SCIENTIFIC COMMITTEE
TWENTY-SECOND REGULAR SESSION
11-19 August 2026
Apia, Samoa (Hybrid)

STOCK ASSESSMENT OF YELLOWFIN TUNA IN THE WESTERN AND CENTRAL PACIFIC OCEAN: 2026

WCPFC-SC22-2026-SA-WP05

11 August 2026

Submitted by A. Magnusson[1], T. Teears[1], J. Hampton[2], N. Davies[3], K. Kim[1], C. Castillo Jordan[1], T. Peatman[4], J. Scutt Phillips[1], P. Hamer[1]

[1] Oceanic Fisheries Programme of the Pacific Community
[2] Thunnidae Consulting
[3] TeTakina Ltd
[4] Shearwater Analytics

NOTE: This paper is a final draft version as it has not had a thorough proofread, but all the essential content is included. A rev 1 will likely be issued with edits and cosmetic improvements in a few days time.

Contents

1 Executive Summary

This paper describes the 2026 stock assessment of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. An additional three years of data were available since the previous assessment in 2023, and the model extends through to the end of 2024. The assessment uses the same 5 region spatial structure, but with revised fisheries definitions. Readers are also referred to SC22-SA-IP09 (Hampton et al., 2026b) which includes exploration of a simpler single region model. New developments to the stock assessment, many of which have been in response to comments on the 2023 assessment by SC19, the yellowfin tuna assessment peer review (Punt et al., 2023), and discussion at the 2026 Pre-assessment Workshop, include:

- Revised fisheries definitions to better suit the 5 region spatial structures.
- Conversion of weight data to lengths outside the model and addition of length data for longline fisheries, with stricter size composition data filtering and enhanced reweighting approach (Peatman et al., 2026b).
- Global sdmTMB CPUE analysis used for regional scaling, with new MFCL regional scaling penalty (Teears, 2026; Davies et al., 2026a).
- Separate sdmTMB CPUE analyses for each model region, removes need for shared selectivity and catchability among CPUE indices (PAW and yellowfin tuna peer review recommendation) (Hamer, 2026; Teears, 2026).
- Time varying CV on CPUE regional indices.
- External M-scale assumption rather than internal estimation (PAW recommendation to improve model stability).
- Tag reporting rate penalties informed by increased tag seeding (Peatman et al., 2026a).
- Substantial increase in CAAL data, including north Pacific samples (Wakefield et al., 2026; Hasegawa et al., 2026).
- Application of the orthogonal polynomial recruitment parameterisation (Davies et al., 2026a).
- Uncertainty ensemble rather than orthogonal grid.

This assessment is supported by the analysis of catch and effort data for longline fisheries to provide regional abundance indices (Teears, 2026), revised analysis of tagger effects and tag reporting rates (Peatman et al., 2026c,a), enhanced and improved size composition data analyses and preparation, include the weight data to length conversion work (Peatman et al., 2026b), tag mixing simulation modeling (noting it was conducted in full but deemed not ready to apply to the assessment at this stage) (Scutt Phillips et al., 2026) and developments to MFCL software (Davies et al., 2026b). An extensive process of stepwise model development was conducted to transition from the 2023 to the 2026 assessment model and new data. The stepwise models can be explored at this link: https://ofp-sam.shinyapps.io/yft-2026-stepwise/

The 2026 yellowfin stock assessment estimated the median $SB_{\text{recent}}/SB_{F=0}$ across the 48 model uncertainty ensemble, including estimation uncertainty, to be 0.44, where *recent* is the period 2021–2024. No model estimates from the ensemble were below the LRP of $SB_{\text{recent}}/SB_{F=0} = 0.20$. Median $F_{\text{recent}}/F_{\text{MSY}}$ is 0.70, with no models indicating $F_{\text{recent}} > F_{\text{MSY}}$. The ensemble indicates the stock is not overfished, nor subject to overfishing based on the stock status reference points used by the WCPFC.

Across the 48 models with suitable convergence in the uncertainty ensemble, the most important factors when evaluating stock status uncertainty for depletion ($SB_{\text{recent}}/SB_{F=0}$) were the assumed values for the natural mortality scale (most influential), the choice of whether to activate or deactivate the estimated reporting rate adjustments within the mixing periods (second most influential) and the tag mixing period assumption (third most influential). The data weighting divisors for the length data also had some influence on $F_{\text{recent}}/F_{\text{MSY}}$.

The general conclusions of this assessment are as follows:

General trends

- The spawning potential of the stock has become more depleted across all model regions until around 2010, after which it has, overall, become more stable, but with differences among model regions. The most notable region level trend is the recent declining trend estimated for region 2.
- Average fishing mortality rates for juvenile and adult age-classes have increased throughout the period of the assessment until around 2010, after which they have stabilised, but with high inter-annual variability, particularly for juveniles.
- Juveniles continue to experience considerably higher fishing mortality than adults.

Stock status management quantities

- The median depletion from the uncertainty grid for the recent period (2021-2024; $SB_{\text{recent}}/SB_{F=0}$) is estimated at 0.44 (80 percentile range, including estimation uncertainty, 0.38–0.53, min 0.30, max 0.60).
- Zero of the 48 models in the ensemble had $SB_{\text{recent}}/SB_{F=0} <$ LRP 0.2.
- CMM 2025-02 contains an objective to maintain the spawning biomass depletion ratio at or above the average $SB_t/SB_{t,F=0}$ for 2012–2015. Across the 48 ensemble models this average was 0.424. The corresponding average for 2021–2024 was 0.411, giving a recent-to-2012–2015 spawning depletion ratio of 0.97; stock status was therefore slightly below the objective.
- Recent (2020–2023) median fishing mortality relative to fishing mortality to achieve MSY

$F_{\text{recent}}/F_{\text{MSY}}$ was 0.70 (80 percentile range, including estimation uncertainty 0.61–0.80, min 0.52, max 0.87).

- The key stock status estimates from the 2026 assessment are similar to the 2023 assessment. Based on the uncertainty ensemble presented the stock would be considered 'not overfished' (less than 20% probability of being below the LRP), and 'not undergoing overfishing' based on the median $F_{\text{recent}}/F_{\text{MSY}} < 1$.

Recommendations

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:

1. Closer examination of data conflict.
2. Evaluation of alternative regional structure, including single-region models.
3. Exploration of fixing vs. estimating the L_1 von Bertalanffy growth parameter.
4. Schedule a workshop focused on tag modeling options.
5. SPC to work with Indonesia through the WPEA project to adjust their historical large fish handline catches.

2 Introduction

This paper presents the 2026 stock assessment of yellowfin tuna (*Thunnus albacares*; YFT) in the western and central Pacific Ocean (WCPO; west of 150°W). Assessment of WCPO yellowfin tuna has been conducted regularly since the late 1990s (Hampton and Kleiber, 2003; Hampton et al., 2005; Langley et al., 2011; Davies et al., 4 08; Tremblay-Boyer et al., 2017; Vincent et al., 2020; Magnusson et al., 2023). The objectives of the 2026 yellowfin tuna assessment are to estimate population level parameters which indicate the stock status and impacts of fishing, such as time series of recruitment, biomass, biomass depletion and fishing mortality. 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 the stock assessment framework MULTIFAN-CL (version 2.2.7.9)[5] Fournier et al., 1998; Hampton and Fournier, 2001; Davies et al., 2026c. MFCL 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"[6] information components (penalties).

Each new assessment of a WCPO tuna stock typically involves updates to fishery catch, effort (for CPUE analysis), and size composition data, updates to tag-recapture data when tagging data is used, implementation of new features in the MFCL modeling software, changes to preparatory data analysis, such as CPUE standardisations, and consideration of new information on biology, population structure and potentially other population parameters. These changes are an important part of efforts to continually improve the modeling procedures and more accurately estimate fishing impacts, biological and population processes and quantities used for management advice. Advice from the Scientific Committee (SC) on previous assessments, and the annual SPC Pre-assessment Workshops (PAW) (Hamer, 2026) guide this ongoing process. Changes to aspects of an assessment can result in changes to the estimated status of the stock and fishing impacts. Each new assessment aims to provide the best possible assessment of stock status with the time and data available.

The assessment uses an 'uncertainty ensemble' of models as the basis for management advice. The uncertainty ensemble is a group of models that are run to capture uncertainty in stock status estimation that relates to uncertainty in biological and structural assumptions, data inputs and data treatment. The uncertainty models are generally based on one-off sensitivity models of a diagnostic model to indicate which uncertainties have notable effects on the estimates of key model parameters and management quantities. The variation in estimates of the management quantities across the uncertainty ensemble represents the uncertainty in stock status and should be considered carefully by managers. Structural or "model" uncertainty is usually more important than the uncertainty in

[5] https://github.com/PacificCommunity/multifan-cl/

[6] Note that any mention of a "prior" in this report does not refer to a prior in the Bayesian sense, though the effect on the parameter estimate is similar, but rather a penalty placed on the likelihood such that the estimated parameter does not deviate too much from the specified "prior" value. The magnitude of the deviation from the "prior" is dependent on the information content of the data and the strength of the likelihood penalty applied.

the estimation of the parameters from individual models, referred to as "estimation uncertainty", however both are taken into account when documenting the uncertainty in management quantities provided by the assessment.

Notable new features of the 2026 assessment are summarised below and this assessment report should be read in conjunction with several supporting papers, specifically the paper on CPUE analyses and other data inputs (Teears, 2026), the paper on size composition data preparations and weighting (Peatman et al., 2026b), the papers on tag reporting rates and tagger effects estimations (Peatman et al., 2026c,a), the paper on tag mixing estimation (Scutt Phillips et al., 2026), the paper on MFCL developments (Davies et al., 2026b) and the paper on conceptual models of bigeye and yellowfin population structure (Hamer et al., 2023). Planning and approaches for this assessment were informed by the discussion at the 2026 SPC PAW (Hamer, 2026). Several other papers are also relevant to this assessment including the single region, no tag data, fleets-as-areas model experiments for yellowfin and bigeye tuna (Hampton et al., 2026b).

