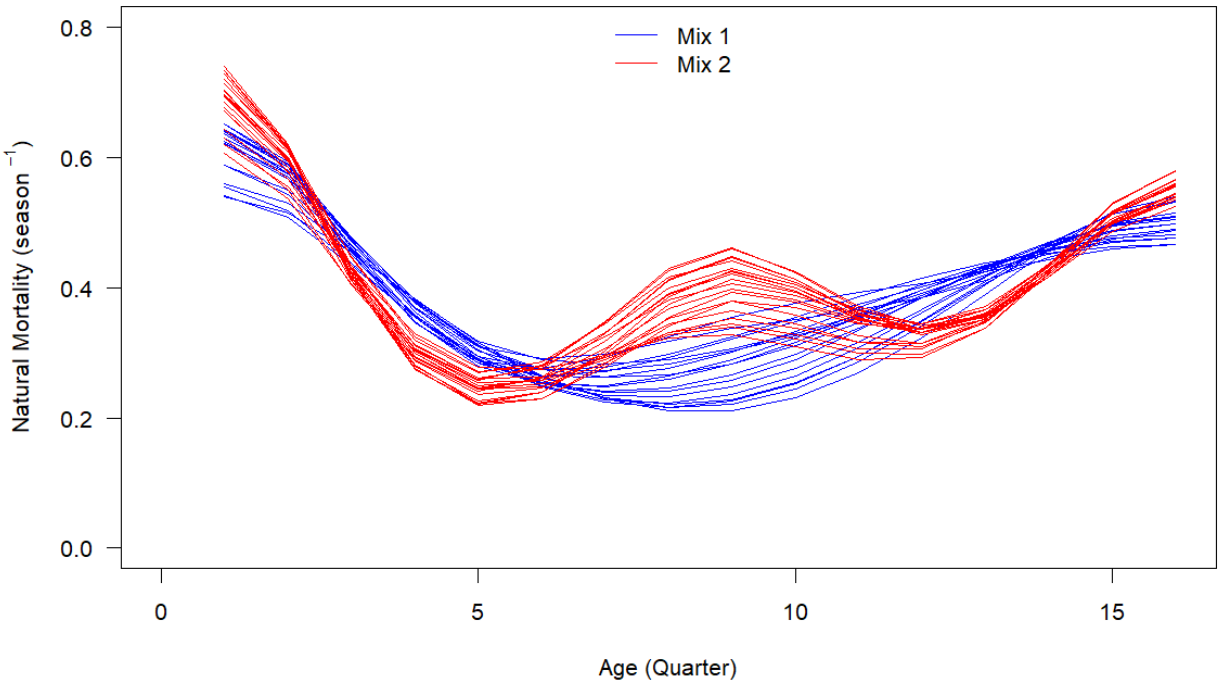


Figure 40: Quarterly natural mortality-at-age as estimated by all models in the structural uncertainty grid by number of mixing periods assumed in the model.

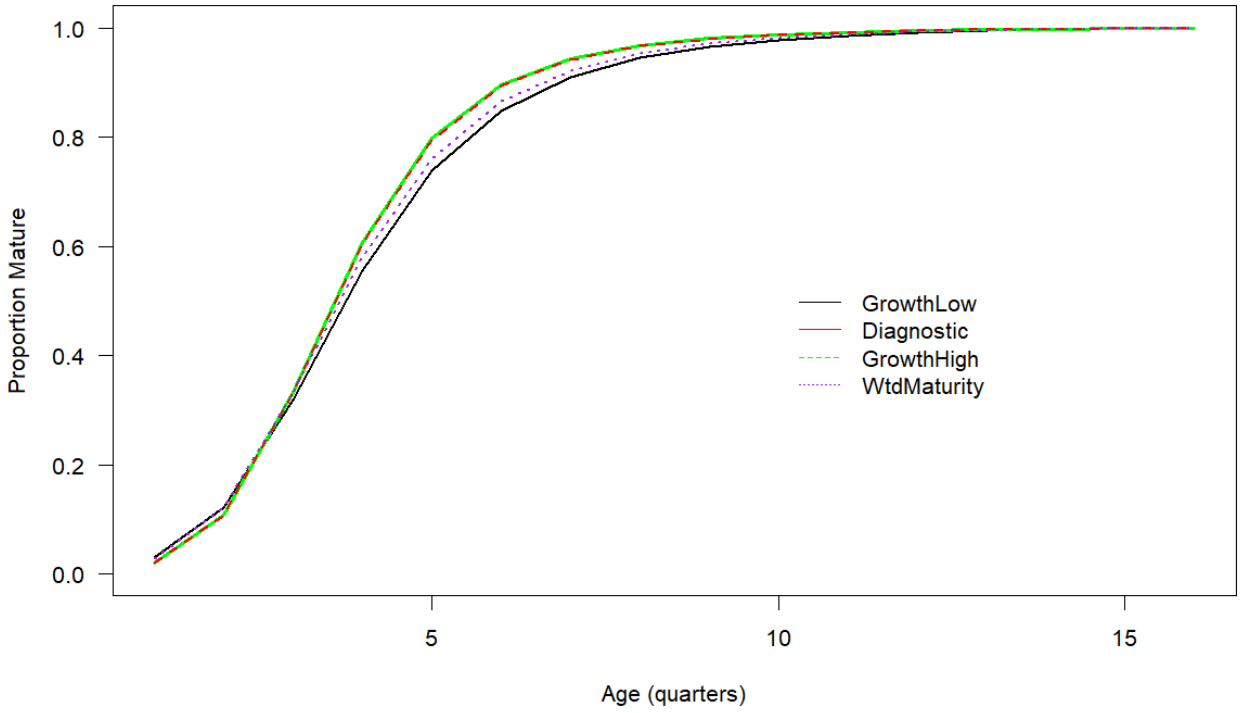


Figure 41: Maturity-at-age as estimated in the models that assumed diagnostic growth (red line), *GrowthLow* (black line), *GrowthHigh* (green line) and diagnostic growth with the regionally weighted maturity-at-length (*WtdMaturity*).

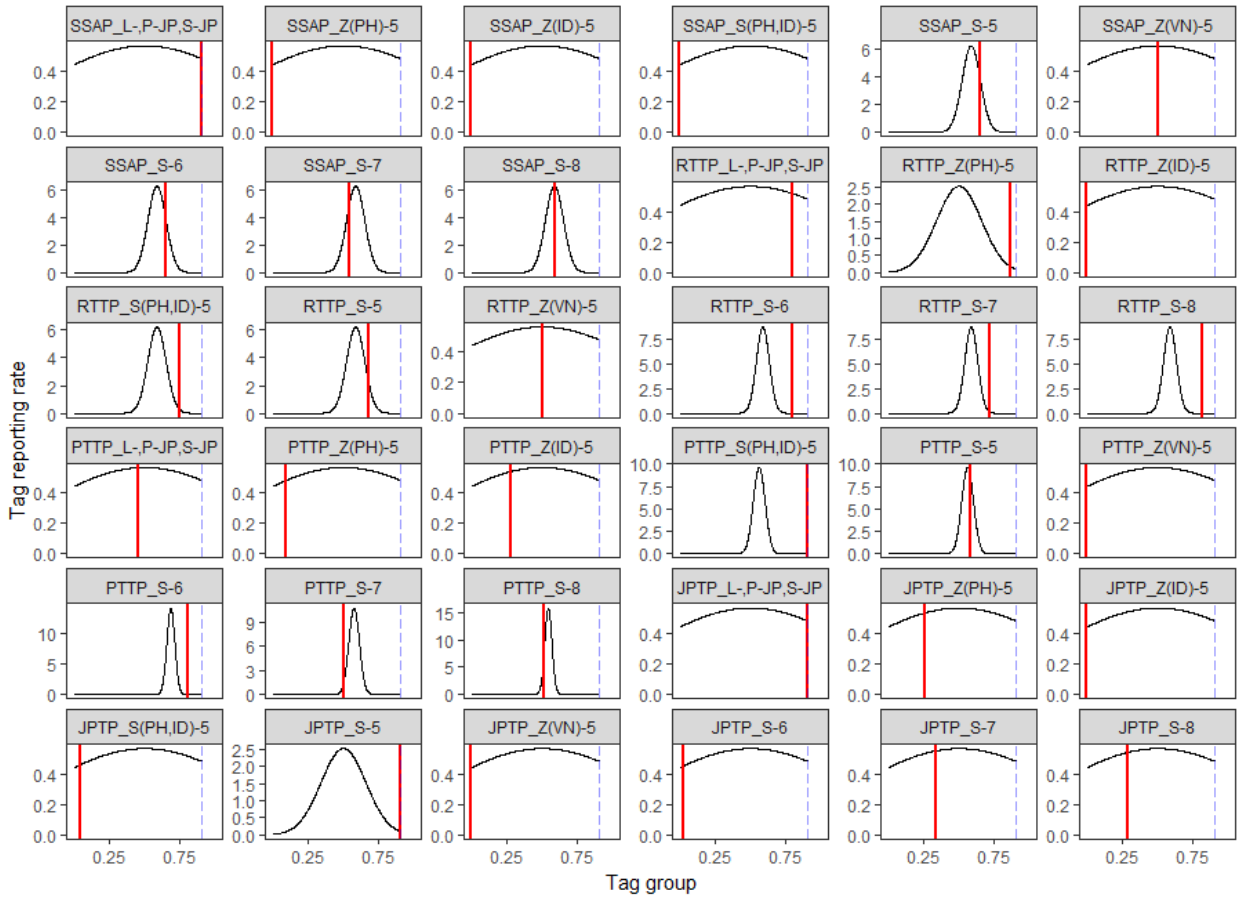


Figure 42: Estimated reporting rates for the diagnostic model (red lines) and the prior distribution for each reporting rate group. The imposed upper bound (0.9) on the reporting rate parameters is shown as a blue dashed line. Reporting rates can be estimated separately for each release program and recapture fishery group but in practice are aggregated over some recapture groups to reduce dimensionality.

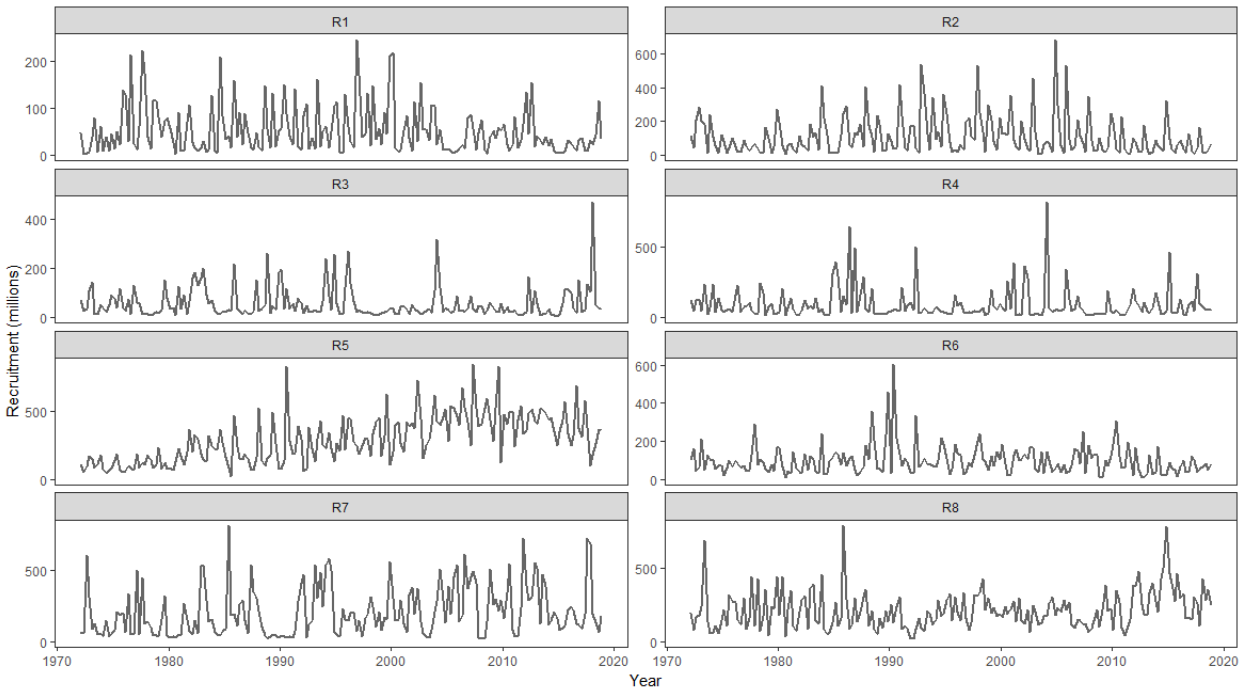


Figure 43: Estimated temporal recruitment by model region from the diagnostic model. Note that the scale of the y-axis is not constant across regions.

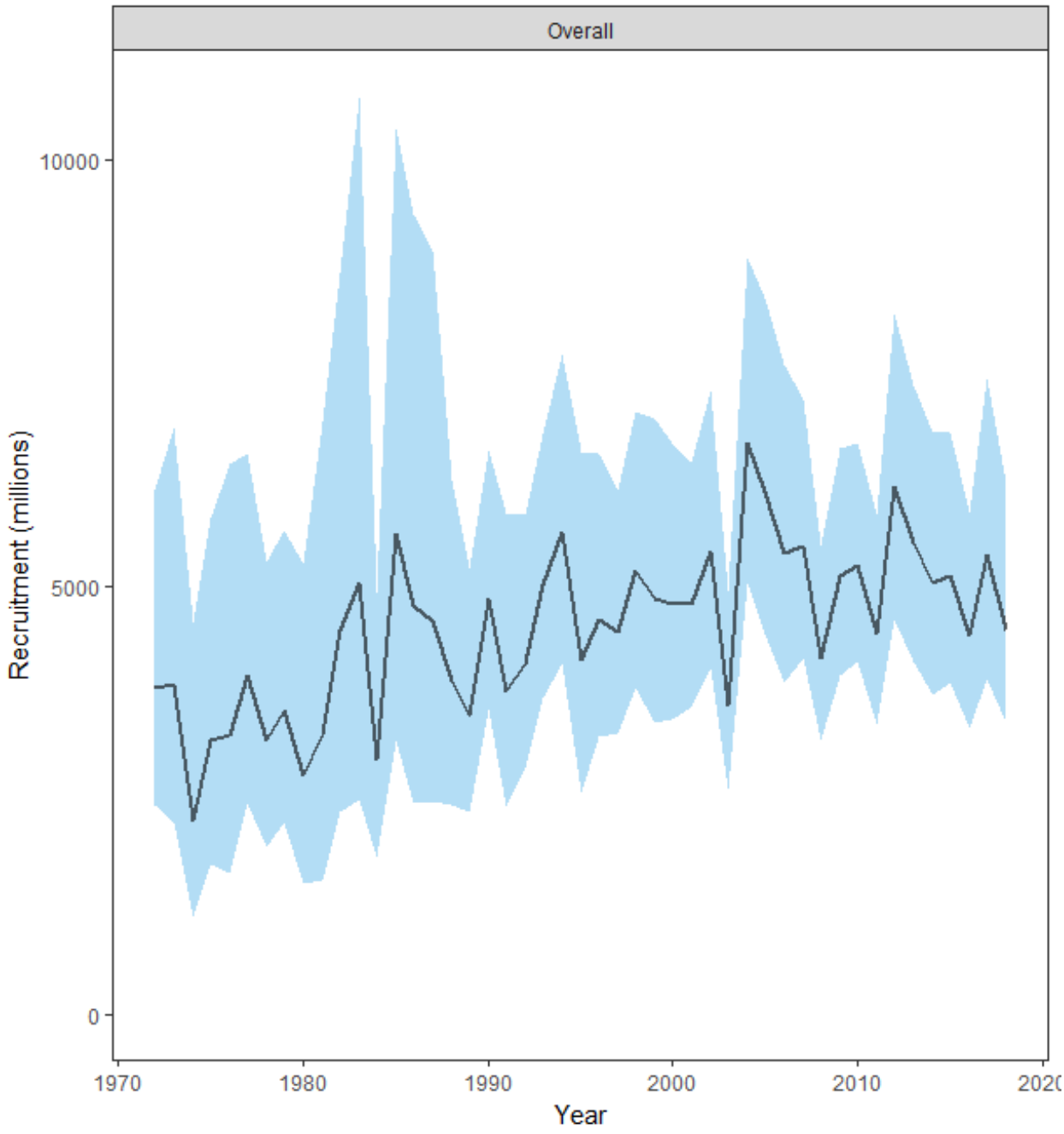
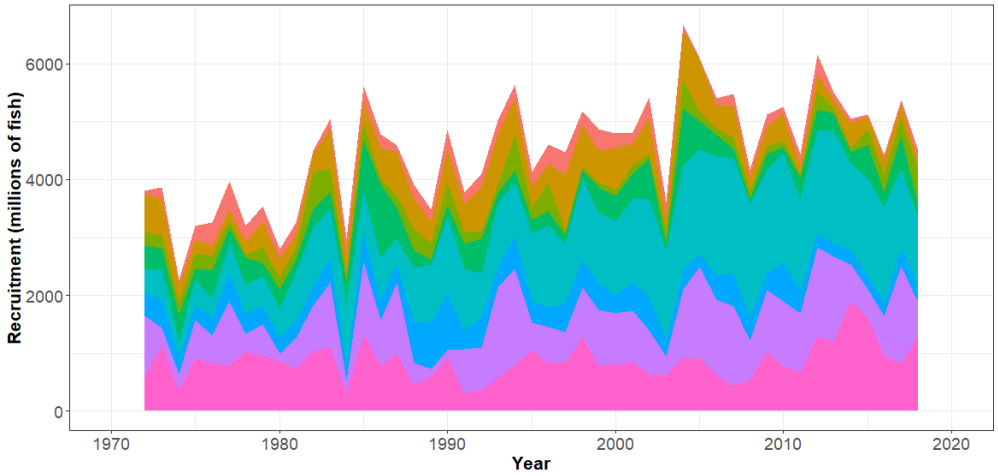
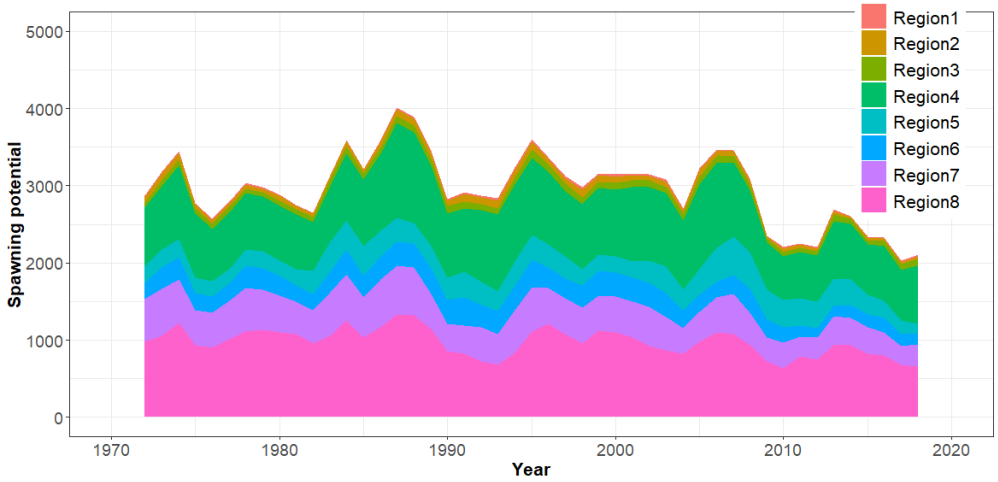


Figure 44: Estimated overall annual recruitment summed across regions where the shaded region is ± 2 standard deviations for the diagnostic model (∼ 95% CI).

(a) Regional recruitment



(b) Spawning potential



(c) Total biomass

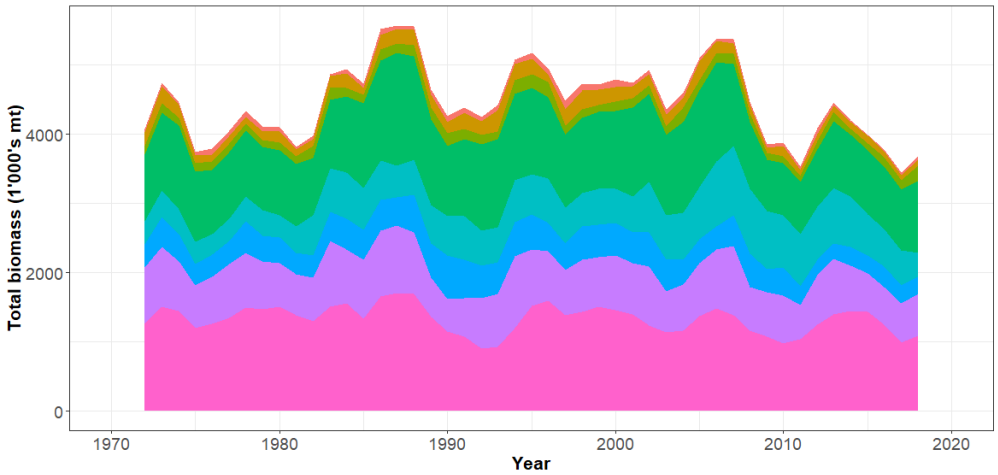


Figure 45: Estimated annual average recruitment, spawning potential and total biomass by model region for the diagnostic model, showing the relative sizes among regions.