Some of the more notable changes to the analysis used in this assessment include the following, which are discussed in more detail in relevant sections of this report.

- Revised fisheries definitions to better suit the 5 region spatial structures.
- Conversion of weight data to lengths outside the model and addition of length data for longline fisheries, with stricter size composition data filtering and enhanced reweighting approach (Peatman et al., 2026b)
- Global sdmTMB CPUE analysis used for regional scaling, with new MFCL regional scaling penalty (Teears, 2026; Davies et al., 2026a)
- Separate sdmTMB CPUE analyses for each model region, removes need for shared selectivity and catchability among CPUE indices (PAW and yellowfin tuna peer review recommendation) (Hamer, 2026; Teears, 2026)
- Time varying CV on CPUE regional indices
- External M-scale assumption rather than internal estimation (PAW recommendation to improve model stability)
- Tag reporting rate penalties informed by increased tag seeding Peatman et al. (2026a)
- Substantial increase in CAAL data, including north Pacific samples (Wakefield et al., 2026; Hasegawa et al., 2026)
- Application of the orthogonal polynomial recruitment parameterisation
- Uncertainty ensemble rather than orthogonal grid

3 Background

3.1 Stock structure

Yellowfin tuna is distributed across the Pacific, Indian and Atlantic Oceans in a continuous band between approximately 45° north and 45° south of the equator (Grewe et al., 2015; Moore et al., 0 10). Genetic studies suggest that populations in the three major oceans are largely separate (Pecoraro et al., 8 12; Moore et al., 0 10), although connectivity between yellowfin spawning areas in the Indian Ocean and populations in the Atlantic Ocean near south Africa has been detected (Mullins et al., 8 12). Grewe et al. (2015) showed strong genetic differences between samples from Baja California in the eastern Pacific, Tokelau in the central Pacific and the Coral Sea in the western Pacific. Similarly, Pecoraro et al. (8 12) found genetic differences between populations in the far eastern and western Pacific Ocean. Evidence of finer scale genetic structure of yellowfin in the western and central Pacific Ocean is less clear and varies between different studies (Aguila et al., 2015; Pecoraro et al., 8 12; Anderson et al., 9 07; Evans et al., 2019). An earlier genetic study by Ward et al. (1994) proposed the existence of eastern and western Pacific sub-populations separated at around 150°W. Observations of the distribution of yellowfin tuna larvae indicate that spawning occurs broadly throughout the central and western tropical Pacific, with spawning all year in the equatorial region, and seasonally in warmer months to the north and south (Nishikawa et al., 1985; Ijima and Jusup, 3 04). Studies using otolith chemistry have suggested that populations at sub-regional scales may be sourced predominantly from local spawning (Wells et al., 2 08; Rooker et al., 6 05; Proctor et al., 2019). The most recent otolith chemistry study provided good evidence that majority of small (30–40 cm) juvenile yellowfin captured around Japan originated from very small juveniles ($<$ 10 cm) sampled further south in the western Pacific tropical region (Satoh et al., 6 29). This suggests dispersal/movement from the equatorial western Pacific spawning areas via the western boundary current (Kuroshio) is important for juvenile recruitment around Japan.

The results of genetic studies are broadly consistent with tag/recapture data in suggesting that mixing between the far western and far eastern Pacific Ocean regions is limited (Moore et al., 0 10; Hamer et al., 2023). The extensive tag/recapture data available since 1989 shows that longitudinal movements among the equatorial regions of the central and western Pacific can be extensive but latitudinal movements to and from the tropical/sub-tropical latitudes may be less so. The longitudinal movements and continuous distribution across the Pacific suggests an isolation by distance mechanism is responsible for the genetic differences observed between the west and east Pacific populations. Despite the significant tagging and recent genetic studies, there remains considerable uncertainty on sub-regional population structure in the WCPO, in particular the spatial connectivity between spawning areas, early life stages, and recruitment to size classes vulnerable to fishing in different regions. This is an important area for further research. Reviews of the status of knowledge of yellowfin tuna population structure is available in Moore et al. (0 10); Hamer et al. (2023), and conclude that the weight of evidence from both genetic and non-genetic studies supports the presence of discrete stocks of yellowfin tuna in the EPO and WCPO, as well as the potential for finer-scale spatial structuring within each of these regions. For this assessment, the stock within the domain of the model area (essentially the WCPO, west of 150°W) has been considered as a discrete stock that exhibits the same biological traits (Langley et al., 2011; Davies et al., 4 08; Tremblay-Boyer et al., 2017; Hamer et al., 2023).

Over time, the spatial complexity of the modeling of the yellowfin stock in the WCPO increased. In the 2011 assessment the model domain was divided into 6 regions. After a review of the bigeye assessment in 2012 (Ianelli et al., 2012), a 9 region model was implemented in 2014 for both yellowfin and bigeye (Davies et al., 4 08; Harley et al., 4 08; McKechnie et al., 2014). This 9 region structure continued until the review by Hamer et al. (2023) suggested alternative simplified spatial stratifications that could be considered for yellowfin tuna assessment in the WCPO. The 2023 assessment subsequently adopted a simplified 5 region structure, which had better model performance and convergence properties than the 9 region structure (Magnusson et al., 2023). The same 5 region structure was adopted for this assessment Figure 1.

3.2 Biological characteristics

Yellowfin tuna can reach a maximum fork length (FL) of around 180 cm and live for up to 15 years, although most fish aged to date have been less than 10 years old (Itano, 2000; Farley et al., 2020; Wakefield et al., 2026) and the maximum validated age is 13 years. Growth of the juvenile stage is particularly fast and they can reach a fork length of around 20–30 cm by three months of age and approximately 50 cm by 1 year (Farley et al., 2020). Length at 50% maturity in the WCPO is at around 100–110 cm (Itano, 2000) which equates to around 2 years of age. In this assessment, for the purpose of computing the spawning biomass, we assume a fixed maturity schedule consistent with the observations of Itano (2000) (see Vincent and Ducharme-Barth, 2020 for details). Yellowfin tuna are thought to spawn opportunistically throughout the Pacific in waters warmer than 24°C (Itano, 2000; Reglero et al., 4 03). Larval stages are found widely in surface waters throughout the central and western Pacific (Nishikawa et al., 1985; Servidad-Bacordo et al., 2012; Ijima and Jusup, 3 04) and at least some spawning appears to occur year-round in the WCPO. However, understanding of spatio-temporal variation in spawning fraction is limited. Important areas for spawning are thought to occur in the Banda Sea in Indonesia, the north-western Coral Sea, the eastern and southern Philippines, northeast of Solomon Islands, and around Fiji (McPherson, 1991; Gunn et al., 2002; Servidad-Bacordo et al., 2012; Ijima and Jusup, 3 04). Juvenile yellowfin (several months of age) are prevalent in commercial fisheries in the Philippines and eastern Indonesia (Hare et al., 2025), suggesting this region is important for the juvenile stages, but they are also found more widely in the equatorial Pacific.

Growth parameters can be highly influential in the estimates of management parameters by stock assessment models, and the most recent stock assessments of yellowfin in the WCPO identified growth as an important uncertainty requiring further research (McKechnie et al., 2017; Vincent et al., 2018, 2020). Considerable work was carried our on developing external growth curves from otoliths and otolith-tag integrated growth models for the previous assessments (Eveson et al., 2020; Farley et al., 2020). Those studies resulted in different estimates of growth rate parameters to those previously estimated by the MFCL model based on length composition modal progression. Annual aging of yellowfin using otoliths has recently been validated using the bomb radiocarbon method (Andrews et al., 8 19). The 2023 assessment used the available age-length data as conditional-age-at-length data inputs to the model and estimated growth internally. This was strongly recommended by yellowfin assessment peer review (Punt et al., 2023). Understanding of spatio-temporal variation in growth is limited and is insufficient to consider such effects in the current assessment, nor is it feasible to implement spatially varying growth using the MFCL model framework (but see (Hampton et al., 2026a)). The 2026 assessment will incorporate a substantial increase in the amount of conditional-age-at-length data, both for the equatorial region and the region north of 10° W for which no sample were available in the previous assessments (Wakefield et al., 2026; Hasegawa et al., 2026).

Natural mortality (M) rate of yellowfin tuna varies with size/age (Hampton, 2000). Mortality is highest for the smaller juveniles and estimated to be lowest for the pre-adult stage (50–80 cm FL) 0.6–0.8 yr^{-1} (Hampton, 2000). After reaching maturity it is thought that mortality increases with age, particularly in females. Sex ratios of yellowfin tuna are commonly observed to be biased toward males at larger sizes, and it is thought that this may relate to the higher mortality rates of mature females due to the physiological stresses related to spawning or a combination of this and different growth rates between males and females at older ages (Schaefer, M.B. et al., 1963; Hampton, 2000; Fonteneau, 2002; Sun et al., 2006; Zhu et al., 8 08). For the purpose of computing the spawning potential and mortality schedules, data on sex ratio at length is important. For the previous assessment data on sex ratios collected by observers in the WCPO was incorporated into the estimation of spawning potential and M at age (Vincent and Ducharme-Barth, 2020). We utilise the same data source in this assessment. The 2020 yellowfin assessment conducted a life-history based meta-analysis of natural mortality for yellowfin (Vincent and Ducharme-Barth, 2020) indicating an envelope of potential quarterly M rates of lower 95% confidence interval (0.1100), mean (0.1298) and upper 95% confidence interval (0.1495). More recently natural mortality of yellowfin has been reviewed by Hoyle et al. (3 01), and this review is taken into account when considering options for natural mortality in the current assessment. The 2023 assessment attempted to estimate M internally and produced a best estimate of adult (asymtotic, age 40 quarters) quarterly M of 0.119, consistent with the 2020 meta-analysis (Magnusson et al., 2023).