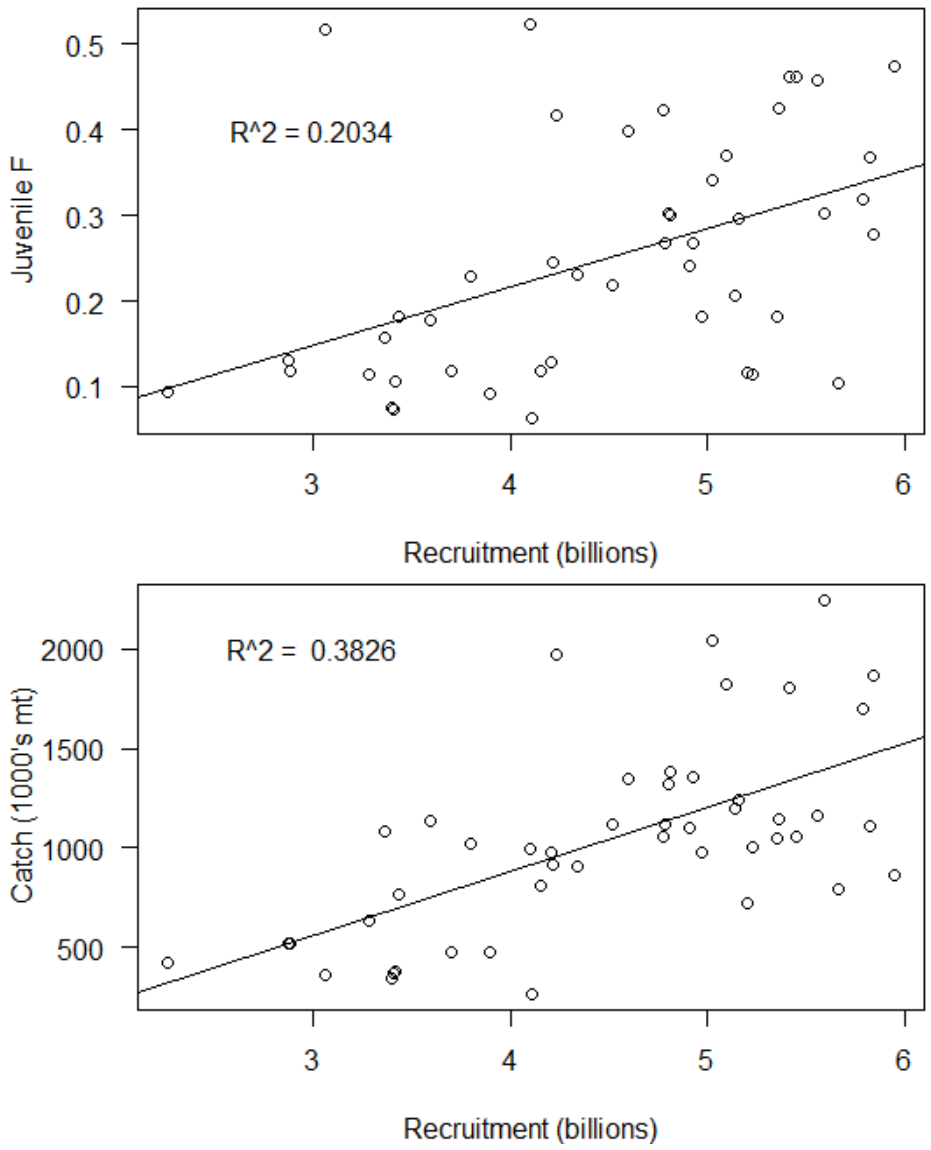


Figure 46: Correlation of estimated recruitment with catch and juvenile fishing mortality for the diagnostic model.

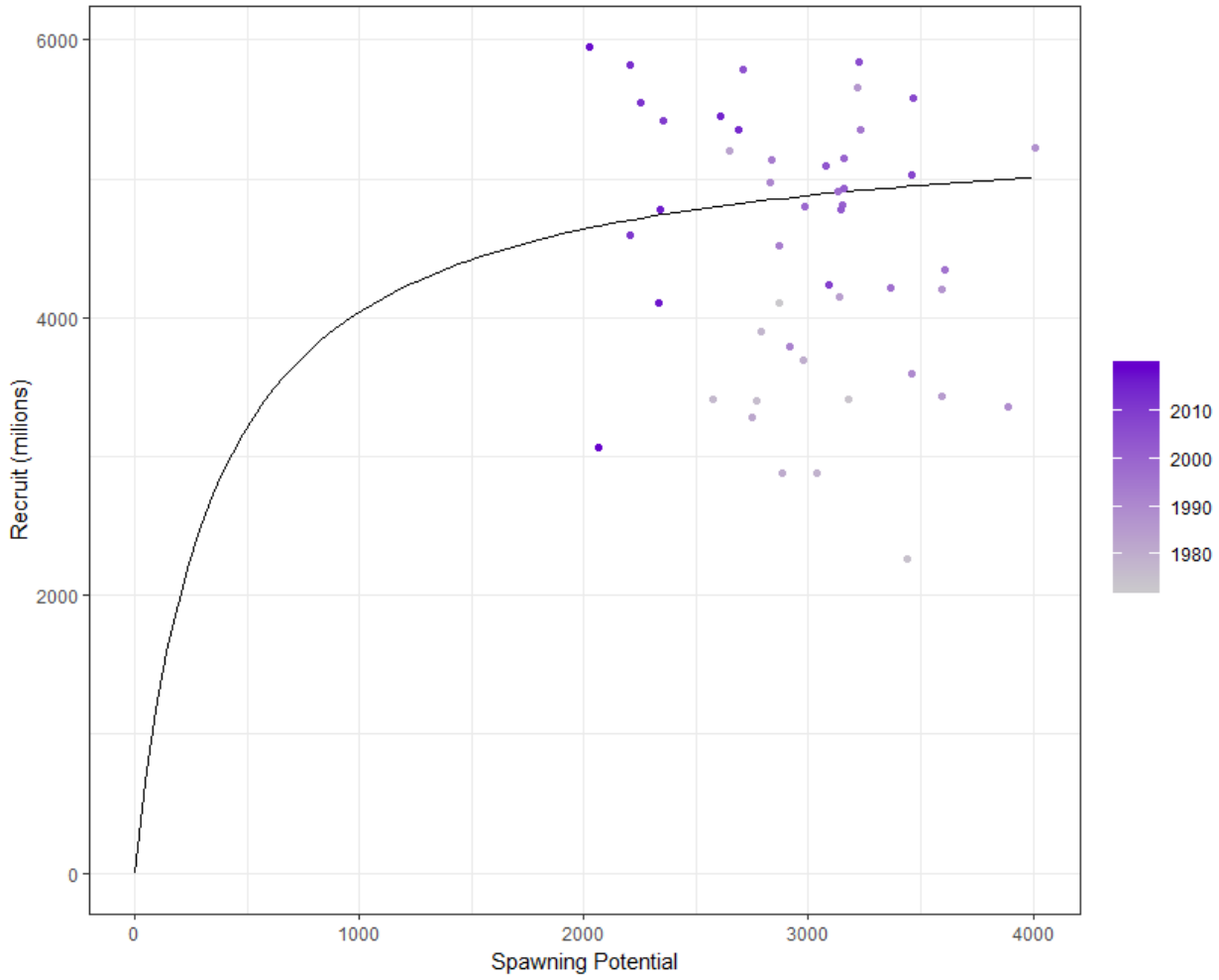


Figure 47: Estimated relationship between recruitment and spawning potential based on annual values for the diagnostic model. The darkness of the circles changes from light to dark through time.

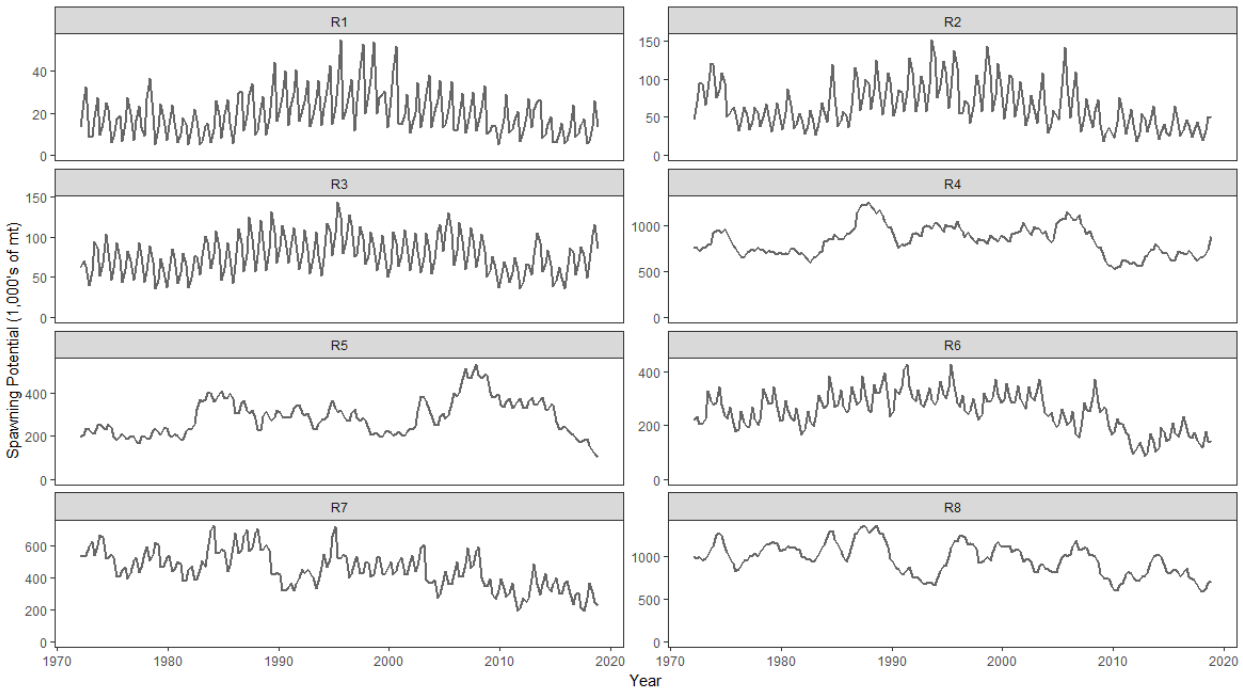


Figure 48: Estimated temporal spawning potential by model region for the diagnostic model. Note that the scale of the y-axis is not constant across regions.

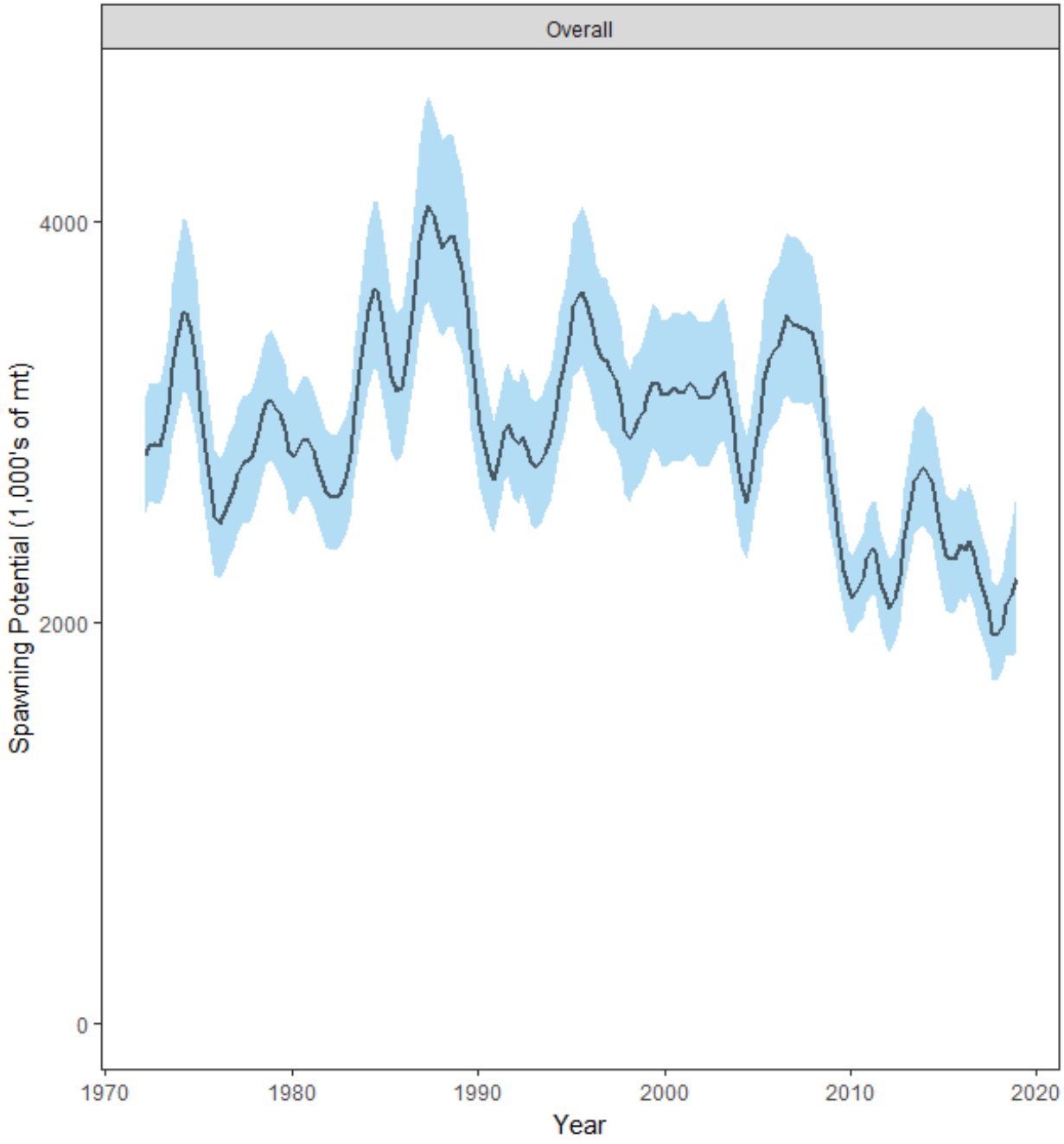


Figure 49: Estimated temporal overall spawning potential summed across regions from the diagnostic model, where the shaded region is ± 2 standard deviations (i.e., ∼ 95% CI).

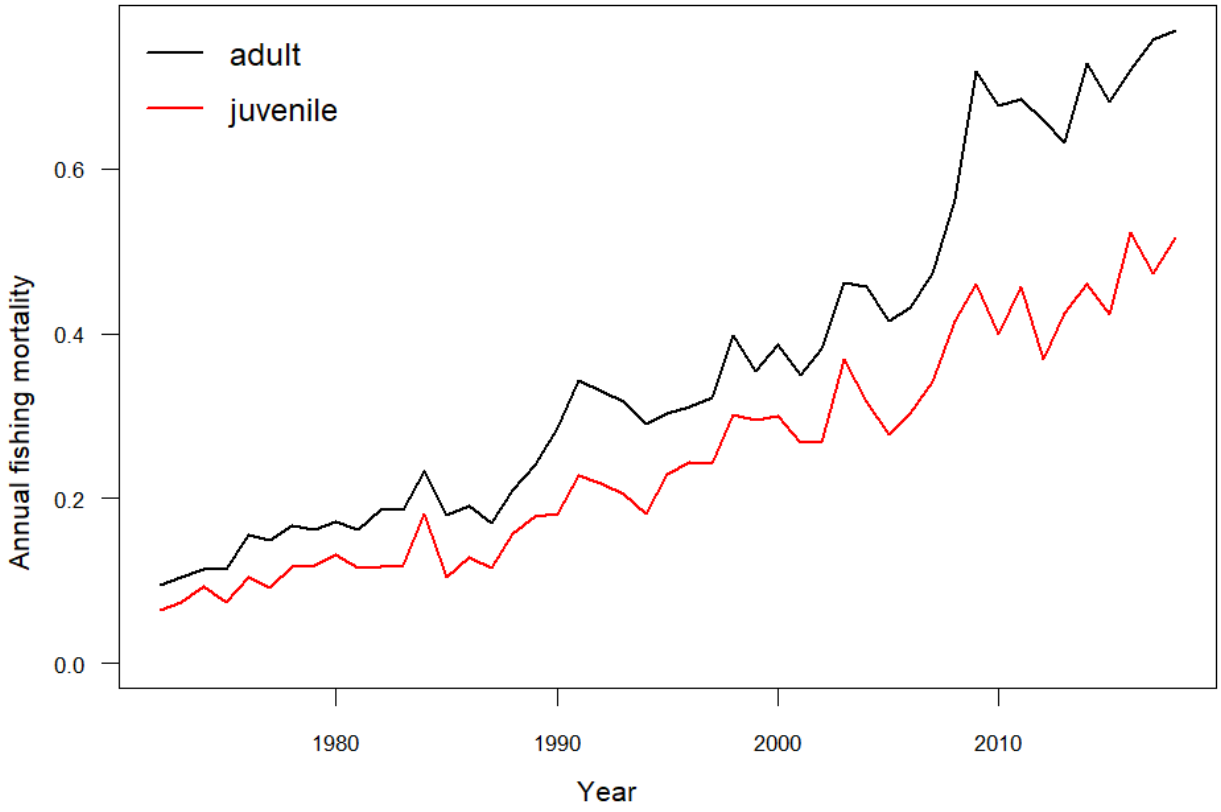


Figure 50: Estimated annual average juvenile and adult fishing mortality for the diagnostic model.

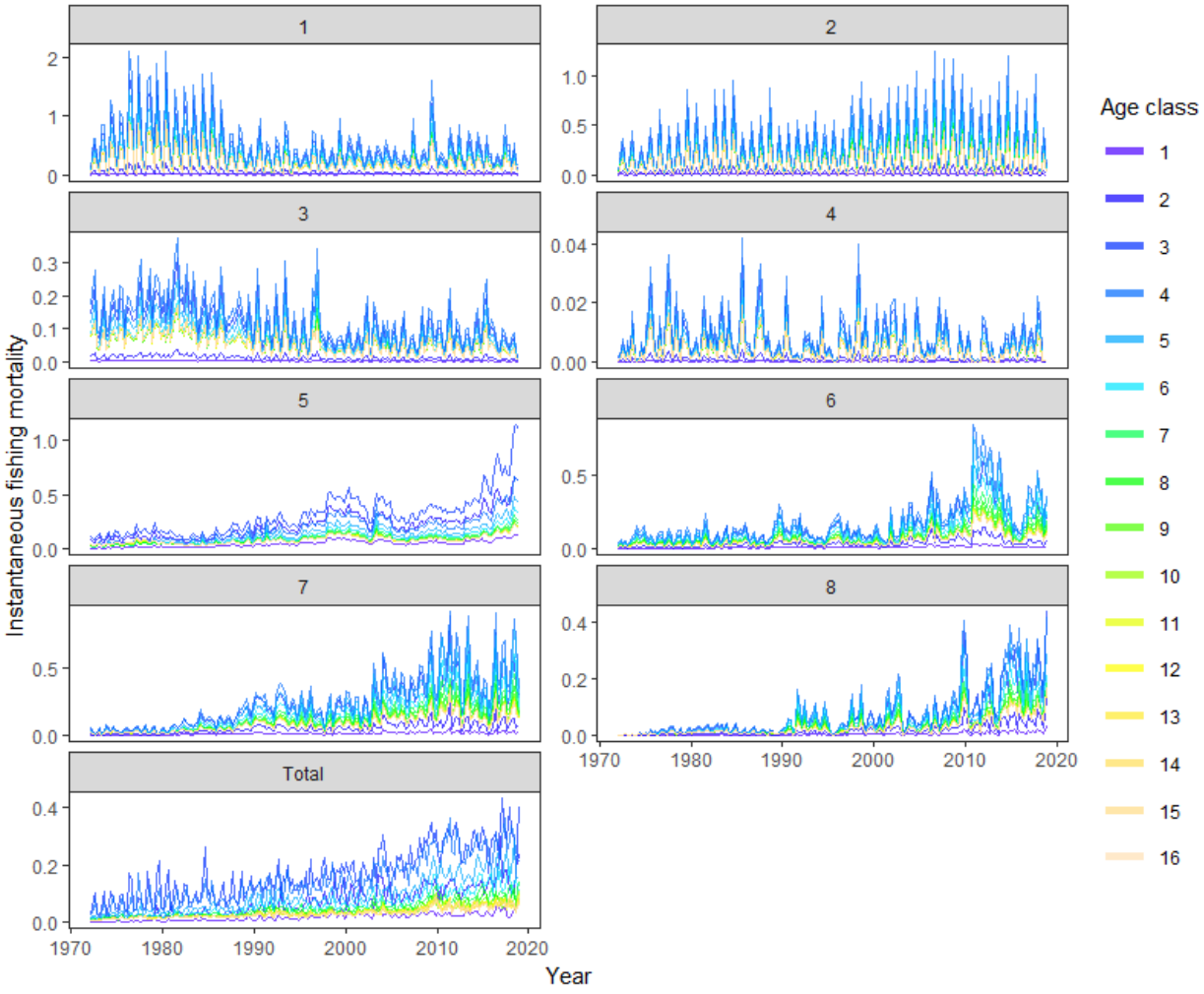


Figure 51: Estimated age-specific fishing mortality for the diagnostic model, by region and overall.

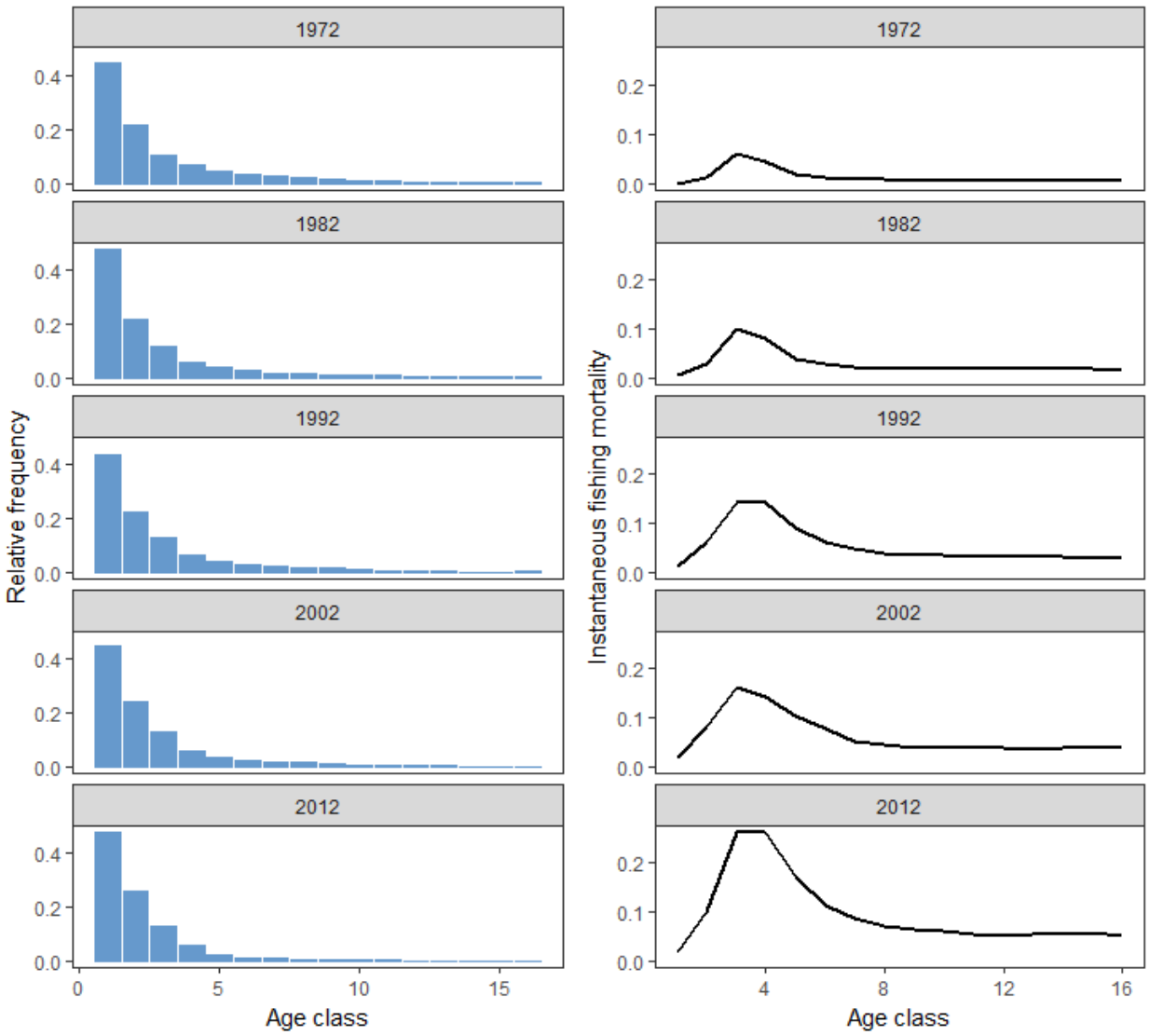


Figure 52: Estimated proportion at age (quarters) and fishing mortality at age (right), by year, at decade intervals, for the diagnostic model.

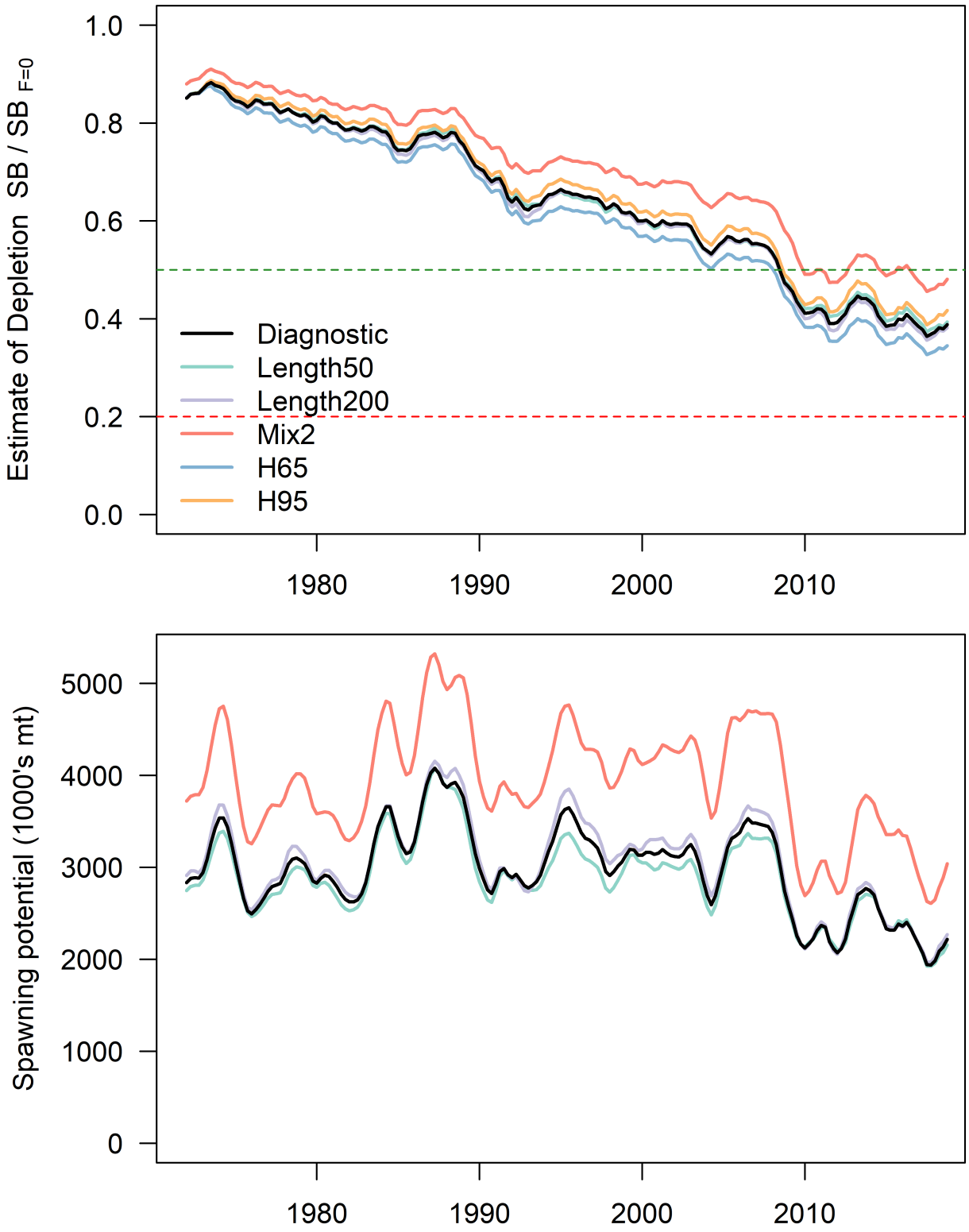


Figure 53: Estimated dynamic depletion of spawning potential (SB) in time t relative to the unfished SB estimated in time t (top panel) and the estimated SB for the first set of one-off sensitivities from the 2019 diagnostic case.

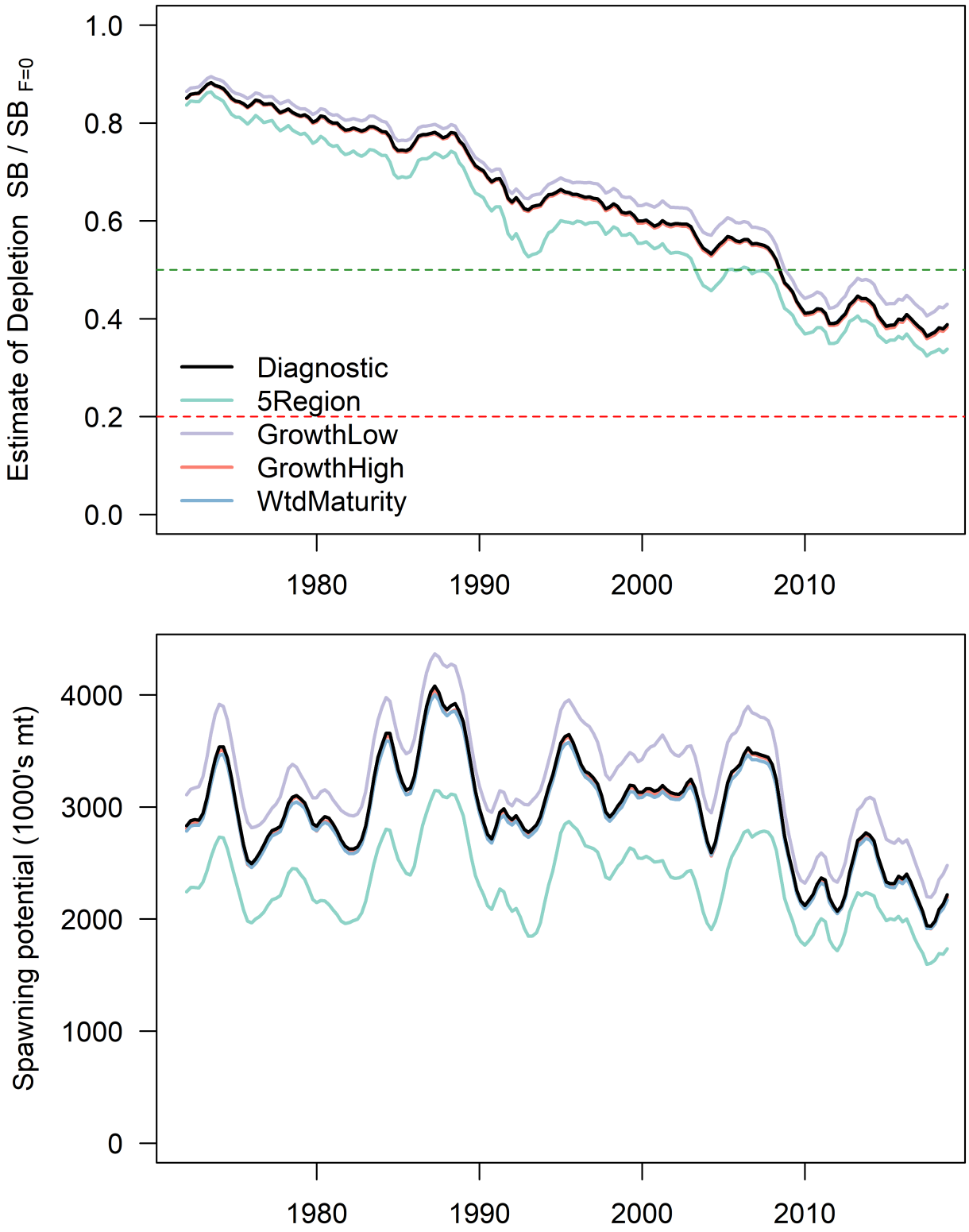


Figure 54: Estimated dynamic depletion of spawning potential (SB) in time t relative to the unfished SB estimated in time t (top panel) and the estimated SB for the second set of one-off sensitivities from the 2019 diagnostic case.

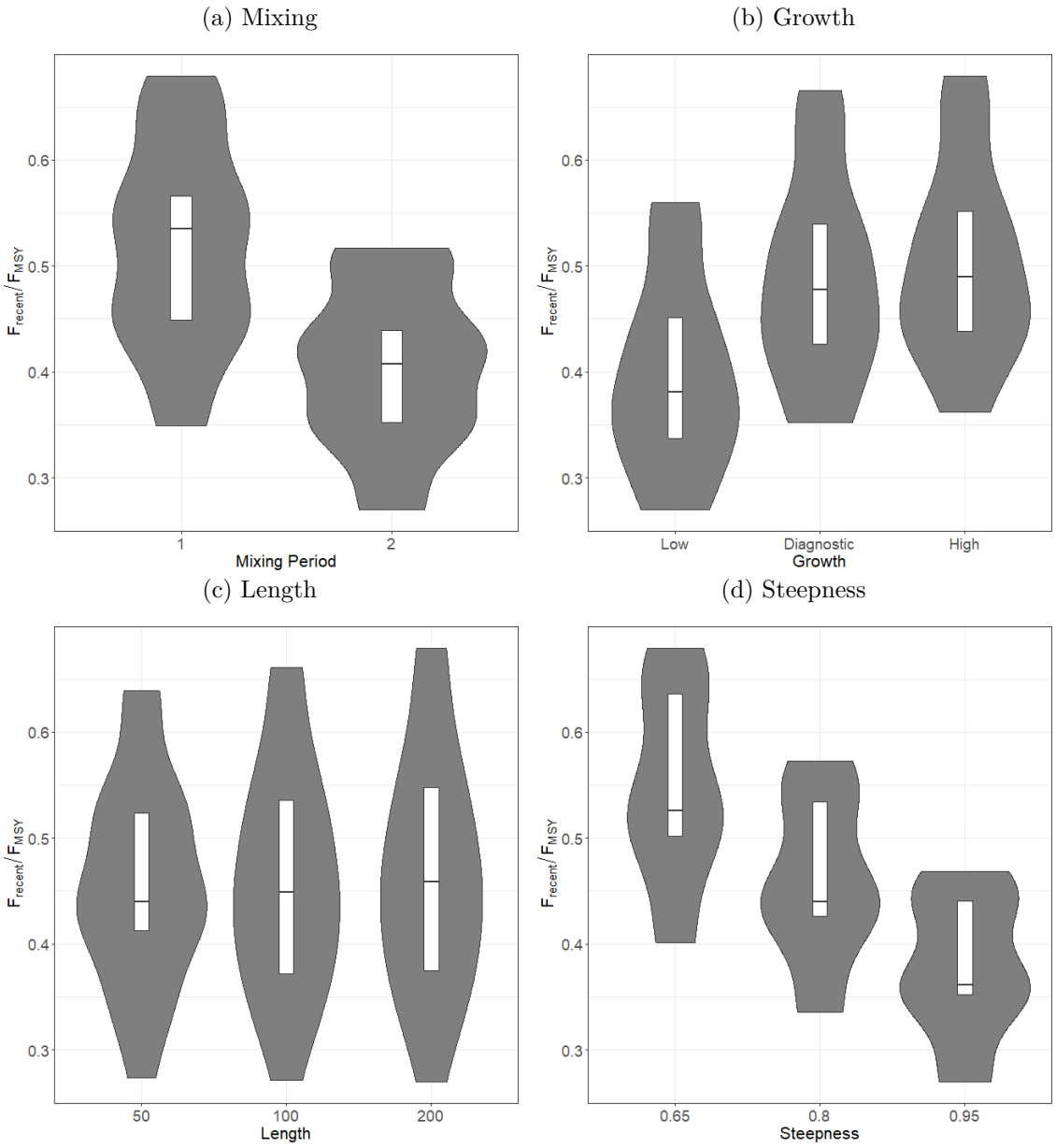


Figure 55: Box and violin plots summarizing the estimated $F_{recent}/F_{\text{MSY}}$ for each of the models in the structural uncertainty grid for the 8 region model. The line in the box is the median of the estimates, while the box shows the 50^{th} percentile. The shaded area shows the probability distribution (or density) of the estimates of all models across the axes of the structural uncertainty grid.

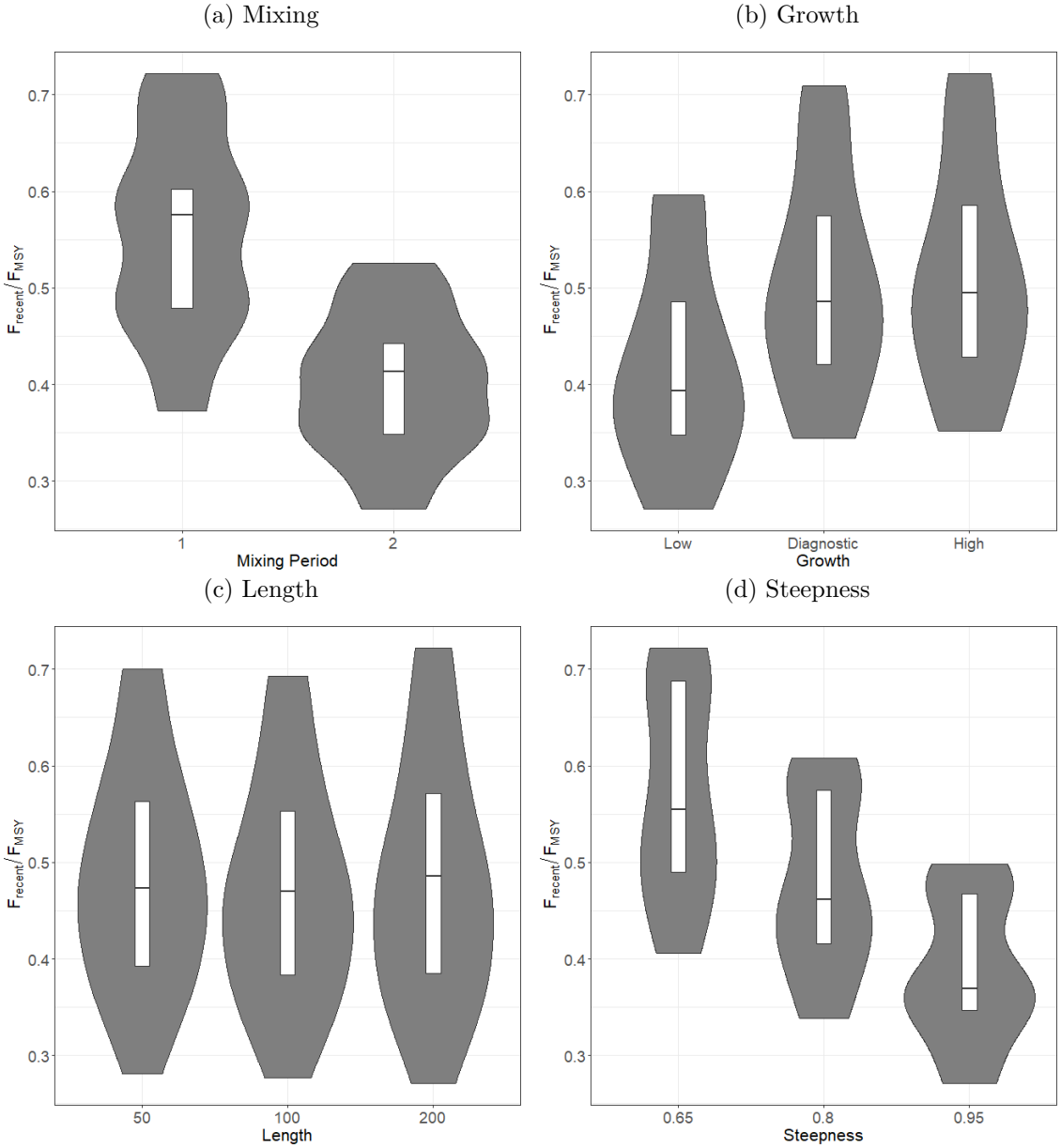


Figure 56: Box and violin plots summarizing the estimated $F_{recent}/F_{\text{MSY}}$ for each of the models in the structural uncertainty grid for the 5 region model. The line in the box is the median of the estimates, while the box shows the 50^{th} percentile. The shaded area shows the probability distribution (or density) of the estimates of all models across the axes of the structural uncertainty grid.