3.3 Fisheries

Yellowfin tuna is an important component of tuna fisheries throughout the WCPO. They are harvested with a wide variety of gear types, from small-scale artisanal fisheries in Pacific Island and southeast Asian waters to large, distant-water longliners and purse seiners that operate widely in equatorial and tropical waters. Purse seiners catch a wide size range of yellowfin tuna, however, smaller yellowfin often dominate catches associated with FADs (fish aggregation devices), whereas the longline fisheries and east Asian handline fisheries (Indonesia/Philippines/Vietnam) takes mostly larger adult fish (Vidal and Hamer, 2020; Vidal et al., 2026; Hare et al., 2026).

The annual yellowfin tuna catch in the WCPO increased from 100,000 mt in 1970 to between 700,000 and 750,000 mt in recent years, mainly due to increased catches in the purse seine fishery and the mixed gear 'other' fishery in Indonesia/Philippines/Vietnam (Hare et al., 2026, 2025). The 2024 catch was approximately 740,000 mt slightly lower than 2023 (Figure 2). Purse seiners harvest the majority of the yellowfin tuna catch (around 50–55% since 2018), while the longline fleet accounted for around 10–15% of the catch in recent years, primarily in the equatorial regions (Figure 3, Hare et al., 2025, 2026). The remainder of the catch is dominated by the domestic fisheries of the Philippines and Indonesia, principally catching smaller individuals using a variety of small-scale gear types (e.g. pole-and-line, ringnet, gillnet, handline and seine net), and large fish on handline (Vidal et al., 2026). Small to medium sized purse seiners based in those countries also catch fish of sizes more typical of the purse seine fisheries elsewhere.

Yellowfin tuna typically represent 15–20% of the overall purse-seine catch in recent years and may contribute higher percentages of the catch in individual sets. Yellowfin tuna are often directly targeted by purse seiners, especially within unassociated schools (free schools) that are comprised of larger yellowfin compared to those associated with FADs (associated sets). Unassociated sets account for the majority of purse seine sets, however, many of these sets fail (skunks) and when considering successful sets only, the numbers of associated and unassociated sets are similar (Hare et al., 2025).

Since 2010, annual catches of yellowfin tuna by longline vessels in the WCPO have varied between approximately 74,000 to 105,000 mt (Hare et al., 2025). The highest longline catch recorded was around 125,000 mt in 1980 (Figure 2). Annual catches from the domestic fisheries of the Philippines and eastern Indonesia area are highly uncertain, particularly prior to 1990. Recent estimates for pole and line and other gears have reached approximately 220,000–260,000 mt in the last 5 years years, and for purse seine 350,000–400,000 mt (Figure 2).

Figure 4 shows the spatial distribution of yellowfin tuna catch in the WCPO for the past 10 years. Most of the catch is taken in western equatorial areas, with less catch by both purse-seine and longline toward the east. The east-west distribution of catch is strongly influenced by ENSO events, with larger catches taken east of 160 °E during El Niño episodes. Catches from outside the equatorial region are relatively minor (5%) and are dominated by longline catches south of the equator and purse-seine and pole-and-line catches in the north-western area of the WCPO (Figure 3 and Figure 4).

Improved catch statistics in recent years for the Indonesian, Philippines, and Vietnamese fleets (Vidal et al., 2026) have resulted from collaborative work between the fisheries agencies of these countries and the SPC. In some instances data are available at the individual fisheries level (e.g., longline or large-fish handline), but often statistics are aggregated across a variety of gears that typically catch small yellowfin tuna, e.g., ring-net, handline, and troll. Data for these fisheries have been included in this assessment.

4 Data compilation

4.1 General notes

Data used in the yellowfin tuna stock assessment using consist of catch, effort (for CPUE analysis), length & weight-frequency data for the fisheries defined in the analysis, and tag-recapture data. Conditional-age-at-length data (CAAL) are also used directly as data in the assessment model, as was recommended by the peer review of the yellowfin tuna assessment (Punt et al., 2023). Improvements in these data inputs are ongoing and readers should refer to the companion papers highlighted at the end of Section 2 for detailed descriptions of how the data and biological inputs were formulated as only brief overviews are provided below. A summary of the data available for the assessment is provided in Figure 5.

4.2 Spatial stratification

The geographical area considered in the assessment corresponds to the WCPO (from 50°N to 40°S between 120°E and 150°W) and oceanic waters adjacent to the east Asian coast (110°E between 20°N and 10°S). The eastern boundary of the assessment excludes the WCPFC Convention area component that overlaps with the Inter American Tropical Tuna Commission (IATTC) area. We began the stepwise model development with the previous 5 region model structure and maintained this structure for all models in the assessment (Magnusson et al., 2023)(Figure 1). The 5 region structure was discussed again at the PAW and was supported to continue for this spatial assessment (Hamer, 2026). The 5 region structure maintains the region around the Papua New Guinea and Solomon Islands area to accommodate the longer residency of yellowfin in these archipelagic waters and the tagging data and associated mixing period assumptions. It has large tropical equatorial, subtropical-temperate northern and southern regions, and a region isolating the Indonesia/Philippines/Vietnam predominantly archipelagic waters. The 5 region structure was adopted in the 2023 assessment, after the previous 9 region structure performed poorly (Vincent et al., 2020). The 5 region structure was supported by the review by Hamer et al. (2023).

4.3 Temporal stratification

The time period covered by the assessment is 1952–2024 which includes all significant post-war tuna fishing in the WCPO. Within this period, data were compiled into quarters (1; Jan–Mar, 2; Apr–Jun, 3; Jul–Sep, 4; Oct–Dec). As agreed at SC12, the assessment does not include data from the most recent calendar year as this is considered incomplete at the time of formulating the assessment inputs. Recent year data are also often subject to significant revision post-SC, in particular the longline data on which this assessment greatly depends.

4.4 Definition of fisheries

MFCL requires "fisheries" to be defined that consist of relatively homogeneous fishing units. Ideally, the defined fisheries will have selectivity and catchability characteristics that do not vary greatly over time and space. For most pelagic fisheries assessments, fisheries are typically defined according to combinations of gear type, fishing method and region, and for some, also flag or fleet. There revised fisheries structure for the 2026 assessment resulted in 36 fisheries, compared to 41 in the 2023 assessment (Table 1). There two fishery types: "index fisheries", that are used for generating indices of abundance (see further below), and "extraction fisheries" that account for the catches removed from the stock. There are 31 extraction fisheries including longline, purse seine, pole and line, handline and various miscellaneous gears fisheries in the Indonesia/Philippines/Vietnam region. There are 5 "index fisheries", one for each model region, all based on longline gear. A graphical summary of the availability of data for each fishery used in the assessment model is provided in Figure 5.

Equatorial purse seine fishing activity was aggregated over all nationalities, but stratified by region and set type, in order to sufficiently capture the variability in fishing operations and selectivity of different purse seine set types. Set types were grouped into associated (i.e. log, FAD, whale, dolphin, and unknown set types) and unassociated (free-school) sets. Additional fisheries were defined for pole-and-line fisheries, handline fisheries, and miscellaneous fisheries (gillnets, ringnets, hook-and-line etc.) in the western equatorial area. In regions 2 and 4 extraction longline fishing was separated into distant water and offshore components to account for the apparent differences in fishing practices and selectivity for these fleets in these regions.

Longline index fisheries: The catch-conditioned approach allows the specification of "index fisheries" that are used to provide standardised CPUE indices of abundance. Index fisheries are akin to "survey fisheries" as described for other software such as Stock Synthesis, and may be the same fisheries as the extraction fisheries, but when used as index fisheries they do not take any catch, and must have effort data to allow modeling of CPUE. For this assessment one index fishery is defined for each model region as a composite fishery composed of longline fisheries operating under various flags in each assessment region (Teears, 2026). In the 2023 assessment a single sdmTMB CPUE standardisation model was conducted across the entire assessment region (i.e. global model), the quarterly predictions were then aggregated and summed by model regions to generate the region specific abundance indices (Teears et al., 2023), that were assumed to have constant catchability and selectivity. This allowed MFCL to use the differences in relative abundance indices among model regions as information on the relative size of the vulnerable population among model regions (i.e. regional scaling). The assumption of shared catchability and selectivity is problematic and can create problems with adequately fitting the size compositions for each index fishery. The yellowfin assessment peer review recommended freeing up these assumptions of shared catchability and selectivity among CPUE abundance indices (Punt et al., 2023). This was discussed in depth at the PAW, where it was recommended to run a global sdmTMB model using the long-term Japanese operational data, use this global model for generating regional abundance indices that are assumed to have shared catchability and selectivity, but only use these for information on relative vulnerable abundance among model regions (i.e. only for regional scaling). For the region specific abundance indices, conduct individual sdmTMB CPUE models for each model region. Separate regional CPUE models could perhaps better account for regional differences in CPUE covariate relationships, and the resultant abundance indices can have their own selectivity without assuming constant catchability among model regions (Hamer, 2026). This approached was adopted and required considerable additional work in the CPUE analyses (Teears, 2026), and also modifications/enhancement to the MFCL code to allow one set of CPUE (i.e. global model) to be used to generate a regional abundance scaling penalty, and the other set of region specific CPUE indices to provide the time series abundance indices for the model to fit to Davies et al. (2026b). Furthermore, specification of the time period from the global indices to use to calculate the regional scale penalty is required. The PAW recommended using a time period where data coverage was most broad across the global model domain, this period was 1965-1969 Teears (2026). The MFCL code was also modified to input global CPUE time series of different start and end dates so the scaling penalty could be determined for different time periods, noting the the penalty is applied only to that time period Section 4.5.3.

The regular longline extraction fisheries are based on the same data set, but are disaggregated into the longline fisheries defined in Table 1. The size composition data (length and weight-frequency) for the extraction fisheries is assumed to represent the actual composition of the removed fish for any space-time strata, and in the data preparation process are weighted by the catch in order to represent the fisheries extractions at the spatial (region) and temporal (quarter) resolution of the model (Peatman et al., 2026b). However, for the index fisheries, while the same aggregation process is conducted, the size data are weighted by standardised CPUE (rather than by catch) so that the size data are more representative of the abundance of the underlying population in each region and time period. Further, because the size data for the index and extraction fisheries are effectively being used twice (but weighted differently), the likelihood weighting for the size composition is adjusted such that the original intended weight (effective sample size) in the likelihood is preserved.

4.5 Catch and effort data

4.5.1 General characteristics

Catch and effort data were compiled according to the fisheries defined in Table 1. Catches by the longline fisheries were expressed in numbers of fish, and catches for all other fisheries expressed in weight (mt). This is consistent with the form in which the catch data are recorded and reported for these fisheries. The catches are aggregated at $5° \times 5°$ and quarterly resolution, with the aggregation process either conducted by SPC, where operational data is available to inform this, or by the particular countries following statistical procedures that are reported to the Commission. For some fisheries, notably those in region 2 — Indonesian/Philippines/Vietnam — operational information on quarterly or spatial patterns in catches is poor so the annual catches are aggregated evenly across quarters and spatial cells. This is done by SPC.

In the catch-conditioned model, effort is not essential but is required (at least for a recent period of time) for projection analyses involving fisheries managed under effort rather than catch controls. The effort data are necessary to derive recent estimates of catchability for running the effort based projections. In this case the main industrial purse seine fisheries operating in the tropical region (i.e., regions 2, 3, 4) are managed under effort control. Effort data for these purse seine fisheries are defined as number of sets specified by set type (associated or unassociated), and are included for the last 12 quarters to facilitate projections. The period of 12 quarters is consistent with previous projections using catch errors models. For this assessment several other fisheries also have effort included to allow effort based projection for management purposes, these are the longline extraction fisheries, with effort measured as numbers of hooks per set, and the Japanese pole and line fishery with effort measured in vessel fishing days.

Total annual catches by major gear categories for the WCPO are shown in Figure 2 and a regional breakdown is provided in Figure 3. The spatial distribution of catches over the past ten years is provided in Figure 4. Discarded catches are estimated to be minor and were not included in the analysis. Catches in the northern region are low and highly seasonal and the annual catch has been relatively stable over much of the assessment period. Most of the catch occurs in the tropical regions (2, 3, and 4).

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

- The steady increase in total yellowfin catch over most of the assessment period, with the highest overall catches reported in the most recent years (2021, at approximately 760,000 mt).
- The steady increase in catch for the domestic fisheries of Indonesia and the Philippines (region 2) since 1970, where mostly small juveniles are taken, and more significant increase in the catches over the last 15 years. Some of this trend can be related to improved information to estimate catches.
- The relatively stable and low catches of yellowfin in the northern and southern temperate regions by longline vessels (regions 1, and 5).
- The development of the equatorial purse-seine fisheries from the mid-1970s, and corresponding increased catches, particularly in equatorial regions, with the purse seine catch recently at 3–5 times higher than the longline catch.
- Large changes in the purse seine fleet composition and the increasing size and likely efficiency of the fleet.
- The notable increase in the handline catch reported for 2024, which now exceed longline. This is the result of changes in how ID provided their catch statistics in 2026, wherebye there

was a change in how yellowfin catch was allocated between the small and large fish handline fisheries. ID determined that a large portion of the ID catch that was previously reported as smaller fish handline gear (part of the 'other' category) should be allocated to the larger fish handline (see Vidal et al. (2026)). There is yet to be a historical adjustment, and this will be important for future assessments. The single terminal year increase in the ID large fish handline catch adds to Philippine catches, it will have minor influence on the current assessment, but the historical time series needs an appropriate adjustment as a priority under the WPEA project. This will impact historical estimates.

4.5.2 Purse seine

For the industrial purse seine fisheries predominantly operating in tropical regions 3, and 4, catch by species within each set type (associated or unassociated) is determined by applying estimates of species composition from observer-collected samples to total catches estimated from raised logsheet data (Hampton and Williams, 2016; Peatman et al., 2021). For the Japanese (JP) fleet for which there is greater confidence in species-based reporting, reported catch by species is used. Purse seine catch for Philippines (PH) and Indonesian (ID) domestic purse seine fisheries, predominantly operating in region 2, was derived from raised port sampling data provided by these countries. We note that the COVID-19 pandemic resulted in low observer coverage of the purse seine fleets from mid 2020 to 2022 and purse seine catch estimates are less reliable over these years.

4.5.3 Longline

For the longline fisheries catches in number of fish by species are derived from raised logbook data or annual catch estimates provided by specific countries. Effort is in terms of hooks per set.

The index fishery CPUE time series for the 2026 assessment were derived from the operational longline dataset for the Pacific region Figure 6. This dataset is an amalgamation of operational level data from the distant-water fishing nations (DWFN), United States, Australian, New Zealand and Pacific-Island countries and territories (PICTs) longline fleets operating in the Pacific basin. It represents the most complete spatiotemporal record of longline fishing activity in the Pacific, spanning from 1952 through to the present and is the result of collaborative ongoing data-sharing efforts from many countries. This data-set was first created in 2015 in support of the Pacific-wide bigeye tuna stock assessment (McKechnie et al., 2015), and has subsequently been used in each bigeye and yellowfin tuna assessment (McKechnie et al., 2017; Vincent et al., 2018; Ducharme-Barth et al., 2020a; Day et al., 2023). Since 2017 spatiotemporal approaches (Thorson et al., 2015) have been used for CPUE modeling in WCPFC stock assessments (Tremblay-Boyer et al., 2017; Ducharme-Barth et al., 2020b; Teears et al., 2023).

Consistent with the 2023 assessment, the CPUE modeling was conducted with sdmTMB (Anderson et al., 2 03). A detailed description of the methods for generating the spatiotemporal abundance indices is provided in Teears (2026). However, some aspects are worth noting here. Firstly, the global analysis used for the regional scaling penalty takes account of habitat viability by not including predictions in cells when SST is below a specified thresholds. The threshold was determined as 19°C based on the probability of positive CPUE observations being greater than 10% (Teears, 2026). Applying the SST thresholds prevents inflating the relative abundance in the northern and southern regions by including spatial cells that have predictions of positive abundance that are unlikely in reality. The approach was a recommendation from the PAW.

The CPUE modeling for the 2026 assessment made some key improvements over the 2023 version, largely following detailed discussion and recommendations from PAW (Hamer, 2026). The 2026 analysis implemented encounter and positive component sub-models separately to allow for differences in model formulas; specifically to allow the number of hooks to be included in the encounter sub-model. The inclusion of vessel ID as a random effect to account for vessel effects was another important improvement in the analysis, and the number of samples included in the sub-sampling of longline CPUE data for region specific indices was increased. Finally effort creep (positive increase in effective effort over time) was included as an adjustment on predicted regional CPUE indices. The decision at PAW was that zero effort creep was implausible, so a range of positive effort creep is explored based on the recommendation from PAW, and no zero options are included in the assessment models for management advice.

Effort creep adjustments recommended at PAW are from 1-3% per year from 1952-1976 when vessel ID information is unavailable, and post 1976, when vessel ID is included, 0.5-1.5% year. Please refer to Teears (2026) for a full description of the CPUE modeling for generating the global and region specific abundance indices.

4.5.4 Other fisheries

Effort data for the ID, PH, and VN surface fisheries and Japanese research longline fisheries are unavailable. However, as these fisheries are not part of the index fisheries, with the catch-conditioned approached effort data are not necessary for these extraction fisheries. Catch estimates for the ID/PH/VN fisheries are derived from various port sampling programmes dating back to the 1960s for ID and the PH, and early 2000s for VN (Vidal et al., 2026).

4.6 Size data

Available length-frequency data for each of the fisheries were compiled into 95 x 2cm size classes from 10–12 cm to 198–200 cm. Weight data were compiled into 200 x 1 kg size classes from 0–1 kg to 199–200 kg. Most weight data were recorded as processed weights (usually recorded to the nearest kilogram). Processing methods varied between fleets requiring the application of conversion factors to convert the available processed weight data to whole weight equivalents. Details of the various processed weight to whole weight conversion factors are provided in Hamer et al. (2026). Note that these conversions occur in the database prior to data extraction for stock assessments (Hamer et al., 2025). Data were either collected onboard by fishers, through observer programs, or through port sampling, see descriptions of data sources in Hamer et al. (2025, 2026). Each size-frequency record in the model consisted of the actual number of yellowfin tuna measured and Figure 5 provides details of the temporal availability of length and weight-frequency data (also see Teears, 2026). Note that a maximum effective sample size of 1,000 is implemented in MFCL when using the robust normal likelihood for size composition data. The effective sample size was further down-weighted as explained in Section 5.5.2. Summaries of the available size composition data by year and fishery are provided in Peatman et al. (2026b); Teears (2026).

An extensive process of preparatory size data processing is required to improve the reliability and representativeness of size composition data. This includes aggregating data according to the spatial model regions and fisheries, filtering out unreliable data sources and low sample sizes, and re-weighting the size frequencies to better represent the catches (extraction fisheries) and relative abundance (index fisheries). Only length data are available for purse seine, pole and line and the other miscellaneous gear fisheries in the Indonesia/Philippines/Vietnam region. For longline fisheries both weight and length data are available. For the 2026 assessment, based on recommendations from previous years (Teears et al., 2023; Day et al., 2023) and discussions at the PAW (Hamer, 2026), it was decided to include weight and length data for longline fisheries due to the increasing coverage of length data in recent years and the decreasing coverage of weight data (Peatman et al., 2026b). This involved the conversion of weight data to lengths external to the model, so that the final models ran entirely on length data, this is the fist time to apply this approach. A key principle used in constructing the size composition inputs was not to use weight and length data at the same time for a fishery, even if it was available. The decision was made to start replacing the weight data (weight converted to length) with actual observed length data if and when the length data had greater coverage than the weight data. The preparation of size data across the longline, purse seine, pole and line and other domestic fisheries (Indonesia/Philippines/Vietnam), including the integration of longline length data and the associated weight-length conversion process is a significant amount of work, and is fully described in the companion paper by Peatman et al. (2026b). Preparation of size data inputs has increased substantially in recent year, with increased review and scrutiny of these data and now takes up a more significant time component of data inputs preparation.