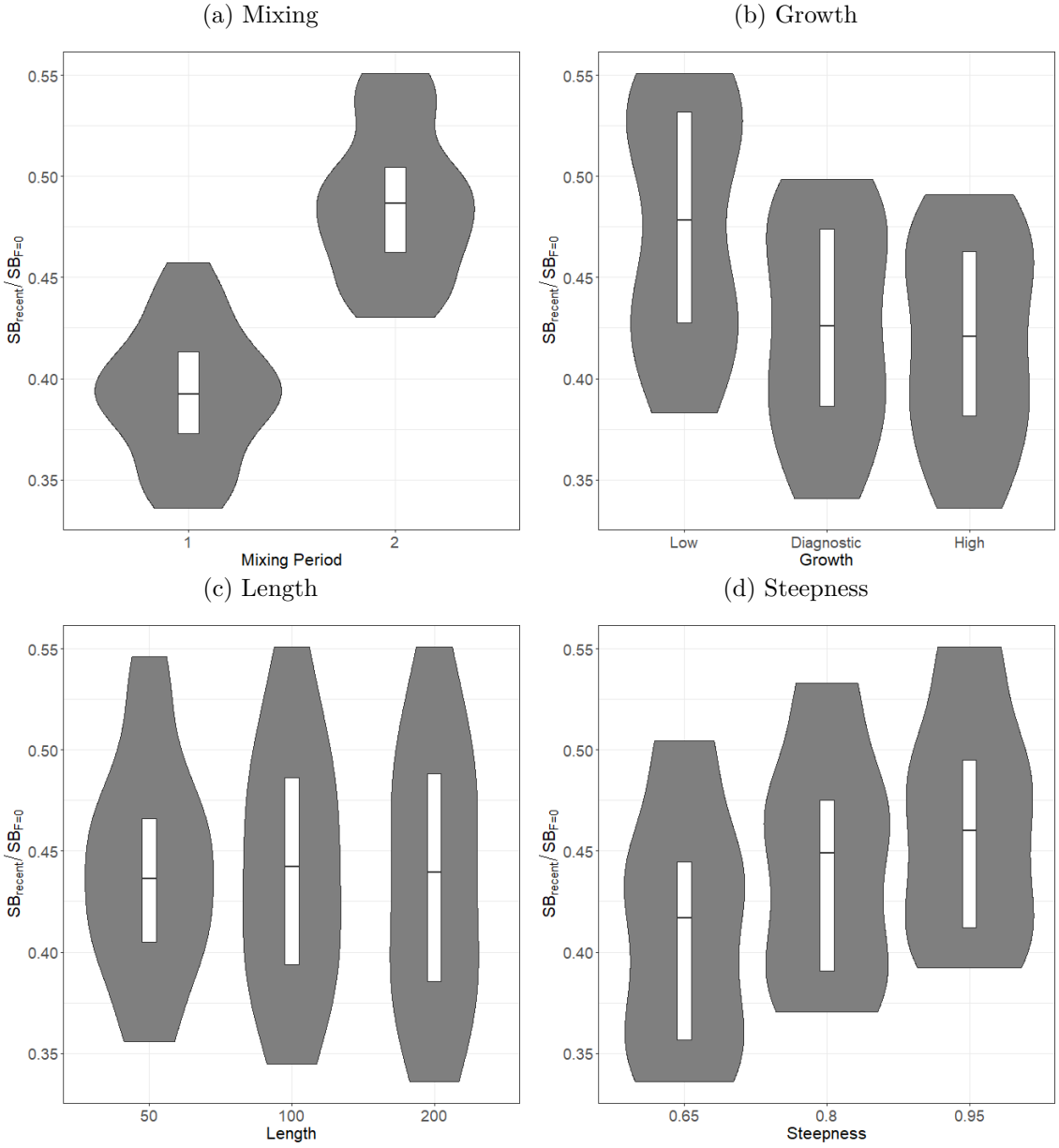


Figure 57: Box and violin plots summarizing the estimated $SB_{recent}/SB_{F=0}$ for each of the models in the structural uncertainty grid for the 8 region model. The line in the box is the median of the estimates, while the box shows the 50^{th} percentile. The shaded area shows the probability distribution (or density) of the estimates of all models across the axes of the structural uncertainty grid.

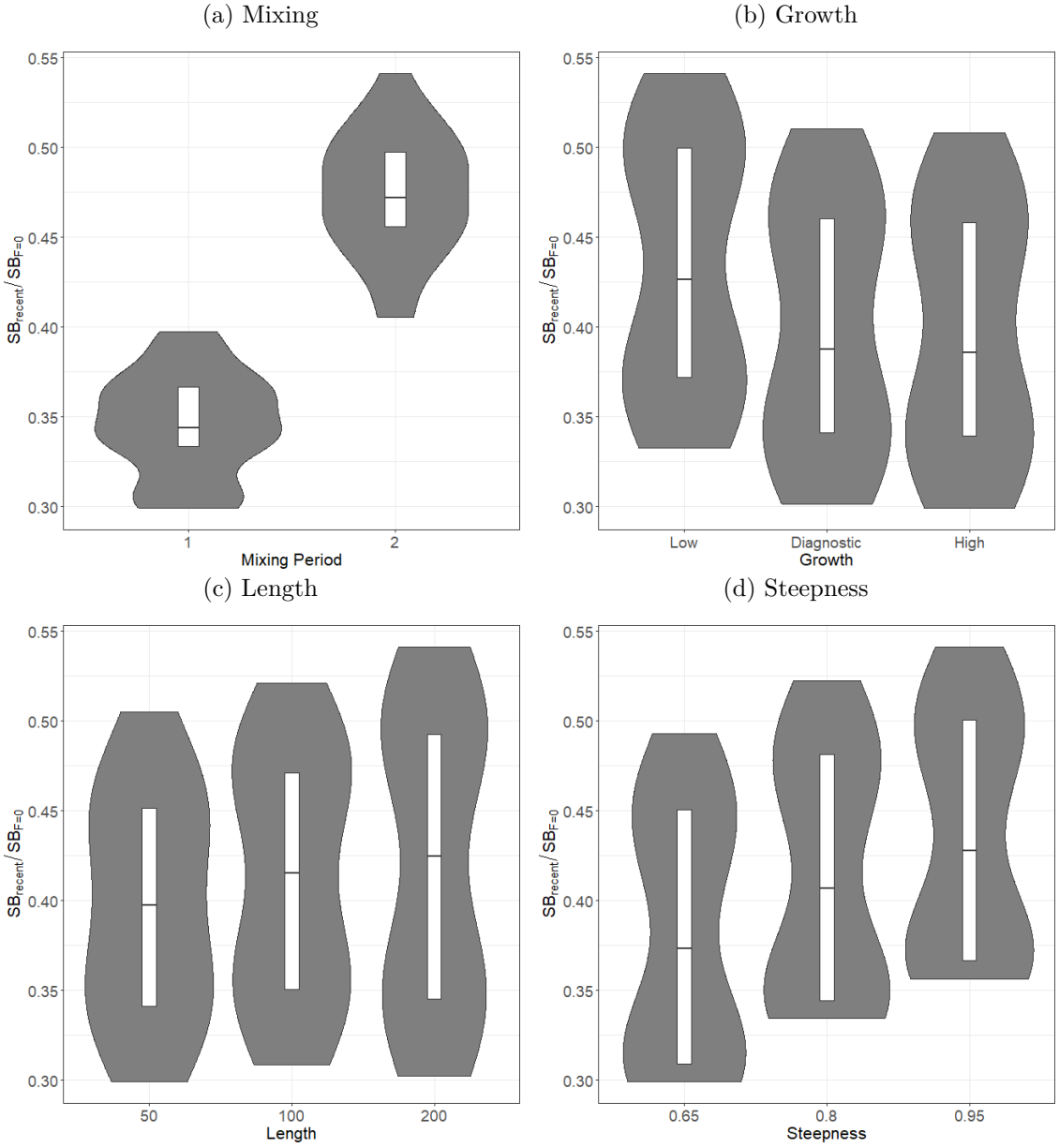


Figure 58: Box and violin plots summarizing the estimated $SB_{recent}/SB_{F=0}$ for each of the models in the structural uncertainty grid for the 5 region model. The line in the box is the median of the estimates, while the box shows the 50^{th} percentile. The shaded area shows the probability distribution (or density) of the estimates of all models across the axes of the structural uncertainty grid.

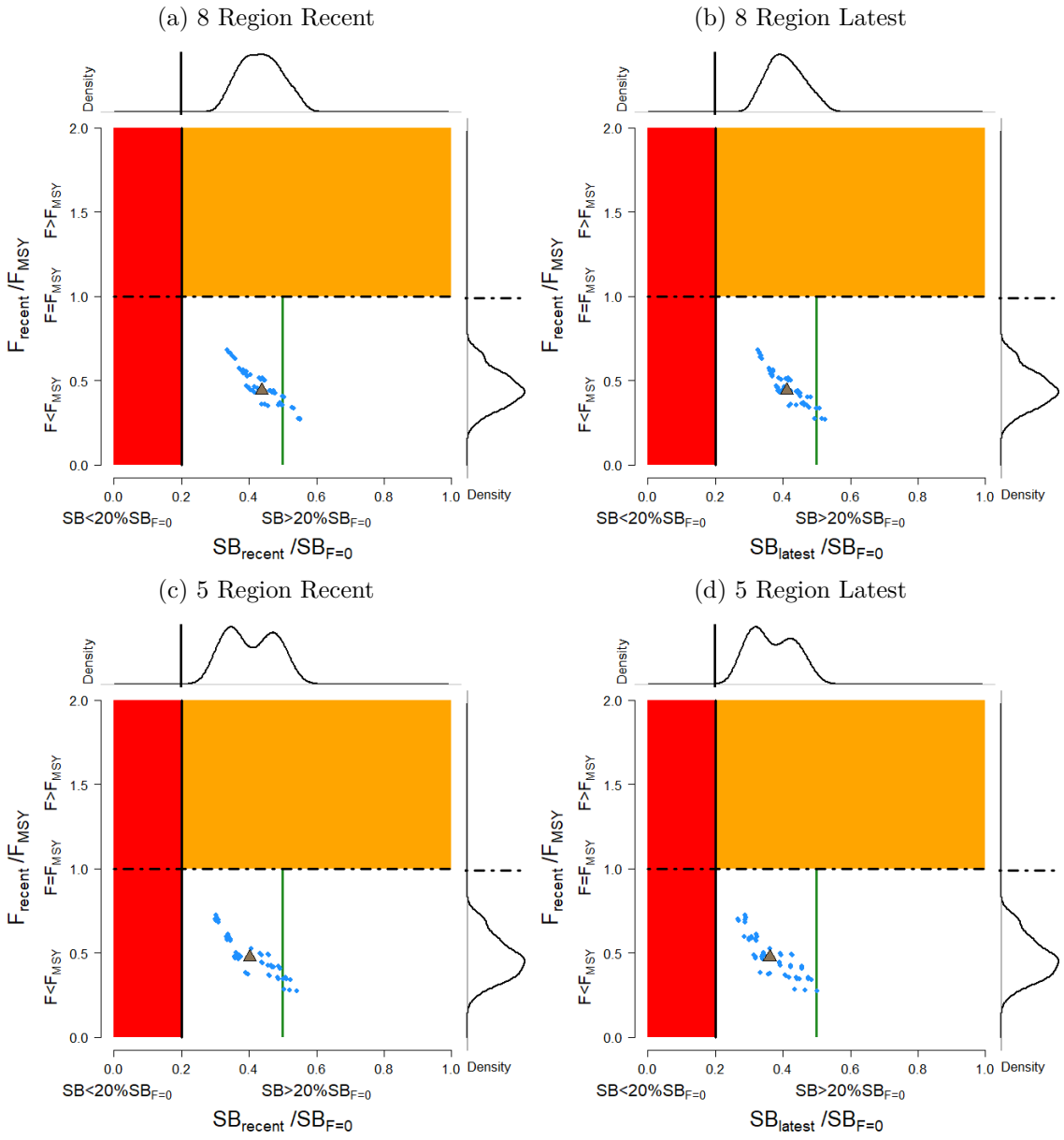


Figure 59: Majuro plots representing stock status in terms of spawning potential depletion and fishing mortality. The red zone represents spawning potential levels lower than the agreed limit reference point which is marked with the solid black line. The orange region is for fishing mortality greater than F_{MSY} (marked with the black dashed line). The green line indicates the target reference point $50\%SB_{F=0}$. The plots summarize the results for each of the models in the structural uncertainty grid for the 8 region and 5 regions models and the brown triangle is the median of the estimates. The top panels are for the 8 region model, while the bottom panels are the 5 region model; the left panels display $SB_{recent}/SB_{F=0}$ and the right panels display $SB_{latest}/SB_{F=0}$.

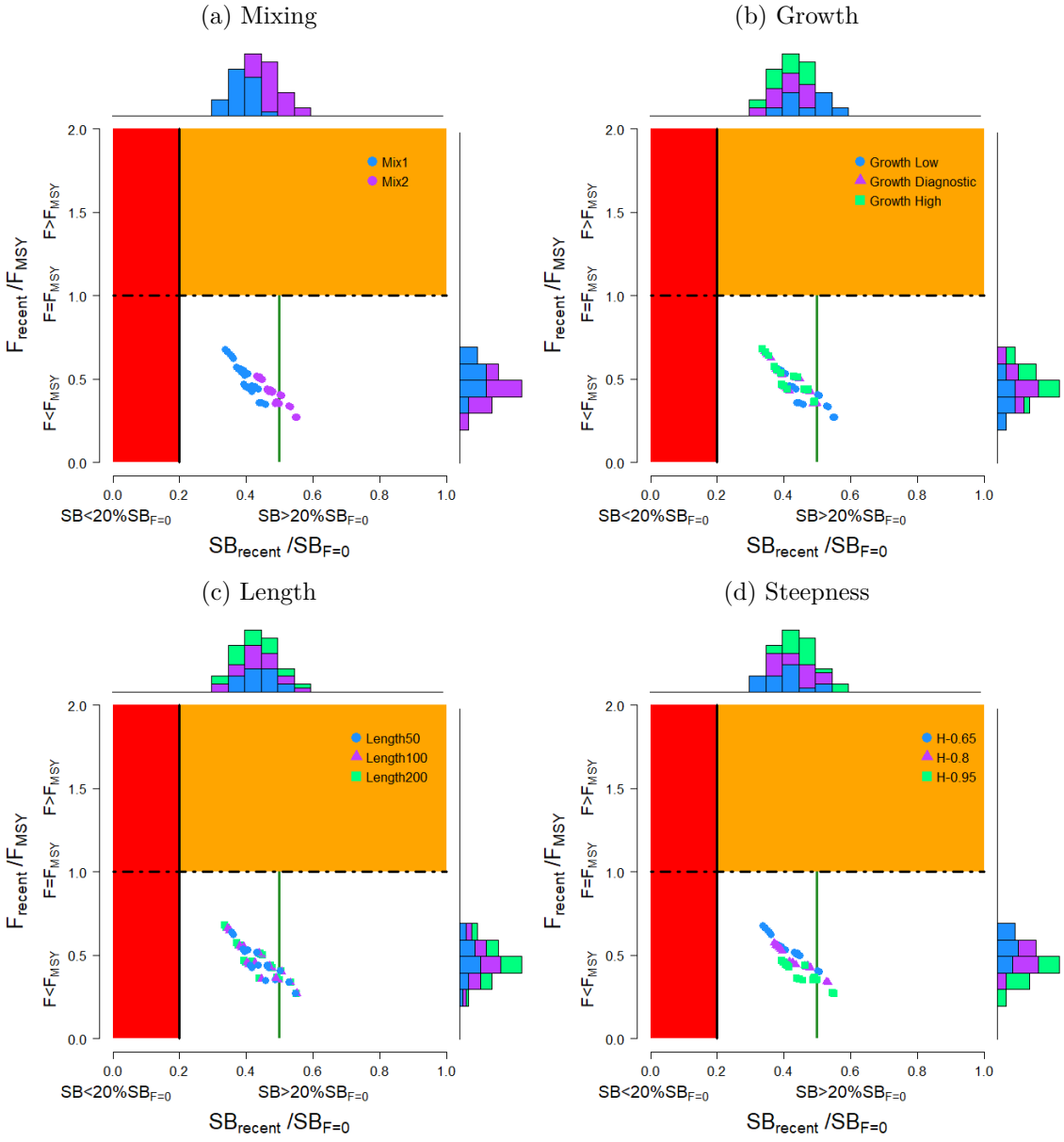


Figure 60: Majuro plots summarizing the results for each of the models in the structural uncertainty grid for the 8 region model (see Figure 59 for explanation of the interpretation of Majuro plots). All panels display $SB_{recent}/SB_{F=0}$ for different axes of the structural uncertainty.

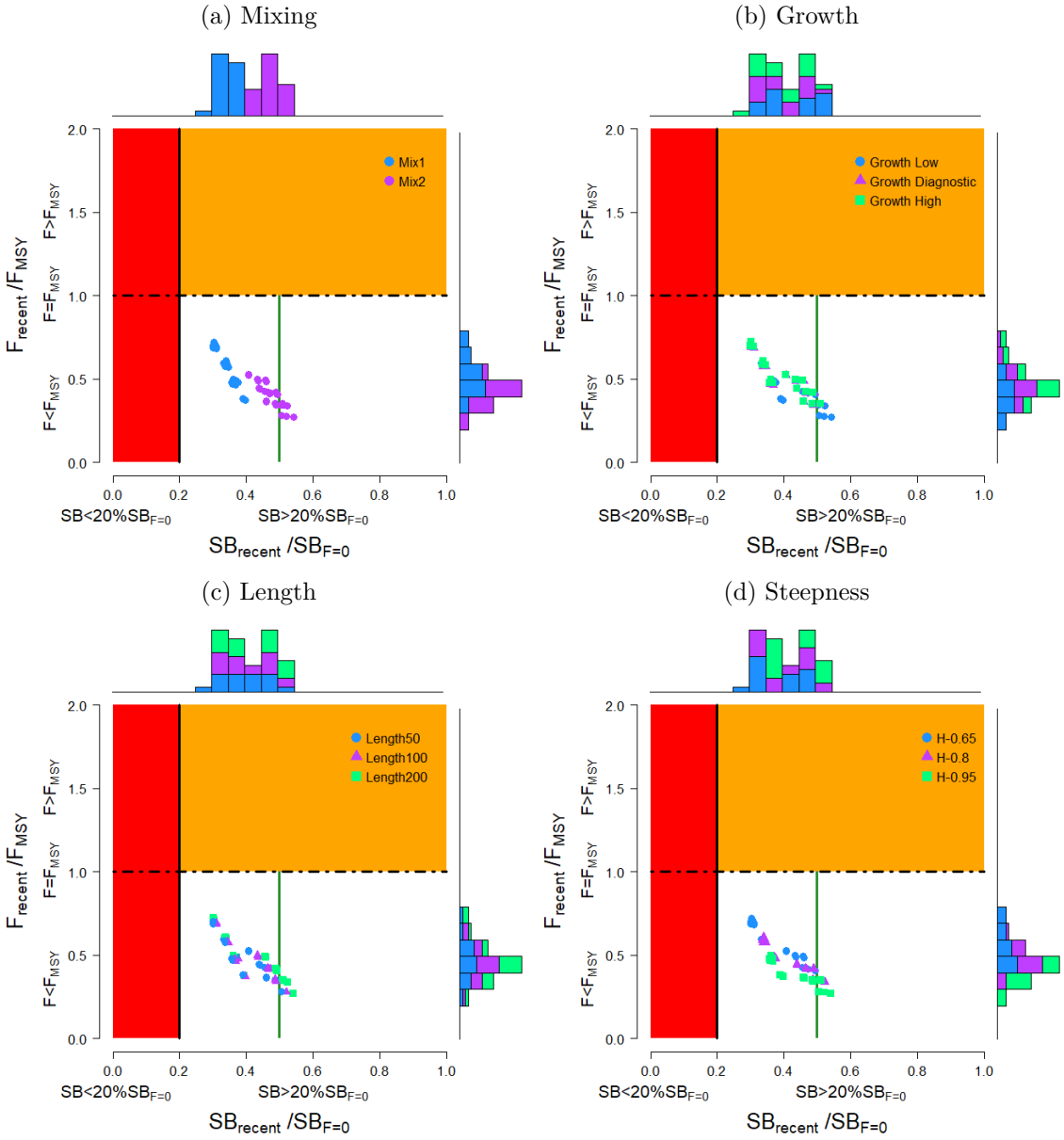


Figure 61: Majuro plots summarizing the results for each of the models in the structural uncertainty grid for the 5 region model (see Figure 59 for explanation of the interpretation of Majuro plots). All panels display $SB_{recent}/SB_{F=0}$ for different axes of the structural uncertainty.

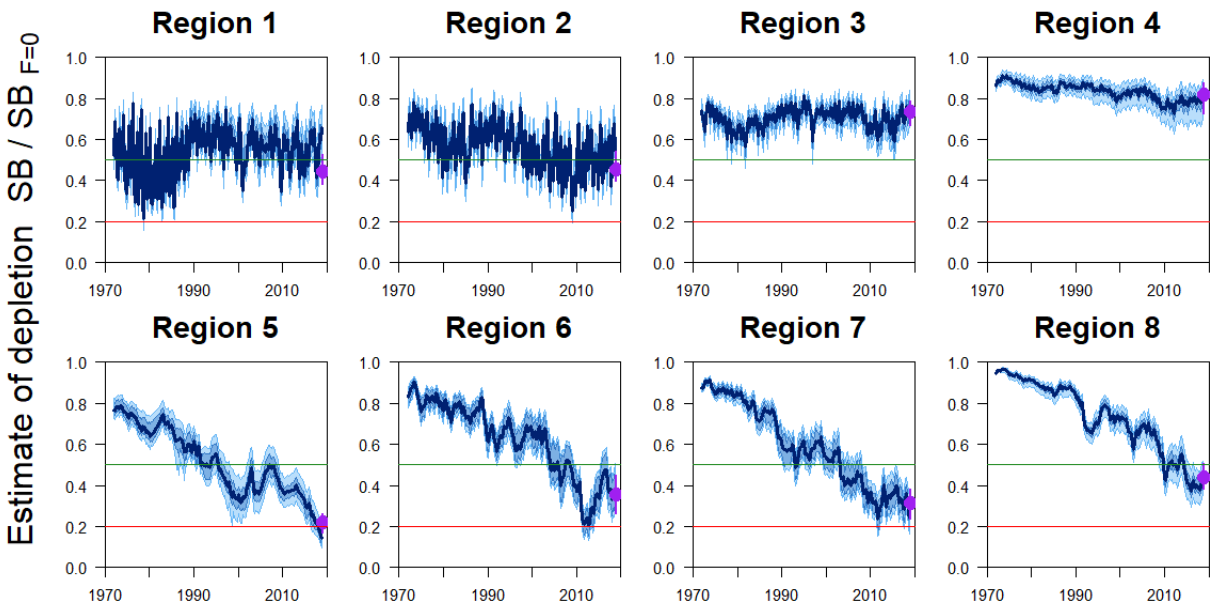


Figure 62: Distribution of time series depletion estimated for in each region of the 8 region models in the structural uncertainty grid. The dark blue line is the median, the blue region represents the 50^{th} percentile range, the light blue is the 80^{th} percentile range, the purple point and error bars are the median and 80^{th} percentile of $SB_{recent}/SB_{F=0}$ for that region.