4.7 Tagging data

A reasonable amount of tagging data is available for yellowfin tuna, although it is mostly constrained to the tropical region (Figure 7, Figure 8). Information on the yellowfin tag data characteristics, the process of constructing the MFCL tagging file, and estimating tag release correction factors (tagger effects) and reporting rate priors (from tag seeding experiments) can be found in Peatman et al. (2026c,a); Teears (2026). A summary of the tagging data and maps displaying tag displacements are also in Teears (2026). Data were available from the Regional Tuna Tagging Project (RTTP) during 1989–92 (including affiliated in-country projects in the Solomon Islands, Kiribati, Fiji and the Philippines), the ongoing Pacific Tuna Tagging Programme (PTTP) that began in 2006. The

Japanese Tagging Programme (JPTP) conducted by NRIFSF and the Ajinomoto Co. Inc, over the period 2000–2020 is also included. The new tagging data for the 2026 assessment comes from PTTP. Historical (1995, 1999-2001) data from the Coral Sea tagging cruises by CSIRO (Evans et al., 2008) are no longer included due their unusual recapture properties. Tags were released using standard tuna tagging equipment and techniques by trained scientists and technicians.

After tagged fish are recaptured, there is often a delay before the tag is reported and the data are entered into the tagging databases. If this delay is significant then reported recapture rates for recent release events will be biased low and will impact estimates of fishing mortality and population scale in the terminal time periods of the assessment. For this reason, tag releases from the PTTP are included only up until the end of 2022, although there have been more tags released later than 2022 (Figure 8). The 2023 assessment included tag data until 2019. Significant numbers of tagged yellowfin have been released in recent (2020 and 2021) tagging cruises focused on bigeye and yellowfin in the central Pacific (model region 4). Recaptures from these releases are available for the 2026 assessment (Teears, 2026). Tags were released using standard tuna tagging equipment and techniques by trained scientists and technicians. Tags have been returned from a range of fisheries, having been recovered onboard or via processing and unloading facilities throughout the Asia-Pacific region.

The numbers of tag releases input to the assessment model are adjusted for a number of sources of tag loss, unusable recaptures due to lack of adequately resolved recapture data, estimates of tag loss (shedding and initial mortality) due to variable skill of taggers (i.e., tagger effects), and estimates of base levels of tag shedding and tag mortality. These adjustments are described in more detail in Peatman et al. (2026c); Teears (2026). These procedures were first described in Berger et al. (4 08) and McKechnie et al. (2016).

For incorporation into the assessment, tag releases were stratified by release region, time period of release (quarter) and the same size classes used to stratify the length-frequency data. The likelihood penalties or “priors” used for the reporting rates of the grouped tag return fisheries have been updated relative to those used in the previous assessment based on the analysis of tag seeding experiments (Peatman et al., 2026a). Tag reporting rate groupings are required as not all reporting fisheries have tag seeding experiments. These reporting rate grouping decisions necessarily require some expert judgement. For this assessment we have also excluded tag release groups with 5 or less recaptures from the estimation of reporting rates, as we felt there was insufficient information to inform model estimation of the reporting rates.

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 populations (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 Davies et al. (2026c), and summarised 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 relative to a series of reference points.

5.2 Population dynamics

The model partitions the population into spatial regions, depending on the specified spatial stratification, and 40 quarterly age-classes. The last age-class comprises a "plus group" in which mortality and other characteristics are assumed to be constant. The population is "monitored" in the model at quarterly time steps, extending through a time window of 1952–2024. The main population dynamics processes are as follows.

5.2.1 Recruitment

Recruitment is defined as the appearance of age-class 1 quarter fish (i.e. fish averaging $\sim$ 24 cm FL) in the population. Bigeye tuna spawning does not typically follow a strong seasonal pattern, especially in the tropical waters where most of the spawning occurs. Rather it occurs sporadically all year and is thought to be influenced by food availability (Schaefer et al., 2005; Sun et al., 2006; 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.

In previous yellowfin tuna assessments, recruitment was estimated using recruitment deviates as year-quarter and region-specific deviates from the mean recruitment (referred to as the 'mean+ plus deviates' approach). The 2023 yellowfin tuna assessment consisted of 1,670 parameters (of 1,920 in total) relating to recruitment, one for each year/season/region (some constrained by sum-to-zero conditions). There has been a long-standing concern that the mean plus devs approach poses a risk of over-parameterisation, with resulting poor statistical performance, including difficulties with convergence. To address these concerns, MFCL has an orthogonal polynomial recruitment (OPR) feature in which:

- A linear hierarchical structure for year, region, season, and region-season interaction levels is implemented using a Gram-Schmidt orthogonalisation basis matrix;
- A user-specified number of degrees is estimated for each of the four levels to control the extent of time-series variation; and, a common effect can be used for a specified number of terminal years of a specified level.

The system gives a large degree of user control over recruitment variability, ranging from the simplest (constant total recruitment by year equally allocated to seasons and regions, i.e. 1 parameter) to a 'fully saturated' system equivalent to the Rec-Devs approach (1600 parameters for the current assessment). We undertook initial testing of alternative OPR specifications (using AIC) of many alternative OPR structures using the 2023 assessment data. We also drew on previous work conducted for the 2025 skipjack assessment (Teears et al., 2025).

Finally, we tested various OPR specifications in order to judge what level of simplification was required for stable model behaviour and convergence, including obtaining a positive definite Hessian. Based on this, we adopted the following OPR specification:

- A close to fully saturated year effect (1 parameter per year);
- A single average seasonal effect (i.e. the overall seasonal distribution of recruitment is constant over time) per region;
- A regional effect varying over time with a 10th order polynomial (i.e. the way the overall annual regional distribution varies over time); and
- A region-season interaction effect also described by a 10th order polynomial, allowing some flexibility in the variation of the seasonal distribution of recruitment by region and the regional distribution by season over year.

Initial application of this approach to the 2026 assessment data estimated unrealistically high recruitments towards the end of the time series (this also occurred with the mean+deviates version). To control this, we set a common recruitment pattern for the terminal 2 years of the assessment and constrained this level via a penalty to be approximately equivalent to the arithmetic mean historical recruitment. This resulted in 235 recruitment parameters being estimated. This is a considerable reduction from the 1,742 parameters that would be required to be estimated under the mean+deviates approach. The recruitment pattern and other stock-assessment related quantities estimated using the Rec-Devs and OPR approaches were similar.

Spatially-aggregated (over all model regions) recruitment was assumed to have a weak relationship with annual mean spawning potential via a Beverton and Holt stock-recruitment relationship (SRR) with a fixed value of steepness (h). Steepness is defined as the ratio of the equilibrium recruitment produced by 20% of the equilibrium unexploited spawning potential to that produced by the equilibrium unexploited spawning potential (Francis, 1992; Harley, 2011). Typically, fisheries data are not very informative about the steepness parameter of the SRR parameters (ISSF, 2011); hence, the steepness parameter was fixed at a moderate value (diagnostic model: 0.80) and the sensitivity of the model results to the value of steepness was explored by setting it to lower (0.65) and higher (0.95) values. 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 1962–2023 to prevent the earlier recruitments (which may not be well estimated), and the terminal recruitments (which are not freely estimated), from influencing the relationship.

The SRR was incorporated mainly so that yield analysis and population projections could be undertaken for stock assessment purposes, particularly the determination of equilibrium- and depletion-based reference points. We therefore applied a weak penalty (equivalent to a CV of 2.2) for deviation from the SRR so that it would have negligible effect on recruitment and other model estimates (Hampton and Fournier, 2001), but still allow the estimation of asymptotic recruitment. This approach was recommended (recommendation 20) by the 2011 bigeye assessment review (Ianelli et al., 2012).

5.2.2 Initial population

The population age structure in the initial time period in each region was assumed to be in equilibrium, and the initial fishing mortality was assumed to be zero ($Z = M$). As noted above, the population is partitioned into quarterly age-classes with an aggregate class for the maximum age (plus-group, 40 quarters). The aggregate age class makes it possible for accumulation of old fish, which is likely in the early years of the fishery when exploitation rates were very low.

5.2.3 Growth

The standard assumptions for WCPO assessments fitted in MFCL were made concerning age and growth: i) the lengths-at-age are normally distributed for each age-class; ii) the standard deviations of length for each age-class are a log-linear function of the mean lengths-at-age; and 3) the probability distributions of weights-at-age are a deterministic function of the lengths-at-age and a specified weight-length relationship[7]. These processes are assumed to be regionally and temporally invariant.

In the previous assessments several approaches to growth estimation were explored, including an external otolith based growth curve with a fixed Richards growth curve based on high-readability otoliths (Farley et al., 2020), a fixed external Richards growth curve based on the same high-readability otolith dataset but with the addition of tag-recapture growth increment data (Eveson et al., 2020), an internal MFCL growth estimation using a conditional-age-at-length (CAAL) dataset from the otolith dataset including daily and annual ages (Farley et al., 2020), which also made use of the modal progressions apparent in various size composition data (Vincent et al., 2020).The peer review of the yellowfin assessment (Punt et al., 2023) carefully considered the approaches to estimating growth. They concluded that based on the otolith sampling protocols, with otoliths selected according to length (i.e., attempting to achieve sufficient numbers of otolith samples across the full length range, plus bias in otolith readability with age), that any external growth curves would likely be biased. They strongly recommended that for this type of otolith sampling regime that age data should be conditional on length and the growth estimated internally, using conditional-age-at-length data within MFCL. This approach was applied in the 2023 assessment and is continued for this assessment to estimate the von Bertalanffy growth curve internal to the MFCL model,

[7]The length-weight relationship has been updated for the current assessment based on an analysis of current and historical port sampling data under WCPFC Project 90 with bias correction applied (Hamer et al., 2025)

considering also information on growth from size composition data. We also note the substantial enhancement of the CAAL data set through the efforts of SPCs recently established schlerochronoly lab (Wakefield et al., 2026) and the contributions of Japan's FRA (Hasegawa et al., 2026).

5.2.4 Movement

Movement was assumed to occur instantaneously at the beginning of each quarter via movement parameters that connect regions sharing a common boundary. Note that fish can move between non-contiguous regions in a single time step due to the "implicit transition" computational algorithm employed (see Hampton and Fournier, 2001 and Davies et al., 2026c for details). Movement is parameterized by a pair of bi-directional parameters at each region boundary. Movement is therefore possible in both directions across each regional boundary in each of the four quarters. Each of these coefficients is estimated independently resulting in 56 estimated movement parameters for the 5 region spatial structure (2×no.region boundaries (7)×4 quarters). There are limited data from which to estimate long-term, annual variation in movement or age-specific movement rates. As such, the estimated seasonal pattern is assumed to be fixed across years and the movement parameters are invariant with respect to age. A "prior" of 0 is assumed for all movement parameters, and a low penalty is applied to deviations from the "prior".

5.2.5 Natural mortality

The instantaneous rate of natural mortality (M) consists of an average over age classes and age-specific deviations from the average M. Average M can be estimated internally by the model or specified as a fixed value. Internal estimation can be constrained by providing a prior value and a penalty weight for deviations from the prior. Age specific deviations from average M can also be estimated or input as a fixed M-at-age function. The later approach was taken in the 2020 assessments for both bigeye and yellowfin (Ducharme-Barth et al., 2020a; Vincent et al., 2020) where M-at-age was calculated using an approach applied to other tunas in the WCPO and EPO (Harley and Maunder, 2003; Hoyle, 2008; Hoyle and Nicol, 2008). The yellowfin peer review had some difficulty with tracing the sources of data for some of the inputs to the external M-at-age function applied in 2020 that requires fitting a model that depends on empirical data on the sex-ratio at length, a growth curve, a base M assumption for males, assumptions on critical lengths for inflections and a multiplier that determines the linear decline in M for young ages to the base M, plus a length at which female mortality is assumed to begin to increase. There are uncertainties in all these empirical data and assumptions, and the data for sex ratio is limited especially for larger and older fish. While the peer review generally supported the method, the construction of the external M-at-age function, notably the limitations in some of the available data required to estimate it, the complexity of the calculations and the various assumption required are problematic. Furthermore, the external M-at-age function requires re-estimation if different growth curves are applied, further complicating use of this method. After the yellowfin peer review, a Tuna Stock Assessment Good Practices workshop run by the "Center for the Advancement of Population

Assessment Methodology" (CAPAM) was held in March 2023 and provided a strong endorsement for applying a simpler Lorenzen functional form for estimation of M-at-age for tuna. Other CAPAM reviews of approaches for estimating M have also, where possible, favoured estimating M within the model, "let the data speak for themselves", while being careful to review model diagnostics and plausibility of model estimated M against external estimates. For the 2023 assessment M-at-age was estimated internal to the model with a Lorenzen function (Magnusson et al., 2023). However, the model was not stable to Jittering, and there were concerns expressed by SC19 that instability may be partly related to the internal estimation of both growth and M. Discussions at the PAW recommended reverting back to fixing the adult/asymptotic M (i.e. M-scale) initially at the best estimate from the 2023 diagnostic model, with the Lorenzen M-at-age shape being determined by the internally estimated growth curve (inverse) (Hamer, 2026).

In this assessment we continue to use the Lorenzen functional form for M-at-age (with M-at-age being inversely proportional to the mean length at age) but to fix the asymptotic value of M, i.e. the M for age 40 quarters, and treat this fixed M as an uncertainty in the uncertainty ensemble used for management advice.

5.2.6 Reproductive potential

The reproductive potential ogive is an important component of the assessment structure as it translates model estimates of total population biomass to the relevant management quantity, spawning potential biomass (SB). Assumptions about maturity do not affect the process of fitting the model, only the reference point values. The approach for calculating the reproductive potential at length ogive in this assessment is the same as the 2020 assessment (Vincent et al., 2020). MFCL estimates the reproductive potential at age ogive internally from the externally calculated reproductive potential ogive at length. This length-based ogive is converted internally to the reproductive potential-at-age using a smooth-spline approximation (Davies et al., 2019). This allows for a more natural definition of reproductive potential as the product of three length-based processes: proportion females-at-length[8] (sex-ratio), proportion of females mature-at-length[9], and the fecundity-at-length of mature females[10] (Figure 36). Another added benefit is that this reproductive potential ogive is growth invariant. As for the 2023 assessment we have not included spawning fraction in the reproductive potential calculation as the data are still not adequate for yellowfin in the WCPO.

5.3 Fishery dynamics

5.3.1 Selectivity

Selectivity has typically been estimated as a function of age within MFCL. Recent developments in MFCL, based on recommendations from the yellowfin peer review (Punt et al., 2023), have

[8]For the current assessment, female sex-ratio-at-length was calculated from Regional Observer Program data in SPC's holdings through to 2018, consistent with the previous assessment as few data have been collected since 2018.

[9]Taken from Farley et al. (2017) as in the previous assessment.

[10]Taken from Sun et al. (2006) and standardized per kg of body weight at length as in the previous assessment.

developed an option to estimate length-based selectivity, which was applied in the recent skipjack assessment (Teears et al., 2025). We attempted the length-based selectivity option, however, despite various configurations and model runs we could not achieve better/satisfactory results in terms of data fits than the traditional age-based selectivity. With the limited time for deeper exploration we decided to continue with age-based selectivity for the 2026 assessment.

Estimated selectivity curves at age are included in Figure 32. These selectivity curves can be transformed to equivalent selectivity by length and these are shown in Figure 33. Selectivity is modeled using a cubic spline. This allows for greater flexibility than assuming a functional relationship with age (e.g. a logistic curve to model monotonically increasing selectivity or double-normal curve to model fisheries with reduced selectivity for both the youngest and oldest fish), and requires fewer estimated parameters than modeling selectivity with separate age-specific parameters. This provides a form of smoothing, and the number of parameters for each fishery is the number of cubic spline 'nodes' that are deemed sufficient to characterize selectivity over the age range. The number of nodes may vary among fisheries to allow for reasonably complex selectivity patterns, however we found that 4 nodes was sufficient for all except two fisheries which had 5 nodes.

In all cases, selectivity is assumed to be time-invariant and fishery-specific. However, a single selectivity function cab be "grouped" or "shared" among a group of fisheries that have similar size compositions or were assumed to operate in a similar manner. This grouping of selectivity facilitates a reduction in the number of parameters estimated and allows selectivity to be set for fisheries with either limited or missing size composition data. Selectivity groupings are indicated in Table 1.

The age-based selectivity functions are penalized such that selectivity of age-classes that are similar in size will have similar selectivities for a given fishery or group of fisheries. Additionally, the assumption was made that at least one fishery within each of the model regions would have selectivity that was penalized to be increasing as a function of length in order to prevent the accumulation of an invulnerable, cryptic biomass within the model. In the previous assessment, this asymptotic selectivity setting was applied to the five index fisheries, but in the current assessment it is applied to one longline fishery, 5.LL.OS.2. Other fisheries can be estimated to have asymptotic selectivity, but they are not restricted to have that shape.

5.4 Dynamics of tagged fish

Tagged fish are modeled as discrete cohorts based on the region, year, quarter and age at release for the first 30 quarters after release. Subsequently, the tagged fish are pooled into a common group. This is to limit memory and computational requirements.

5.4.1 Tag reporting

In theory, tag-reporting rates can be estimated internally within the model. In practice, experience has shown that independent information on tag-reporting rates for at least some fisheries is required for reasonable model behavior to be obtained. We provided reporting rate priors (penalties) for reporting groups (combinations of tagging programme and fishery) that reflect independent estimates of the reporting rates and their variances from tag seeding (Peatman et al., 2026a). For the RTTP and PTTP, relatively informative "priors" were formulated for the equatorial purse seine fisheries given that tag seeding experiments were focused on purse seiners. We also make some assumptions regarding fisheries that have similar characteristics to those with independent estimates, but increase the prior variance for these fisheries. For others where there was very little information to inform priors and variance, uninformative priors were allocated, or in cases where there were 5 or less tag returns, tag reporting rates for these groups were not estimated. In these cases, the small numbers of tag returns were removed from the tagging data input file and the reporting rate for these groups set to a fixed value of zero.

In the 2023 assessments reporting rate adjustments were activated during mixing and post-mixing periods. In the 2025 skipjack assessment is was recognised that applying reporting rate adjustments during the mixing periods can cause issues, particularly when lots of tags are returned in the mixing period, if adjustments for under reporting (i.e. reporting rates of around 0.5 for RTTP and PTTP) are made these can remove many more tags, and leave low numbers and even zero tags remaining in the tag population after the mixing period, despite there still being observed recaptures after the mixing period. In cases were reporting rates are thought to be very high shortly after tag release, which is considered highly likely in some tag releases events, applying estimated reporting rate adjustments that greatly reduce tag numbers remaining after the mixing period could significantly bias the fishing mortality higher, and the population scale lower. Conversely not applying reporting rate adjustments will pass biased lower tag recaptures (more live tags remaining) to the post-mixing period, resulting in biasing fishing mortality lower and population scale higher. In the diagnostic model stepwise development we deactivated the reporting rates during the mixing periods, but this was a somewhat subjective judgement decision. We therefore also included the activation of reporting rate during the mixing period in the sensitivity analysis and subsequently in the uncertainty ensemble. This aspect of modeling tag data requires further review and is discussed in the recommendations.

All reporting rates were assumed to be time-invariant, and have an upper bound of 0.99. For this assessment, as previously noted, we did not estimate reporting rates for tag reporting groups with five or few tag return. 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 awareness and follow-up. Consequently, fishery-specific tag reporting rates that are also specific to individual tagging programs were estimated.