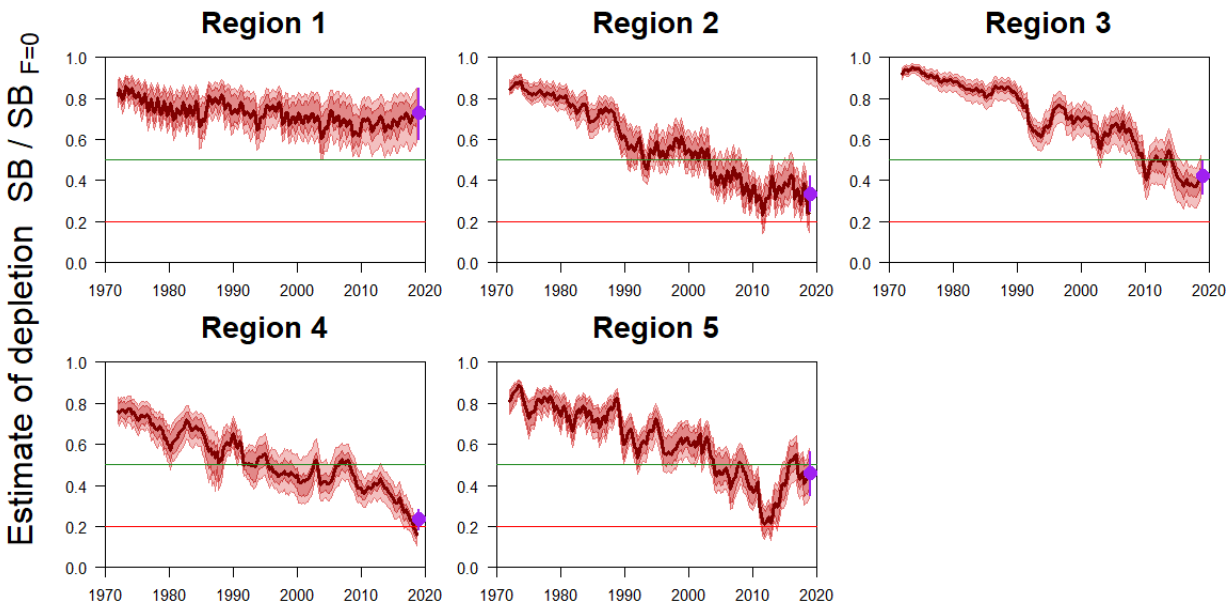
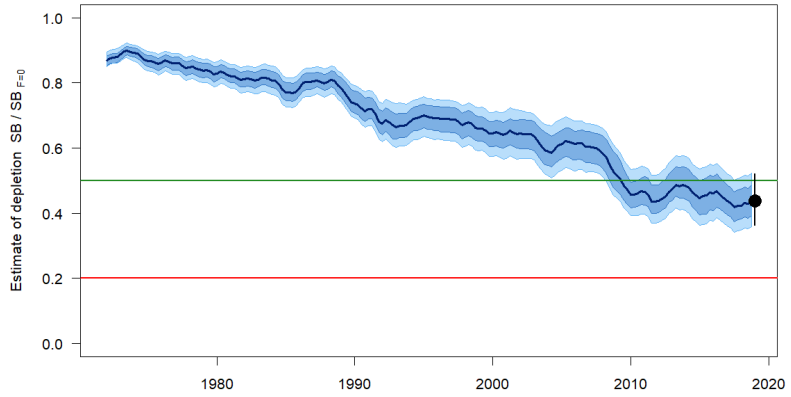
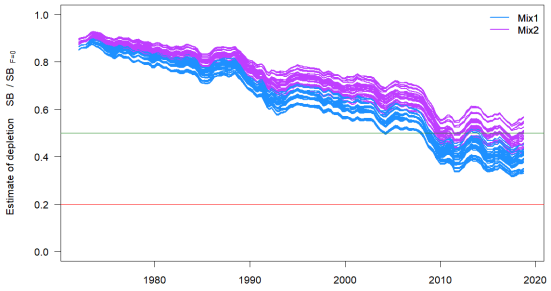


Figure 63: Distribution of time series depletion estimated for in each region of the 5 region models in the structural uncertainty grid. The dark red line is the median, the red region represents the 50^{th} percentile range, the light red is the 80^{th} percentile range, the purple point and error bars are the median and 80^{th} percentile of $SB_{recent}/SB_{F=0}$ for that region.

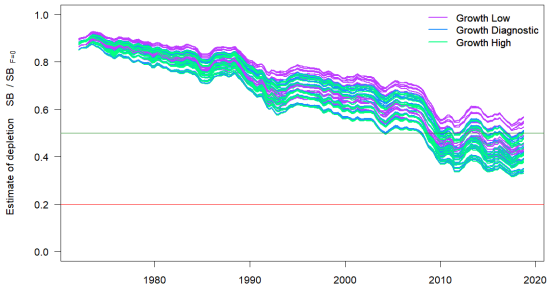
(a) 8 Region Quantiles



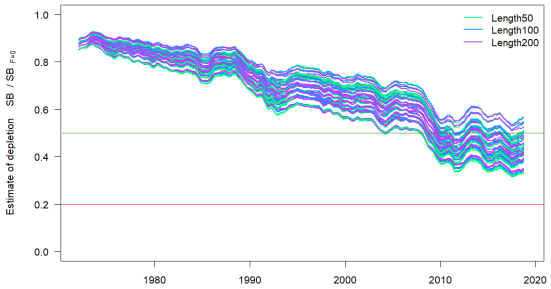
(b) Mixing



(c) Growth



(d) Length



(e) Steepness

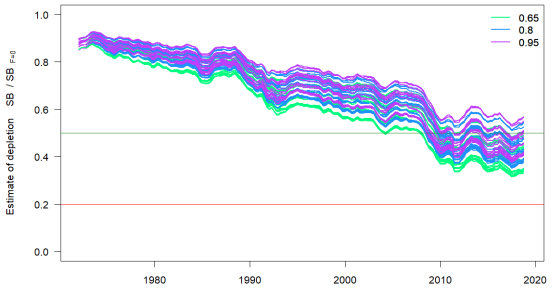
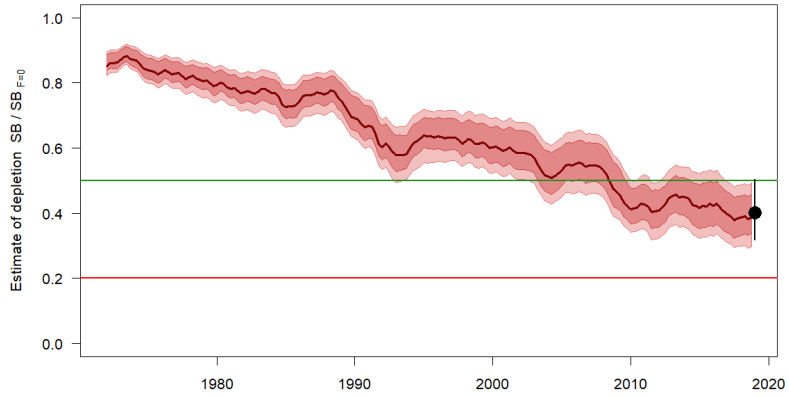
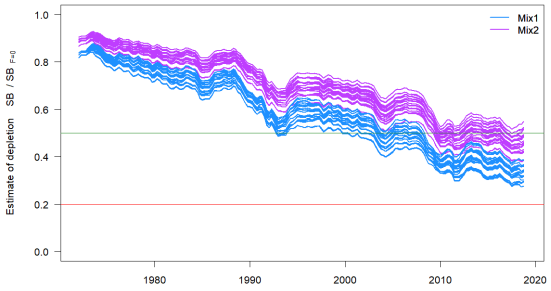


Figure 64: Plots showing the trajectories of fishing depletion of spawning potential for the model runs with the 8 region structure included in the structural uncertainty grid (see Section 6.3 for details of the structure of the grid models). Panel 64a shows the median, 50% quantile, and 80% quantile of instantaneous depletion across the structural uncertainty grid and the point and error bars is the median and 80% quantile of estimates of $SB_{recent}/SB_{F=0}$. The remaining 4 panels show the axes used in the grid, with the color denoting the level within the axes for each model.

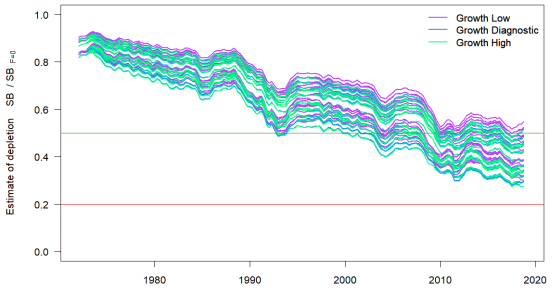
(a) 5 Region Quantiles



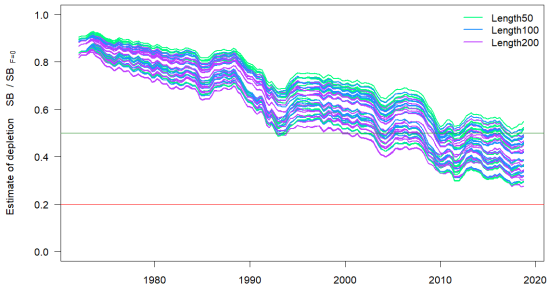
(b) Mixing



(c) Growth



(d) Length



(e) Steepness

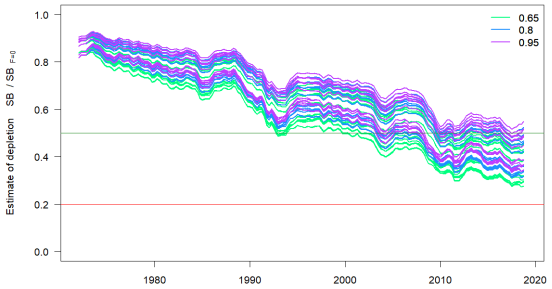


Figure 65: Plots showing the trajectories of fishing depletion of spawning potential for the model runs with the 5 region structure included in the structural uncertainty grid (see Section 6.3 for details of the structure of the grid models). Panel 65a shows the median, 50% quantile and 80% quantile of depletion across the structural uncertainty grid and the point and error bars is the median and 80% quantile of estimates of $SB_{recent}/SB_{F=0}$. The remaining 4 panels show the axes used in the grid, with the color denoting the level within the axes for each model.

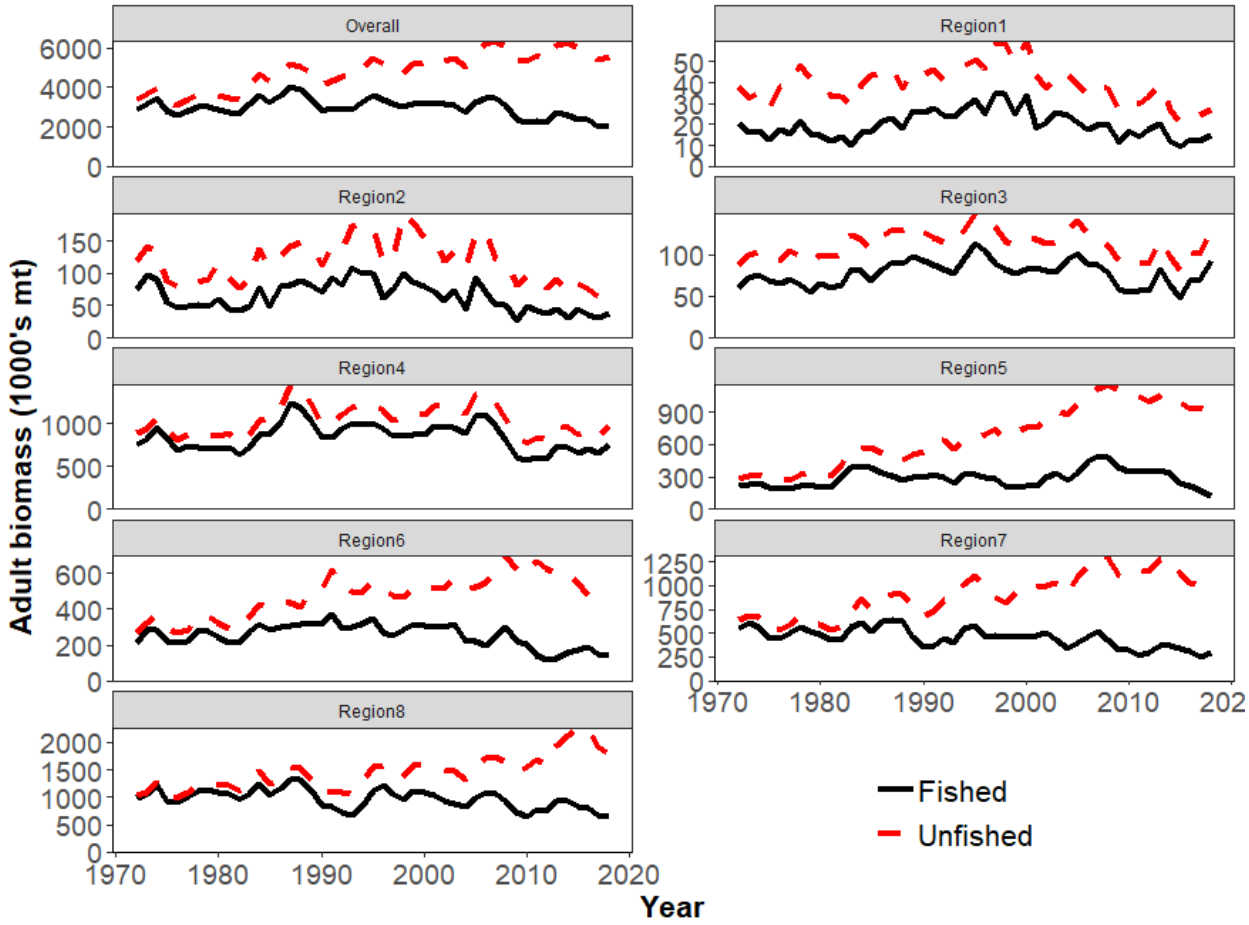


Figure 66: Comparison of the estimated annual spawning potential trajectories (lower solid black lines) with those trajectories that would have occurred in the absence of fishing (upper dashed red lines) for each region and overall, for the diagnostic model.

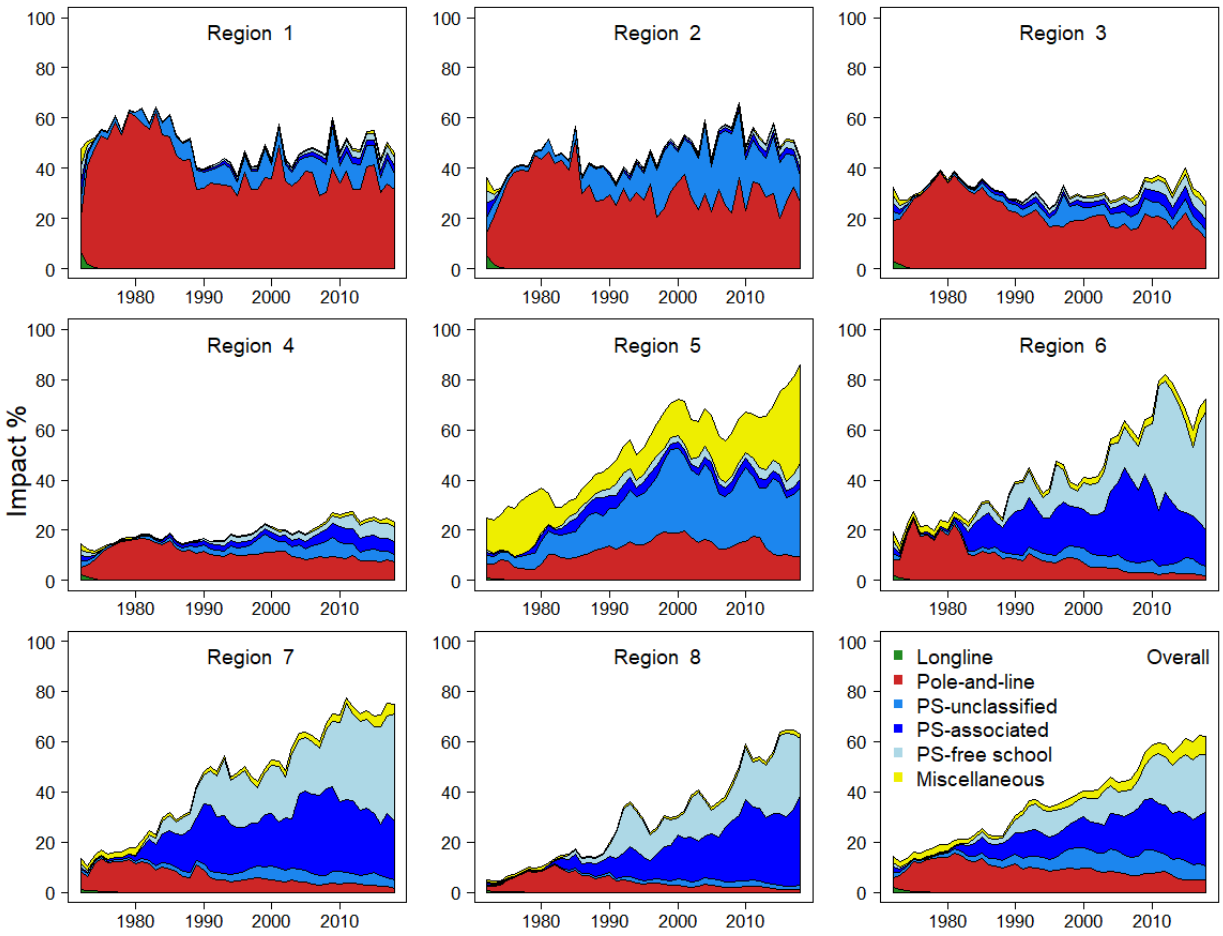


Figure 67: Estimates of reduction in spawning potential due to fishing (fishery impact = $1-SB_{latest}/SB_{F=0}$) by region, and over all regions (lower right panel), attributed to various fishery groups for the diagnostic model.

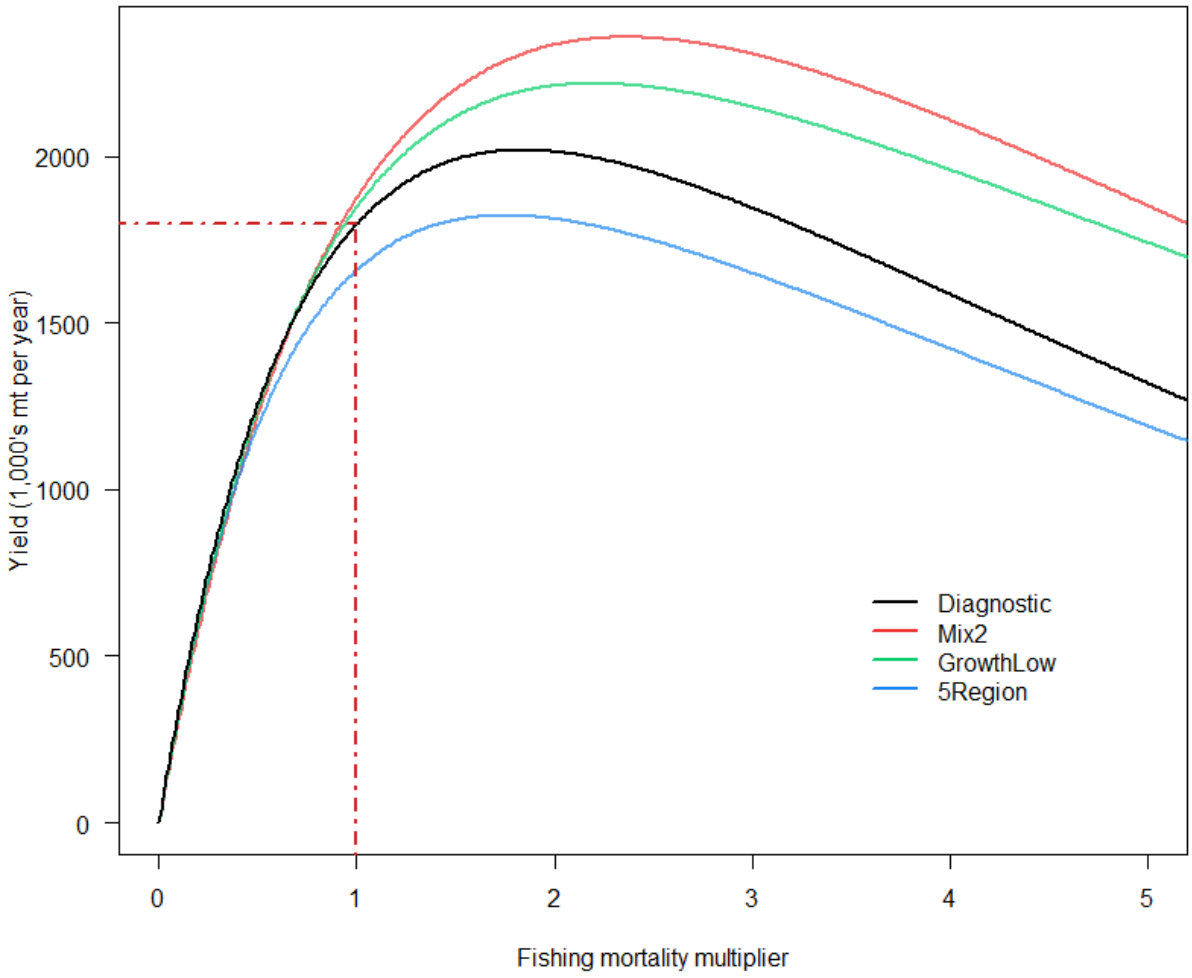


Figure 68: Estimated yield as a function of fishing mortality multiplier for the diagnostic model and a few of the one-off sensitivity models. The red dashed line indicates the equilibrium yield at current fishing mortality.

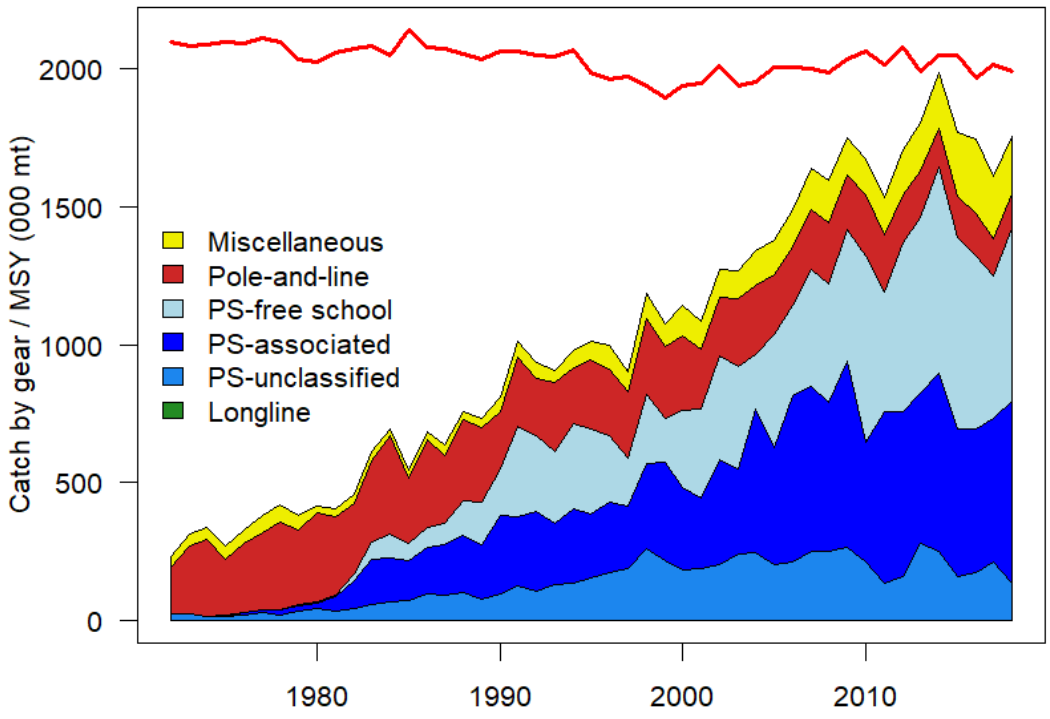
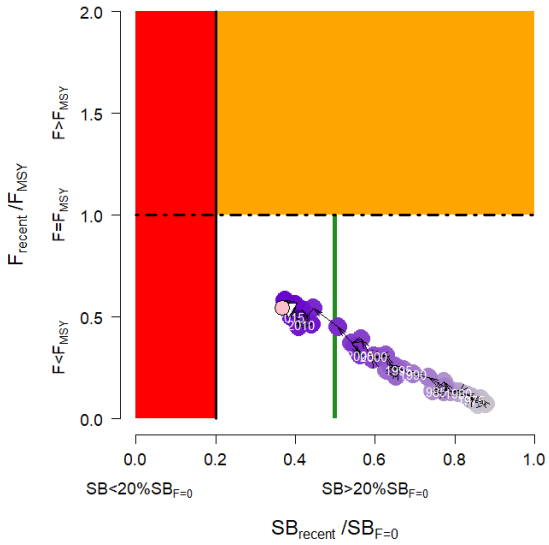
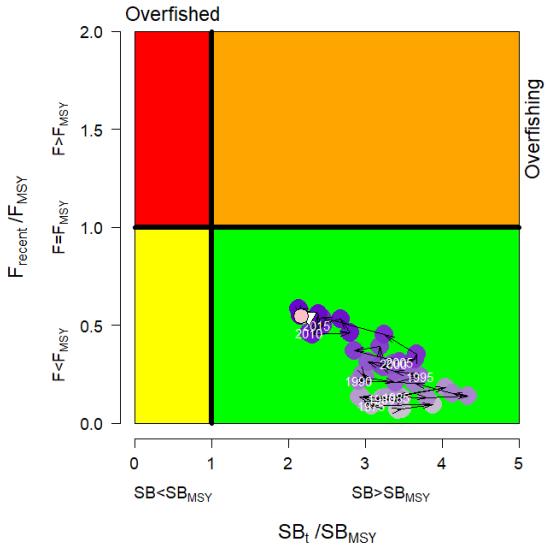


Figure 69: History of the annual estimates of MSY (red line) for the diagnostic model compared with annual catch by the main gear types. Note that this is a "dynamic" MSY which is explained further in Section 5.7.4.

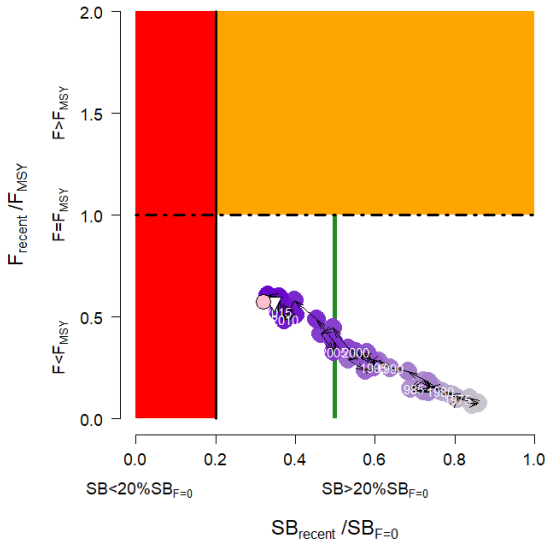
(a) Majuro 8 Region



(b) Kobe 8 Region



(c) Majuro 5 Region



(d) Kobe 5 Region

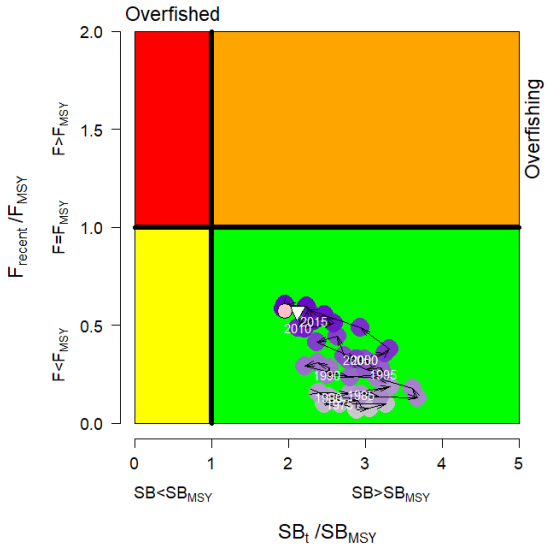


Figure 70: Majuro plot (a) for the diagnostic model (see Figure 59 for explanation of the interpretation of Majuro plots). The pink circle is $SB_{latest}/SB_{F=0}$ and the white triangle is $SB_{recent}/SB_{F=0}$, which are both detailed in Table 4. The equivalent Kobe plot for the diagnostic model (b) and the 5 region Majuro (c) and Kobe (d) are provided for comparison.

(a) 8 Region model

(b) 5 Region model

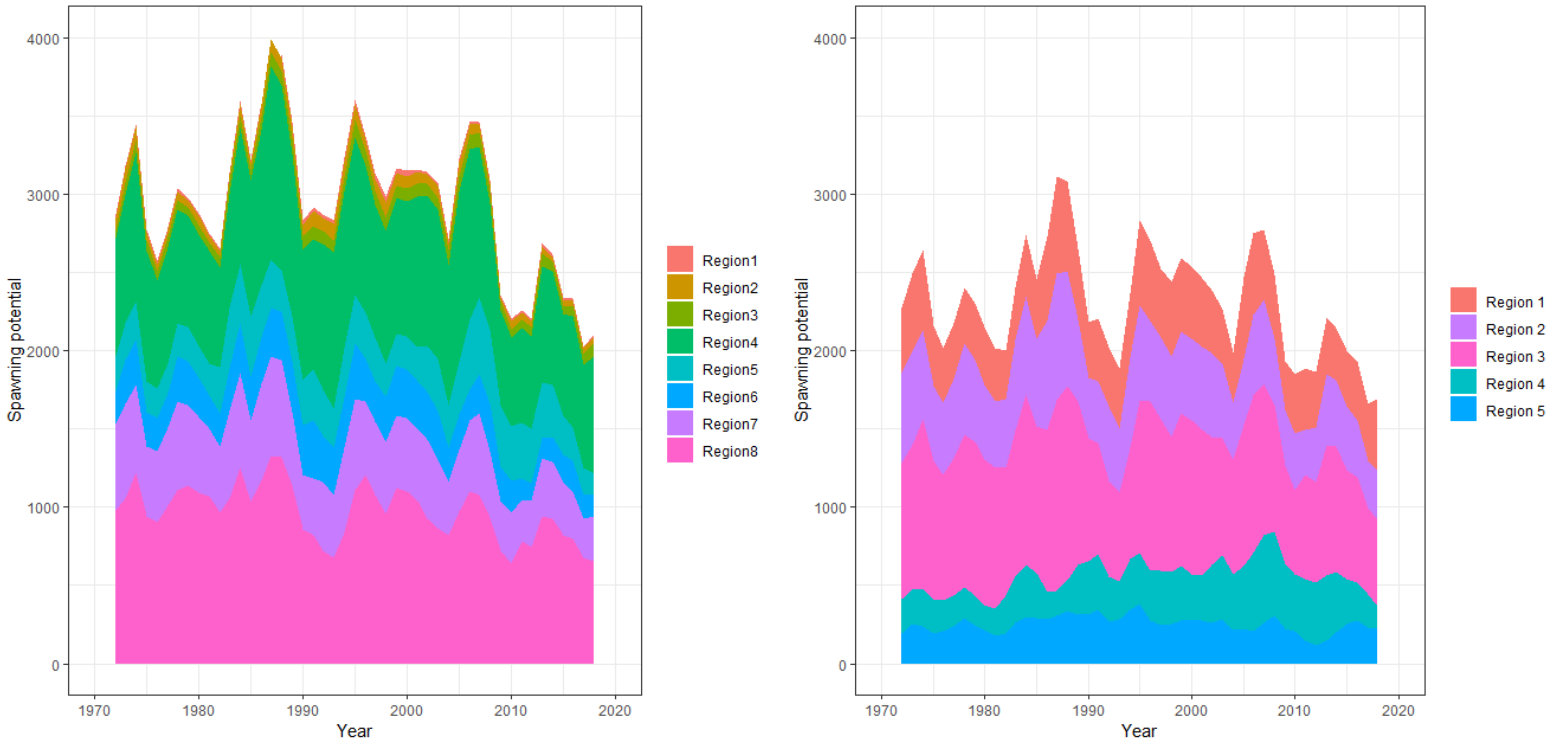


Figure 71: Comparison of the biomass estimated by region for the 8 region and 5 region diagnostic models.

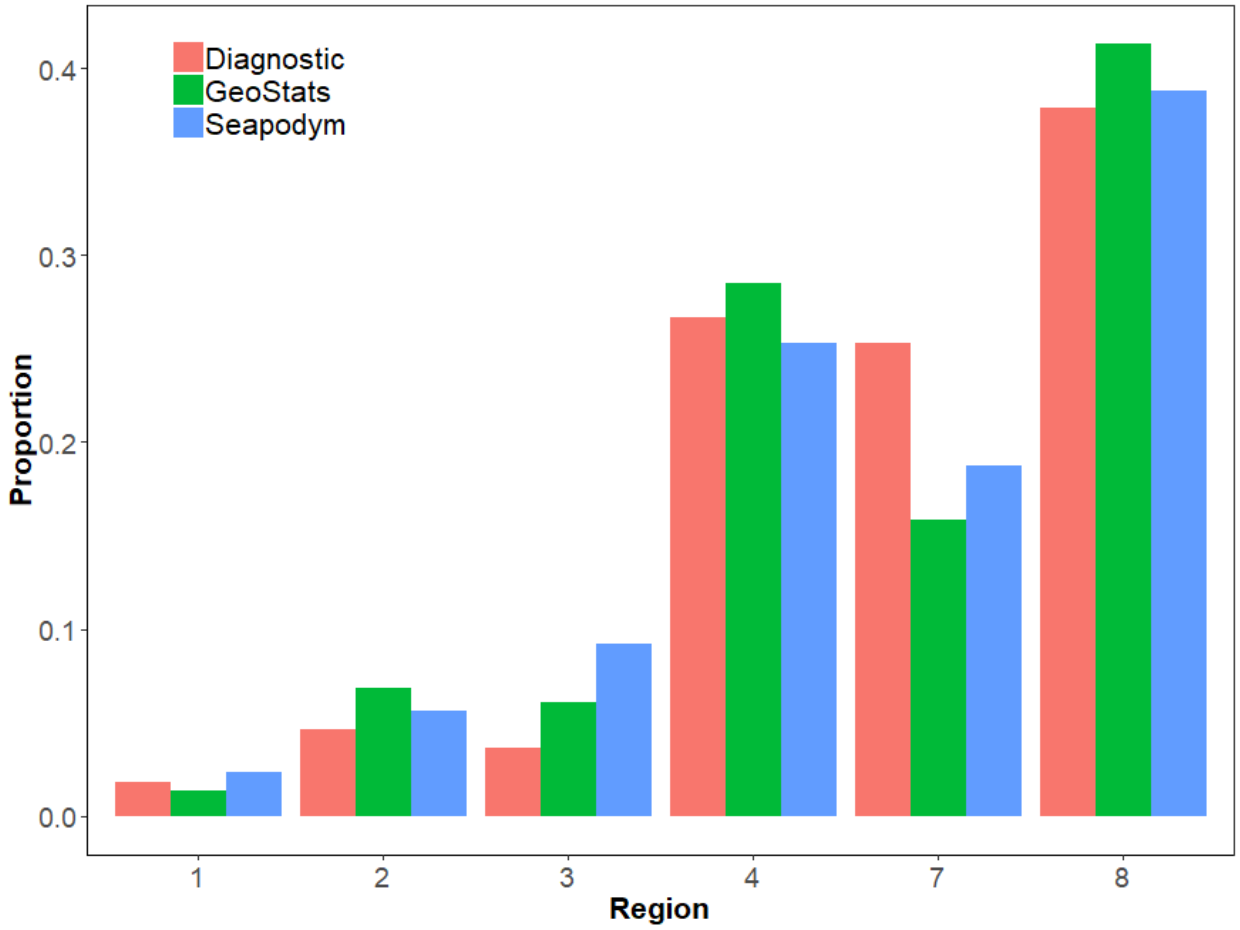


Figure 72: Proportion of average total biomass in Regions 1-4, 7, and 8 relative to the average sum of these regions from the 2019 diagnostic model for 1972 to 2018, from the geostatistical CPUE analysis of the Japanese pole-and-line for 1972 to 2018, and the SEAPODYM model from 1979 to 2010.

Appendix A

A.1 Likelihood profiles

The approach for calculating a likelihood profile of the total population scaling parameter (*totpop*) is outlined in Section 5.6. The profile reflects the loss of fit over all the data, i.e. the overall objective function value, caused by changing the population scale from that of the maximum likelihood estimated value. A range of fixed values were used until the best fit for each data source was found. The likelihood profile for the diagnostic model is shown in Figure A1 and the 5 region model is shown in Figure A2. Both profiles displays significant declines in the total log-likelihood as the total biomass estimate moves further away from the maximum value from the model. A likelihood profile for the overdispersion parameter (Tau) by fixing the parameters at different values is presented in Figure A3. The likelihood profiles for each fishery in the 8 region model are presented in Figures A4 to A7. Additional likelihood profiles for the *Mix2*, *Length50*,*Length200*, and *GrowthLow* models are presented in Figure A8.

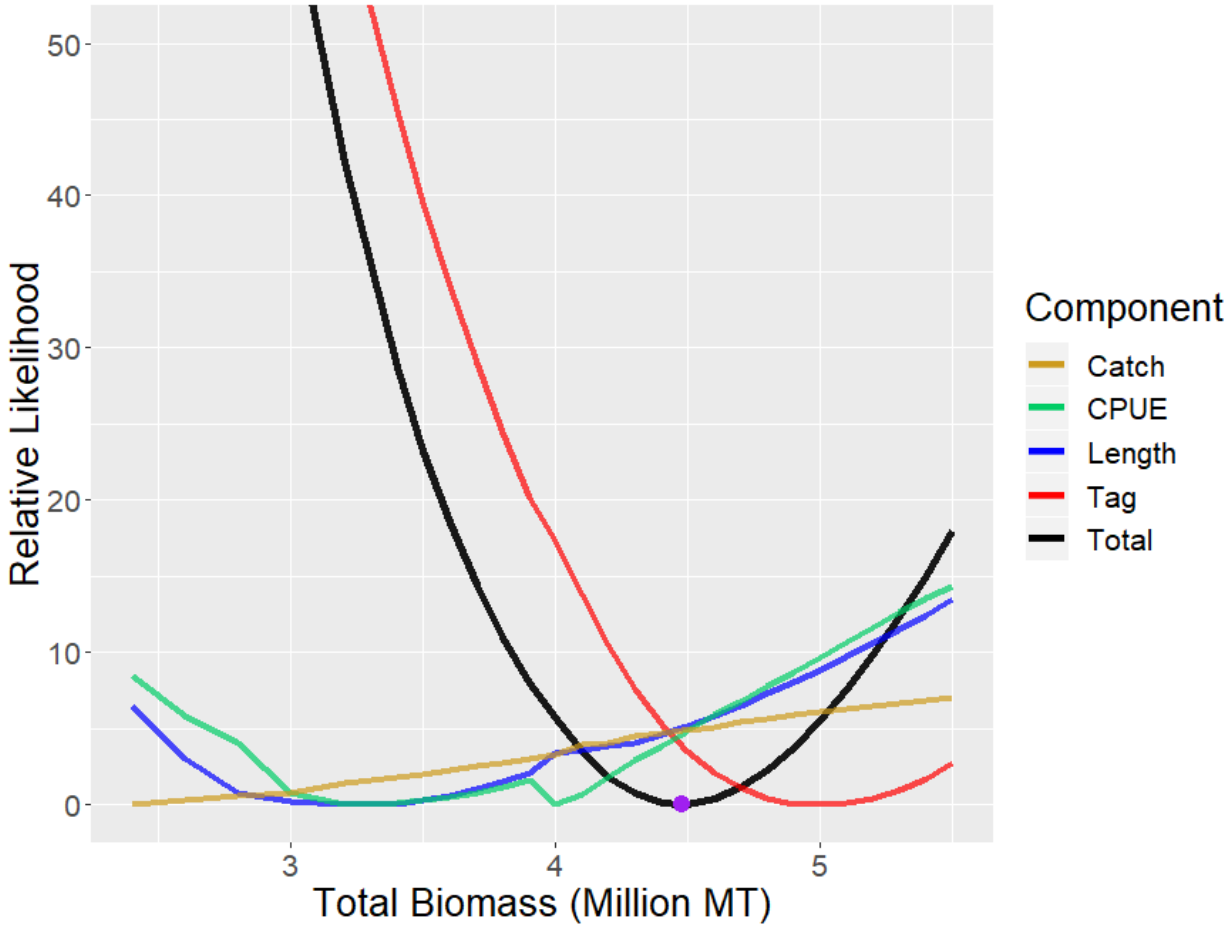


Figure A1: Profile of the total log-likelihood with respect to total biomass in million mt across a range of fixed values for the diagnostic model, where the purple dot indicates the maximum likelihood estimate.