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 with the untagged population to have taken 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 during this period based on the actual recaptures, corrected for tag reporting and tagging effects, rather than use fishing mortalities based on the general population parameters. This, in effect, desensitizes the likelihood function to tag recaptures in the specified "pre-mixed" periods while correctly removing fish from the tagged population that is present after the "pre-mixed" period. Herein we refer to the "pre-mixed" period as the "mixing period".

The allocation of appropriate "mixing periods" in stock assessments that use tag-recapture data is problematic as model estimations are sensitive to this assumption and misspecification can bias estimation of management quantities (Kolody and Hoyle, 2014). Mixing rates may vary depending on release locations and times depending on various factors, including; oceanographic dynamics, contexts of releases (e.g., FADs versus free schools, archipelagic waters versus oceanic), fishing effort distribution and environmental/food conditions that influence movements. The yellowfin peer review discussed mixing period assumptions and supported an approach applied to the 2022 skipjack assessment (Castillo Jordan et al., 2022; Punt et al., 2023). In this approach an external individual based model was used to estimate mixing periods 'specifically' for each release group, taking into account the locational and temporal varying conditions of each release event constituting the group that may result in different rates of mixing of the released fish (Scutt Phillips et al., 2022). It applied the individual based Lagrangian model (Ikamoana) (Scutt Phillips et al., 2018) to simulate movement of individual fish (particles) and quantify the fishing pressure (observed) that individuals experienced across their dispersal trajectories in comparison to a population of simulated untagged particles. This approach was explored in full based on a recent SEAPODYM yellowfin model. However, ultimately there were remaining issue with the new yellowfin SEAPODYM model estimates of implausible high fishing mortality levels that meant the mixing estimation from the Ikamoana simulations were not reliable, and were therefore not used in this assessment (see (Scutt Phillips et al., 2026)).

In the previous yellowfin assessment the diagnostic model assumed that tagged yellowfin gradually mix with the untagged population at the region-level and that this mixing process is complete by the end of the second quarter after the quarter in which the fish were releases (ie., “mixing period” assumption is two quarters). A sensitivity was also include whereby the tag mixing period assumption was one quarter (Magnusson et al., 2023). For the 2023 assessment a qualitative assessment of tag mixing based on tag recaptures scaled to the purse seine catches plotted as numbers of tags recaptured per 100 mt of catch in 1 × 1 degree grids cells, was used to indicate these likely tag mixing periods (Teears et al., 2023). This provided some support for the 1 and 2 quarter mixing periods, however, for the 2026 assessment we add a more conservative tag mixing period of 3 quarters, with all three options applied in the uncertainty ensemble.

5.5 Likelihood components

There are four data components that contribute to the log-likelihood function for the yellowfin stock assessment: the index fishery CPUE data, the length frequency data, the tagging data and the conditional-age-at-length data.

5.5.1 Index fishery CPUE likelihood

In catch-conditioned models, CPUE indices differ from the regular extraction fisheries in that no catch is extracted and CPUE is modelled directly as a log-normal likelihood contribution. The contribution of each observation to the log likelihood is made up of time-varying CVs and a generic component, or scaled CV. The time-varying CVs are input from the standardisation models (Teears, 2026) where each region was modeled separately. The generic CVs were derived from overall standard error from each region-specific standardisation model.

5.5.2 Length frequency

The distributions for the length frequency compositional data are assumed to be approximated by robust normal distributions, with the variance determined by the input sample size and the observed length frequency proportion. Size composition distributions for each time period are typically assigned input sample sizes lower than the number of fish measured. Lower input sample sizes recognize that (i) length 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 input sample sizes used by the model are capped at 1,000 within MFCL. The 2023 assessment explored a self-scaling method for weighting the size composition data, the Dirichlet-multinomial likelihood (Thorson et al., 2017). The Dirichlet-multinomial likelihood approach suggested high weights for the size composition data, which resulted in deterioration of the fits to the abundance indices. We again used the simpler approach of applying divisors on the input samples sizes in this assessment, and have multiple options in the uncertainty ensemble.

5.5.3 Tag data

We applied the same approach as the previous assessments which set a fixed tau (τ) across all release groups/recapture fisheries. An alternative to this approach is to allow τ to be estimated (either overall or by groups of recapture fisheries) as done in the 2025 skipjack assessment (Teears et al., 2025), this effectively allows τ to vary somewhat to achieve a better overall objective function within the joint likelihood. The τ estimation will be impacted by all the data in the model, and through the negative binomial assumptions it is influenced by the overdispersion in the tag data that it can detect. τ is not simply a penalty weight on the neg-binomial term. It is implicit in the likelihood formulation. We began experimenting with estimating τ late in the assessment process, but needed more time to understand it. The estimation of τ maybe a reasonable approach, but requires further investigation for this assessment. Even later in the assessment process, exploration of the MFCL flag settings and code for fixing τ for the bigeye assessment found that the last several assessments of yellowfin and bigeye assessments had unbeknownst applied a fixed τ of slightly >1 (1.02), close to Poisson. It appears this was a result of a misunderstanding of the flag settings for fixing τ that had carried across from previous assessments. This was not obvious and required digging into MFCL code to discover. At that point while we had settled on a diagnostic model, we had not run sensitivities or an uncertainty ensemble, we therefore extended the stepwise to include a model with τ=2, given τ of 1 is considered undesirable for tag recapture data as it indicates a Poisson. This also provide a step that could show the difference between a model with high weight on the tag data and the eventual diagnostic model with downweighting of the tag data.

5.5.4 Conditional age-at-length data

A further likelihood component involves using conditional-age-at-length (CAAL) data, as recommended by the yellowfin tuna assessment peer review (Punt et al., 2023). These data are included in the assessment and contribute to estimation of the growth parameters, as they provide direct observations of the distribution of ages, generally measured with some error, within each observed length class. For each fishery and time period (quarter) aged fish are separated into observed (measured) length classes, with the distribution of ages tabulated as a function of the known length (ideally with ageing measurement error, but this feature is still to fully developed and tested in MFCL). The measured age distribution within each length interval is assumed to be multinomially distributed, and this forms the basis of the likelihood component for these data. However, despite the otolith data being collected across a range of locations and times the conditional-age-at-length data generally have a degree of non-independence, as is the case for size compositional data, so these conditional age-at-length data were downweighted by a factor of 0.75 for the diagnostic case model.

As mentioned previously additional CAAL data are available for the 2026 assessment, with enhanced coverage of the region north of 10 degrees N. The samples from the northern region, while only covering a subset of the age-length composition (>60 cm FL), showed larger sizes at age than the samples from equatorial or more southern latitudes. Discussion with Japan FRA staff and at the PAW proposed a weighting method for the CAAL samples by regions to try to weight the growth curve estimation to be more representative of higher population density areas. This was mainly proposed for bigeye but is equally applicable to yellowfin. It was decided to use the global CPUE spatial density predictions to calculate region weighting factors to adjust the observed samples sizes so that the relative sample size proportions of the CAAL data among model regions were the same as the relative mean densities among regions from the global sdmTMB model. The method is described in Hasegawa et al. (2026). The method is conditional on the observed sample sizes, so some are upweighted (i.e. weighting factor >1) and some are downweighted (i.e. weighting factor <1). The region specific weighting factors were applied to the CAAL data from each fishery in a region that had CAAL. The weighting factors operate as scalars and simply multiply by the observed samples. A new feature was created in MFCL for implementing these weights (scaling factors) (Davies et al., 2026b), which has the effect of adjusting (scaling) the effective samples sizes for fitting to the CAAL observations. These weighting factors do not change the overall observed samples size, applying additional overall downweighting is required for that.

5.6 Parameter estimation and uncertainty

The parameters of the model were estimated by minimising the sum of the negative log-likelihoods associated with each of the data components plus the negative log of the probability density functions of the priors and penalties specified in the model. The optimisation is continued either until a pre-specified minimum gradient is achieved or until a maximum number of iterations is reached. This optimisation is performed by an efficient algorithm using exact derivatives with respect to the model parameters (auto-differentiation, Fournier et al. (2012)). Estimation was conducted in a sequence of phases, the first of which uses somewhat arbitrary starting parameter values. A bash shell script, or "doitall", implements the phased sequence for gradually estimating more and more parameters for the model. Some parameters were assigned starting values consistent with available biological information. The values of these parameters are provided in the .ini input file or within the MFCL code.

5.7 Diagnostic methods

5.7.1 Convergence

The following criteria were considered to assess model behaviour and convergence for a diagnostic model:

1. Minimum gradient sufficiently small, preferably less than 0.0001.

2. Hessian matrix was positive definite.
3. Models were not overly sensitive to jitter analyses (i.e. minor sensitivity of management quantities and -ve log-likelihood).
4. The number of parameters estimated on or close to bounds was small, preferably zero.
5. The number and composition of pairs of highly correlated (-0.9 > r > 0.9) estimated parameters.

Obtaining PDH solutions in models with high numbers (i.e. >1,000) of estimated parameters can be challenging, and failure is often due to one or a few very small negative eigenvalues that have little influence on the estimations of key management quantities. While a PDH solution is an important diagnostic to meet for a diagnostic model, this criterion may not be essential (or practicable) for all models in the uncertainty ensemble. While it is possible to simply filter out any model that does not achieve a PDH from the uncertainty ensemble, it is worth also considering the number of -ve eigenvalues, what parameters these relate to, the maximum gradient and the plausibility of the estimated management quantities. The purpose of the uncertainty ensemble is to provide a distribution of plausible management quantities that captures a range of structural, biological and or data uncertainties, rather than to fine tune each individual model to achieve the best possible diagnostics, which in any case would not be possible in the time available. Further, the structural, biological and data uncertainties are typically much more consequential than the parameter estimation uncertainties.

5.8 Other diagnostics

Other diagnostics applied to the diagnostics model, besides the Hessian status discuss above, include:

Jitter analysis Jitter analyses were performed on the diagnostic model as a diagnostic for model stability and to ensure convergence to a global solution. Jitter analysis entails randomly perturbing parameter estimated from the diagnostic model and refitting the model to the data. This 'jittering' can be conducted multiple times, 40 jittered models were produced using the same procedure as in the 2023 assessment (Magnusson et al., 2023).