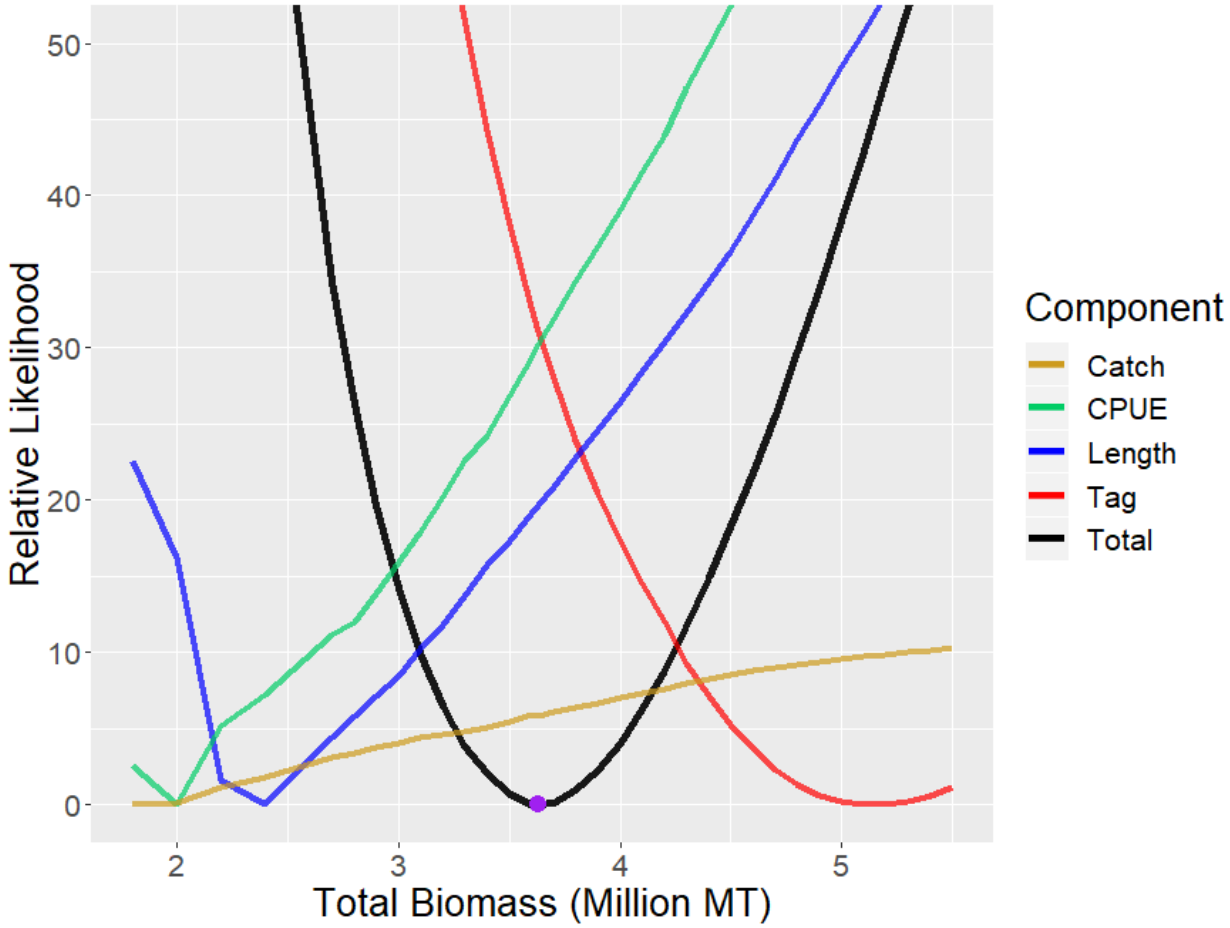


Figure A2: Profile of the total log-likelihood with respect to total biomass in million mt across a range of fixed values for the 5 region model, where the purple dot indicates the maximum likelihood estimate.

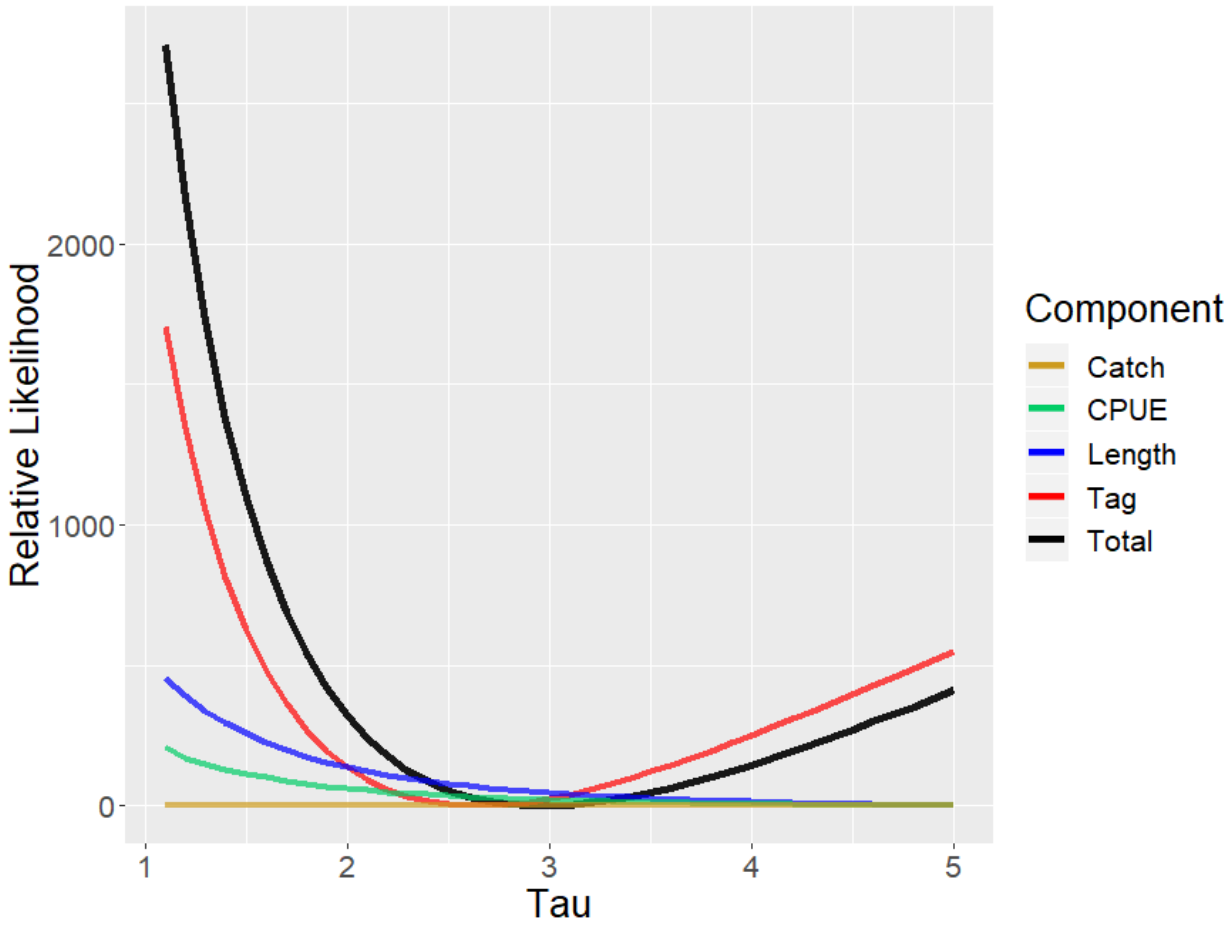


Figure A3: Profile of the total log-likelihood with respect to the tag overdispersion parameter (τ) across a range of fixed values for the diagnostic model.

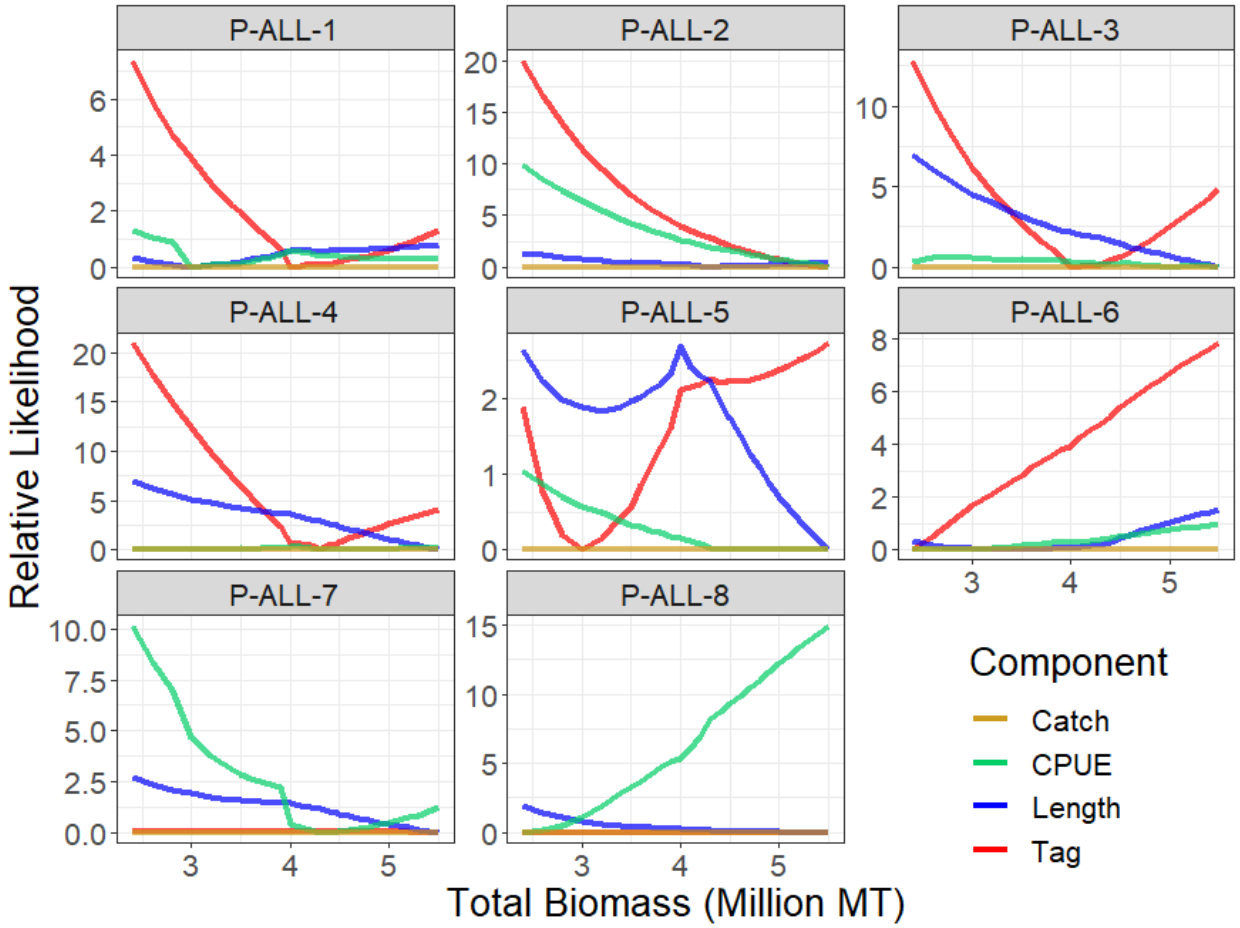


Figure A4: Profile of the total log-likelihood with respect to total biomass in million mt for the pole-and-line across a range of fixed values for the diagnostic model, note the large difference in scale between fisheries.

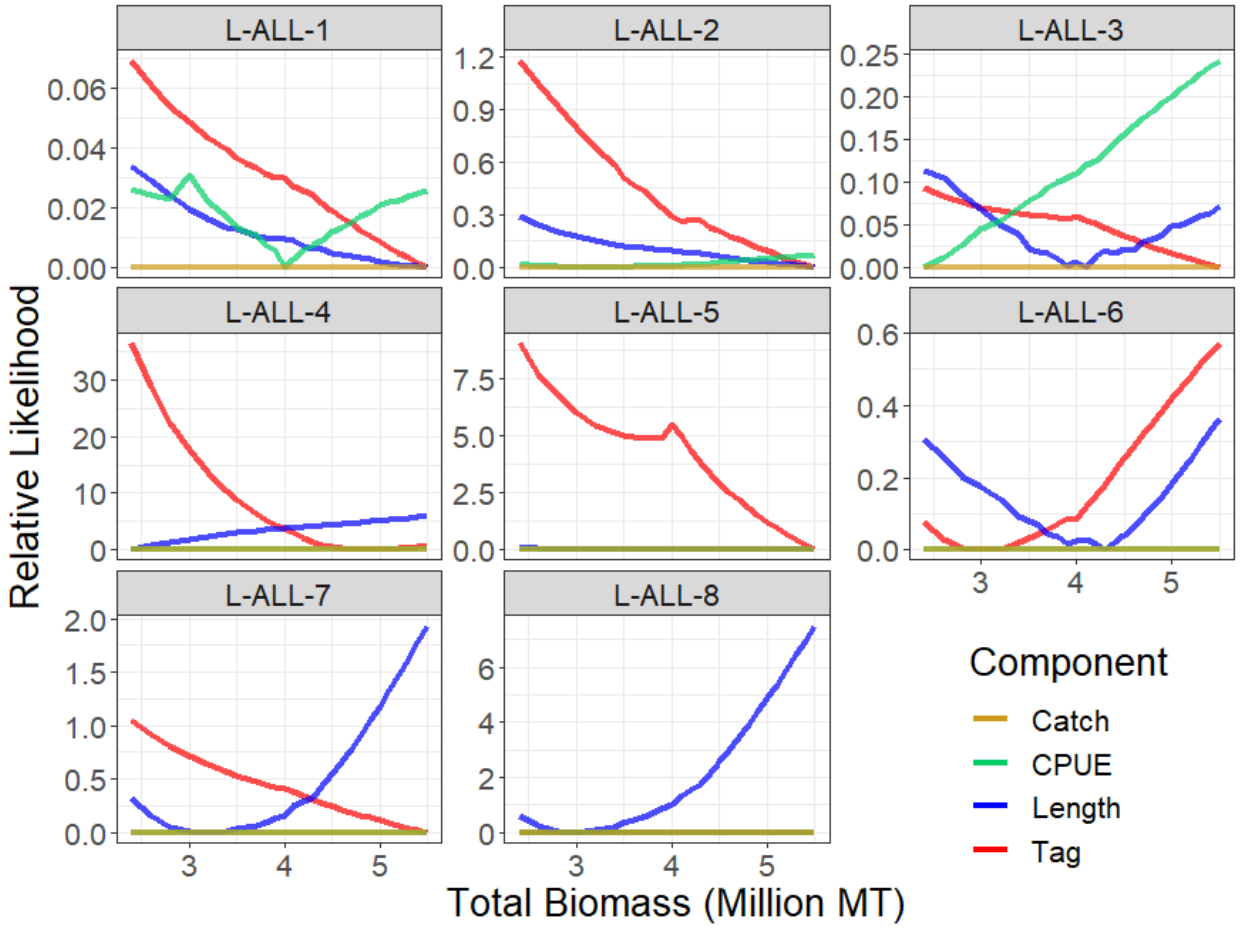


Figure A5: Profile of the total log-likelihood with respect to total biomass in million mt across a range of fixed values for the longline fisheries in the diagnostic model, note the large difference in scale between fisheries.

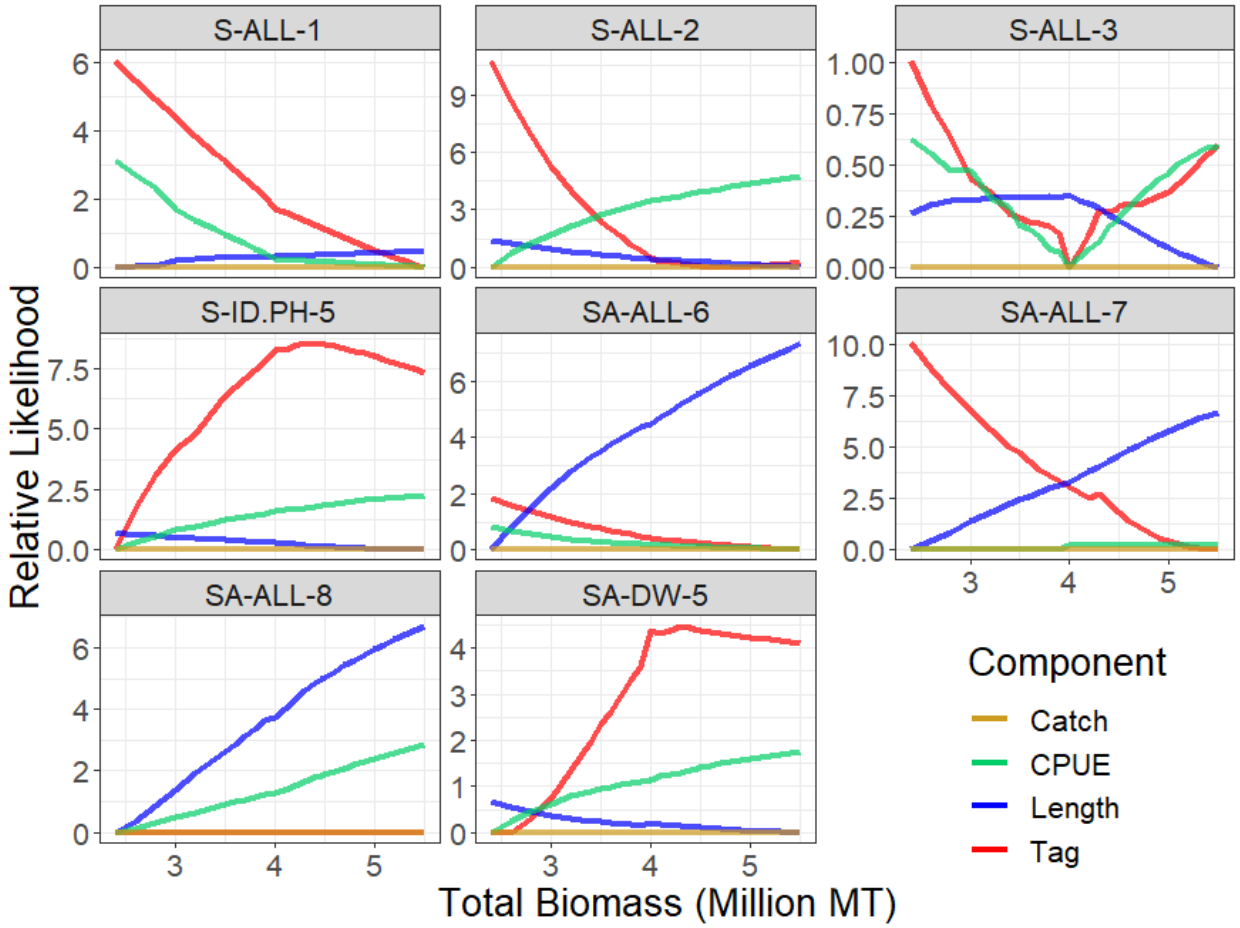


Figure A6: Profile of the total log-likelihood with respect to the total population scaling parameter for the associated and unclassified purse seine fisheries across a range of fixed values for the diagnostic model, note the large difference in scale between fisheries.

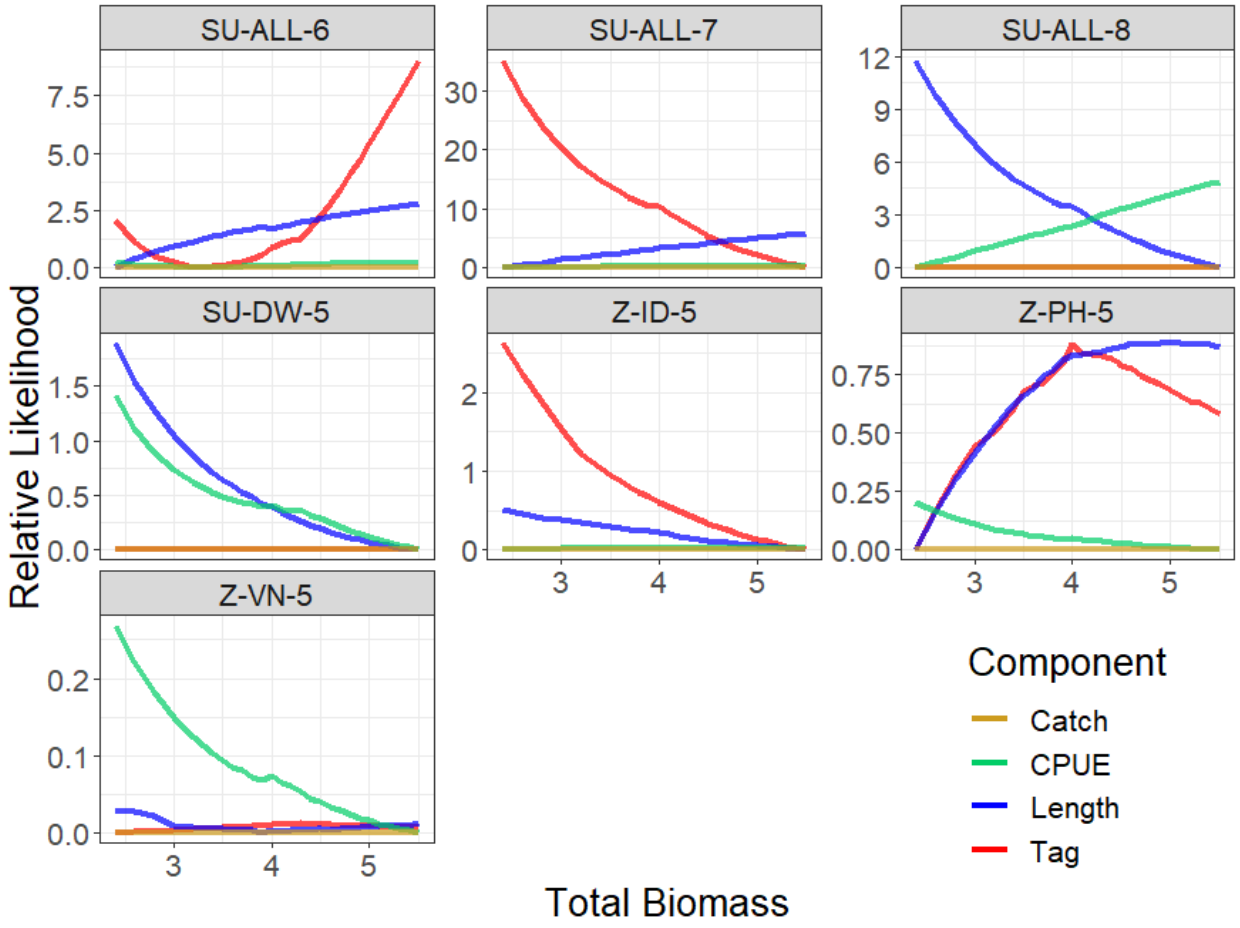
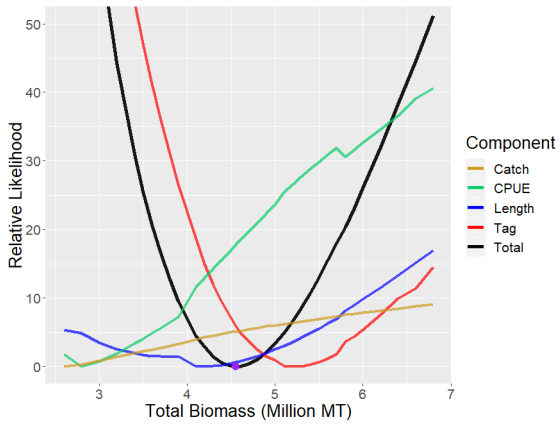
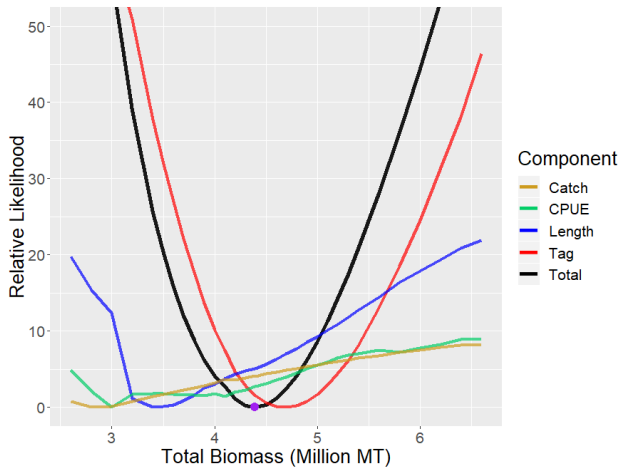


Figure A7: Profile of the total log-likelihood with respect to total biomass in million mt for the unassociated and miscellaneous fisheries across a range of fixed values for the diagnostic model, note the large difference in scale between fisheries.

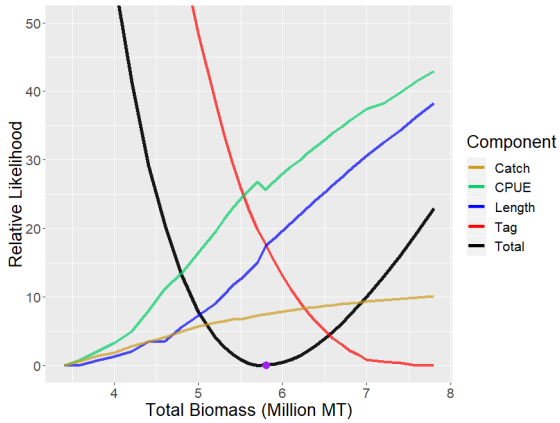
(a) *Length200*



(b) *Length50*



(c) *Mix2*



(d) *GrowthLow*

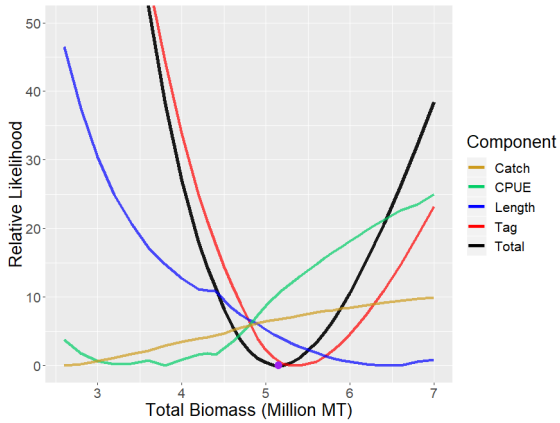


Figure A8: Profiles of total log-likelihood with respect to the total biomass in million mt across a range of fixed values where the purple dot indicates the maximum likelihood estimate, for A8a the *Length200* model, A8b the *Length50* model, A8c the *Mix2* model, and A8d the *GrowthLow* model.

A.2 'Status quo' stock projections for WCPO skipjack tuna

The potential stock consequences of 'status quo' fishing conditions (at average fishing levels over the period 2016 to 2018) were evaluated through stochastic projections, using the uncertainty framework approach previously endorsed by SC:

- Stochastic 30 year projections were conducted from each assessment model within the uncertainty grid developed for the 2019 skipjack assessment.
- For each model, 100 stochastic projections, which incorporate future recruitments randomly sampled from historical deviates, were performed.
- Future recruitment in the projection period was based upon the long-term recruitment patterns (sampled from the period 1982 to 2017).
- Future purse seine effort, and pole and line and other fishery catches, were assumed to remain at the average levels over the period 2016 to 2018.
- The outputs of the projections (median $SB_{2048}/SB_{F=0}$ and F/F_{MSY}, and risk $SB_{2045}/SB_{F=0} <$ LRP) were combined across the 54 models within each regional structure.
- Catchability (which can have a trend in the historical component of the model) was assumed to remain constant in the projection period at the level estimated in the terminal year of the assessment model.

Results of the projections are summarized in Table A1 and Figure A9.

Table A1: Summary of skipjack stock outcomes under 2016–2018 average fishing scenario.

Regional structure	$SB_{2048}/SB_{F=0}$	Risk $SB_{2048}/SBF = 0 <$ LRP	$F_{recent}/F_{\text{MSY}}$
8 region	0.43	0%	0.78
5 region	0.39	2%	0.54

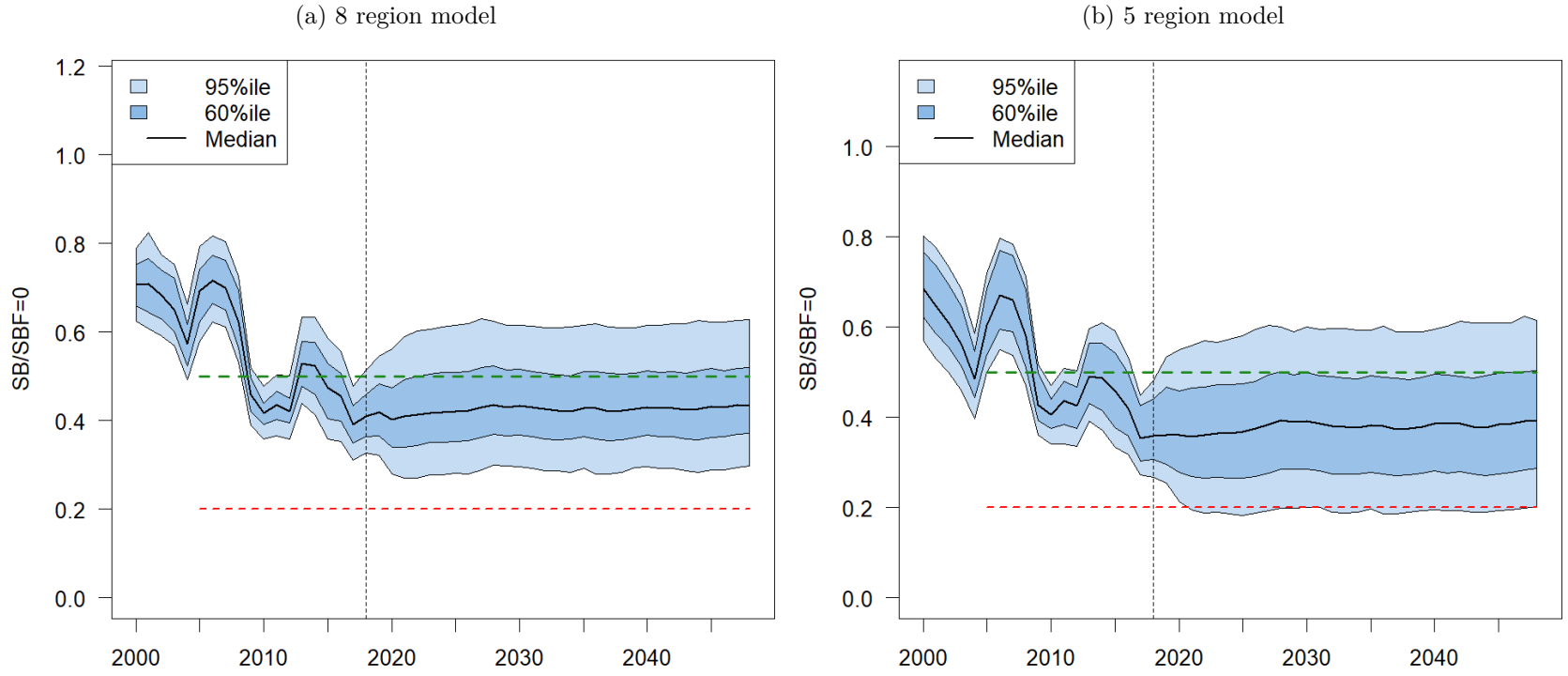


Figure A9: Time series of skipjack tuna spawning biomass ($SB_t/\bar{SB}_{F=0}$, where $\bar{SB}_{t,F=0}$ is the average SB from t-11 to t-1) from the uncertainty grid of assessment model runs by regional structure: 8 regions (left) and 5 regions (right), for the period 2000 to 2018, and stochastic projection results for the period 2019 to 2048 assuming 2016-2018 average fishing levels continue. Vertical dotted line at 2018 represents the last year of the assessment. During the projection period (2019-2048) levels of recruitment variability are assumed to match those over the time period used to estimate the stock-recruitment relationship (1982-2017). The red horizontal dashed line represents the agreed limit reference point, the green horizontal dashed line represents the current interim target reference point.

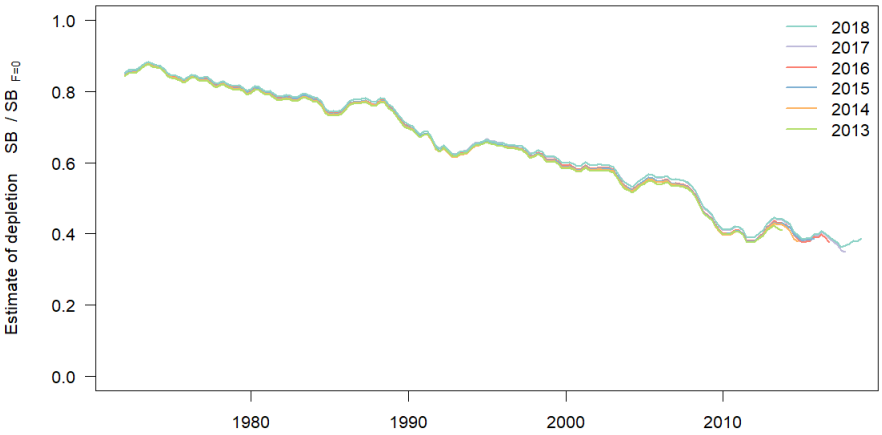
A.3 Retrospective analyses

A.3.1 Removal of recent years data

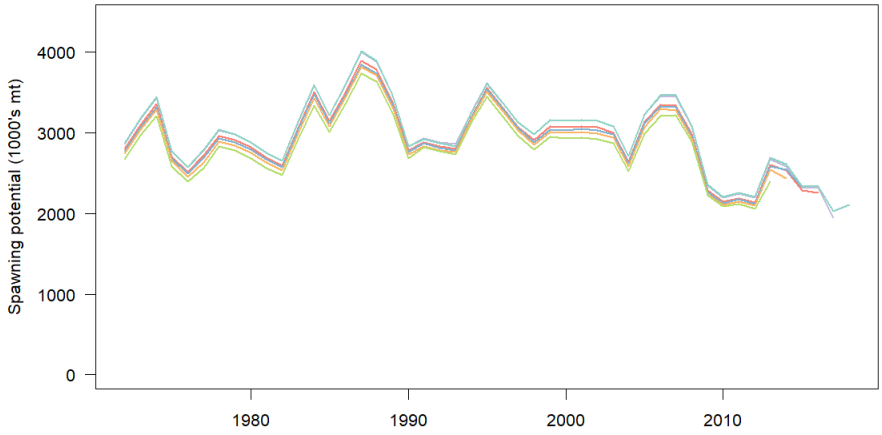
Retrospective analysis involves rerunning the selected model by consecutively removing successive years of data to estimate model bias (Cadrin and Vaughan, 1997; Cadigan and Farrell, 2005). A series of 6 models were fitted starting with the full data-set (through 2018), followed by models with the retrospective removal of all input data for the years 2018 to 2013. The models are named below by the final year of data included (e.g., 2013–2018). A comparison of the spawning biomass, recruitment and depletion trajectories are shown in Figure A10.

The models with each year of data removed sequentially produced estimates of spawning biomass and fishing depletion with very similar temporal dynamics to the full diagnostic model. For further details of the retrospective approach and methods to interpret the results the reader is referred to Scott et al. (2016).

(a) Estimated Fishing Depletion



(b) Estimated Spawning potential



(c) Estimated Recruitment

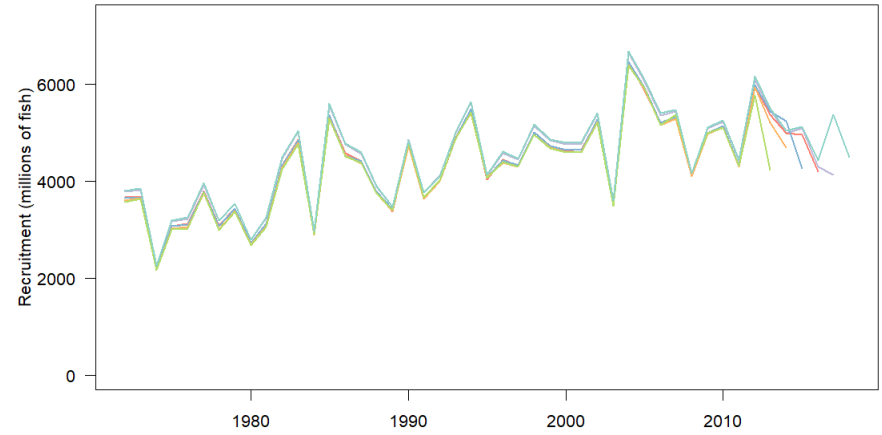
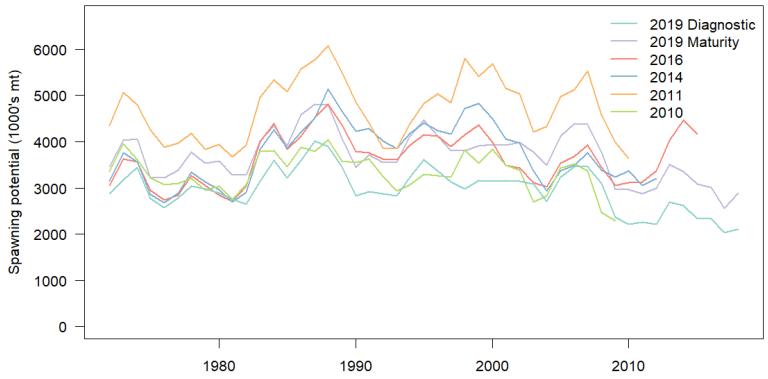


Figure A10: Estimated spawning potential, recruitment and fishery depletion ($SB_t/SB_{t,F=0}$) for each of the retrospective models.

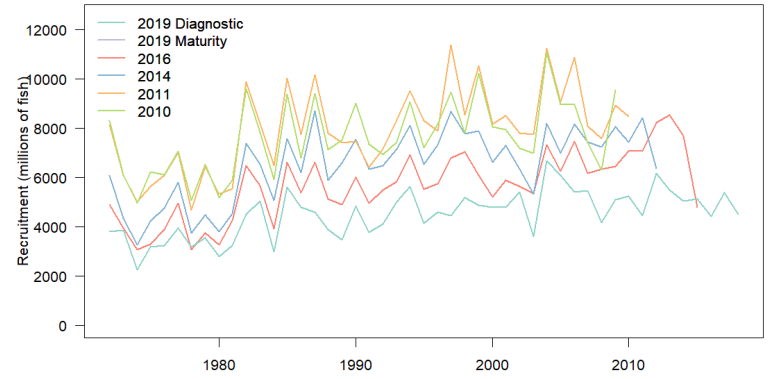
A.3.2 Comparison of results to previous skipjack stock assessments in the WCPO

The 2019 diagnostic model was compared retrospectively to the previous assessments conducted in 2016, 2014, 2011, and 2010. A model was created that was the 2019 diagnostic model but assumed the maturity ogive from the 2016 stock assessment (2019 Maturity). The comparison is summarized with plots of spawning biomass, recruitment and fisheries depletion in Figure A11.

(a) Estimated Spawning Potential



(b) Estimated Recruitment



(c) Estimated Fishing Depletion

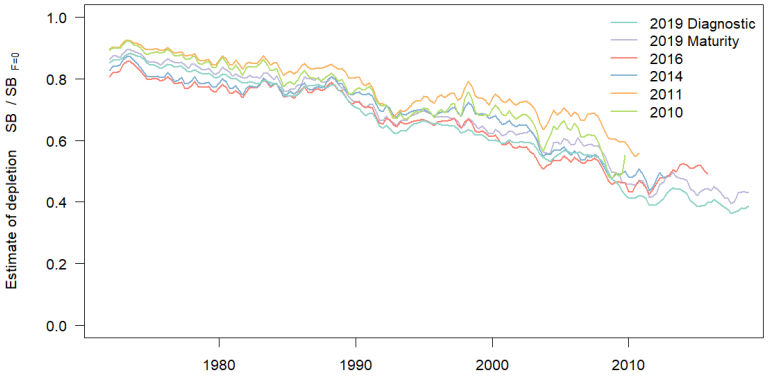


Figure A11: Retrospective comparison of spawning potential and fisheries depletion ($SB/SB_{F=0}$) for the diagnostic/reference case models for the last 5 stock assessments of skipjack in the WCPO and a model that assumed the 2016 maturity for the 2019 diagnostic model (2019 Maturity).

A.4 Sensitivity analyses reference points

In the current assessment, inferences about stock status and recommendations for management advice are based on the structural uncertainty grid, rather than the diagnostic 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 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, along with those of the diagnostic model for context:

Table A2: Summary of reference points for the set of one-off sensitivities.

	Diagnostic	5Region	H0.65	H0.95	Mix2	GrowthLow	GrowthHigh	Length50	Length200	WtdMaturity
C_{latest}	1755264	1748999	1755256	1755261	1755735	1753469	1755478	1754331	1756997	1755264
$Y_{F_{recent}}$	1797600	1660000	1977200	1699200	1876800	1848800	1788400	1784800	1799200	1796800
f_{mult}	1.838	1.749	1.544	2.253	2.360	2.200	1.797	1.903	1.786	1.843
F_{MSY}	0.223	0.222	0.191	0.263	0.219	0.223	0.221	0.228	0.218	0.224
MSY	2018400	1822400	2126000	1995200	2357200	2218400	1993200	2028400	2003600	2018800
$F_{recent}/F_{\text{MSY}}$	0.544	0.572	0.648	0.444	0.424	0.455	0.556	0.525	0.560	0.543
$SB_{F=0}$	5732710	5267910	6290767	5426736	6388850	5875870	5741829	5574425	5930517	5638955
SB_{MSY}	972200	862100	1319000	708700	1175000	952000	980500	936100	1001000	957000
$SB_{MSY}/SB_{F=0}$	0.170	0.164	0.210	0.131	0.184	0.162	0.171	0.168	0.169	0.170
$SB_{latest}/SB_{F=0}$	0.367	0.320	0.335	0.388	0.444	0.403	0.363	0.369	0.364	0.367
$SB_{latest}/SB_{\text{MSY}}$	2.165	1.957	1.596	2.970	2.416	2.489	2.123	2.197	2.155	2.160
$SB_{recent}/SB_{F=0}$	0.384	0.346	0.350	0.405	0.476	0.425	0.379	0.394	0.375	0.384
$SB_{recent}/SB_{\text{MSY}}$	2.263	2.112	1.668	3.105	2.589	2.622	2.218	2.346	2.222	2.261