Model fits to data Plots of observed and predicted CPUE index (and residuals), and LF, tagging recapture and CAAL data were examined.

ASPM ASPM diagnostics (Carvalho et al., 1 08) for estimated by 1) fixing growth and selectivity parameters at their estimated values; 2) removing the LF, tags and CAAL data from the model, leaving only the abundance indices to be fitted; 3) setting all log recruitment deviations to a fixed value of 0; and 4) re-fitting the model estimating only the population scaling parameters. A comparison of biomass, depletion scaling and fishing mortality trends estimated by the ASPM and the full model gives an indication of the extent to which these estimates are informed by the abundance indices. A second version of the ASPM, in which recruitment deviations were estimated, was also run.

Retrospective 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 quarters sequentially excluded (Cadigan and Farrell, 2005). Retrospective analyses with 5 peels for the diagnostic model are presented in (Section 7.3.3). The Mohn's rho statistic, a measure of the average relative bias of retrospective estimates, was computed for recruitment, SB and spawning potential depletion ($SB_t/SB_{t,F=0}$) to indicate whether significant retrospective bias was present in the model.

Likelihood profiles For highly complex population models fitted to large amounts of often conflicting data, absolute estimates of total abundance can be unstable. Therefore, a likelihood profile analysis was undertaken of the marginal penalized likelihood in respect of population scaling, following the procedure outlined by McKechnie et al. (2017) and Tremblay-Boyer et al. (2017). The results of these procedures are presented for the diagnostic model Section 7.4.1). Likelihood profiling is initially conducted on the main data components, i.e. CPUE indices, size data, tag data, CAAL data. If significant conflict is found among the main data components, additional profiling is conducted within each data component to explore the sources of conflict. Profiling may be required at fishery and or regional level within individual data types.

5.9 Stock assessment interpretation methods

5.9.1 Depletion and fishery impact

Many assessments estimate the ratio of recent to initial biomass (i.e. spawning biomass at the beginning of the time period SB_0) as an index of fishery depletion. The problem with this approach is that recruitment may vary considerably over the time series, and if either the initial or recent biomass estimates (or both) are "non-representative" because of recruitment variability or uncertainty, then the ratio may not measure real fishery depletion, but simply reflect recruitment variability. This problem is better approached by computing the spawning potential time series (at the model region level) using the estimated model parameters, but assuming that fishing mortality was zero. Because both the estimated spawning potential SB_t (with fishing), and the unfished spawning potential $SB_{t,F=0}$, incorporate the same recruitment variability, their ratio at each quarterly time step (t) of the analysis, $SB_t/SB_{F=0[t]}$, can be interpreted as an index of fishery depletion. The computation of unfished biomass may include an adjustment in recruitment magnitude to account for the possibility of reduction in recruitment for exploited populations through stock-recruitment effects. To achieve this the estimated recruitment deviations are multiplied by a scalar based on the difference in the SRR between the estimated fished and unfished spawning potential estimates.

A similar approach can be used to estimate depletion associated with specific fisheries (i.e. fisheries impact analysis) or groups of fisheries. Here, fishery groups of interest - longline, purse seine associated sets, purse seine unassociated sets, pole and line and "other" fisheries, are removed in-turn in separate simulations. The changes in depletion observed in these runs are then indicative of the depletion (impacts) caused by the removed fisheries.

5.9.2 Reference points

The unfished spawning potential ($SB_{F=0}$) in each time period was calculated given the estimated recruitments and the Beverton-Holt SRR. This offers a basis for comparing the exploited population relative to the population subject to natural mortality only i.e. $SB/SB_{F=0}$. The WCPFC adopted limit reference point (LRP) for the bigeye stock is $SB_{\text{recent}}/SB_{F=0}$ =0.20. SB_{recent} refers to the average spawning biomass across year 2021–2024, whereas $SB_{F=0}$ refers to the average unfished biomass across the years 2014-2023. There is no agreed WCPFC target reference point for the bigeye tuna stock however CMM 2025-02 states in para 10 "Pending agreement of a target reference point the spawning biomass depletion ratio ($SB/SB_{F=0}$) is to be maintained at or above the average $SB/SB_{F=0}$ for 2012–2015". Stock status was referenced against these points by calculating the reference points; $SB_{\text{recent}}/SB_{F=0}$ and $SB_{latest}/SB_{F=0}$ where $SB_{F=0}$ is calculated over 2014–2023 and SB_{recent} and SB_{latest} are the mean of the estimated spawning potential over 2021–2024, and 2024 respectively (Table 2).

The other key reference point, $F_{\text{recent}}/F_{\text{MSY}}$, is the estimated average fishing mortality at the full assessment area scale over a recent period of time (F_{recent}; 2020–2023 for this stock assessment) divided by the fishing mortality producing MSY which is a product of the yield analysis and is detailed in Section 5.9.3.

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 Davies et al. (2026c).

5.9.3 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 be determined using the formulae given in Davies et al. (2026c), 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.

Fishing mortality-at-age (F_a) for the yield analysis was determined as the mean over a recent period of time (2017–2020). We do not include 2021 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 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 (1952–2024). 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.

5.9.4 Kobe analysis and Majuro plots

For the standard yield analysis (Section 5.9.3), the fishing mortality-at-age, F_a, is determined as the average over some recent period of time (2020–2023). In addition to this approach the MSY-based reference points (F_t/F_{MSY}), and SB_t/SB_{MSY}) are 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.

5.9.5 Stock projections from the structural uncertainty grid

Projections of stock assessment models can be conducted within MFCL to ensure consistency between the fitted models and the simulated future dynamics, and the framework for performing this exercise is detailed in Pilling et al. (2016). Typically, stochastic 30 year projections of recent catch are conducted from each assessment model within the uncertainty grid developed. For each model, 100 stochastic projections, which incorporate future recruitments randomly sampled from historical deviates, are performed.

The results of stock projections are not available in this version, and will added in later revision

6 Model runs

6.1 Developments from the last assessment

The progression of the model development (referred to as the "stepwise") from the 2023 diagnostic model to the 2026 diagnostic model is described here.

6.1.1 Stepwise model development

The major changes incorporated at each step in the diagnostic model development are summarised below, including the model number codes for each step, and a brief description. Each step builds from the previous step, and in most cases retains all the previous changes.

- **[01]** 2023 diagnostic-model refit.
- **[02]** New MFCL executable and INI1007.
- **[03]** Lorenzen natural-mortality scaling fixed to the 2023 diagnostic-model estimate, and bias-corrected length-weight parameters.
- **[04]** Convert to the 36-fishery structure (31 extraction, 5 index).
- **[05]** Size-composition inputs replaced with the reweighted weight data converted to lengths for longline fisheries, data until 2021.
- **[06]** Weight-as-length plus more recent observed-length compositions for longline fisheries, data until 2021. Note that in no instance is originally measured weight and lengths used at the same time for a fishery (Peatman et al., 2026b)
- **[07]** Update all data sets to 2024, except CAAL.
- **[08]** Switch to separate sdmTMB regional CPUE indices and global sdmTMB for regional scaling applying the MFCL regional scaling penalty, global model time period 1952-2024 (all time period), strong recommendation from PAW and yellowfin assessment peer review (Punt et al., 2023; Hamer, 2026).
- **[09]** Update CAAL data, reweighted based the global CPUE sdmTMB relative mean density.
- **[10]** Time-varying CPUE uncertainty.
- **[11]** Growth curve: estimate L1 rather than fixed.
- **[12]** Add effort creep adjustment to the regional CPUE abundance indices,(1% per year prior to 1976 when no vessel ID available, 0.5% post 1976 when vessel ID available, recommended from PAW (Hamer, 2026)
- **[13]** Allow declining right limbs on all except one fishery (F5) selectivity at age
- **[14]** Implement the orthogonal polynomial recruitment.
- **[15]** Deactivate the reporting rate adjustments during mixing.
- **[16]** Tag data downweight tau (τ) set to 2. **Diagnostic2026**

Many more additional models were explored in the development of the diagnostic model, these are not described in this report.

6.2 Sensitivity analyses and uncertainty ensemble

6.2.1 Sensitivities

One-off sensitivity models were explored to understand the sensitivity of the diagnostic model estimations to structural, biological and data uncertainties, and to inform selection of models for the uncertainty ensemble. Each one-off sensitivity model was created by making a single change to the 2026 diagnostic model. Due to the time constraints there was limited time to explore diagnostics on sensitivities.

We primarily focused on the key sensitivities below:

- Steepness: 0.65, 0.95 (diagnostic: 0.8)
- Tag mixing: 1 quarter, 3 quarter (diagnostic: 2 quarter)
- Sample size divisors 10, (diagnositic 20), 40,
- Adult natural morality (M at age 40 quarters): Low =0.077/quarter , Higher M = 0.149/quarter (diagnostic from 2023 = 0.119/quarter)
- Effort creep on longline CPUE indices: 'maximum' scenario suggested from PAW (Hamer, 2026) 2.5% per year 1962–1976, 1.25% year post 1976, (diagnostic: min 1% per year prior 1976, 0.5% year post 1976).
- Reporting rates during mixing period: reporting rates activated (diagnostic: reporting rates not activated)
- Tau 1 v tau 2

6.2.2 Uncertainty characterisation

Typically, three types of uncertainty could be incorporated into the estimates of stock status used for management advice. One involves the statistical uncertainty of the estimates produced by individual models, often referred to as 'estimation' uncertainty. The second involves 'model' uncertainty, which is the uncertainty in the structural and fixed-parameter assumptions underpinning individual models, (e.g., fixed biological parameter such as M, steepness etc.). The third involves data inputs, such as alternative abundance indices, size data treatments or other data inputs, and we include data weighting in the likelihood which alters the influence of data types or data components on the estimation of management quantities.

The 'uncertainty ensemble' provides a distribution of plausible management quantities that captures a range of structural, biological and or data uncertainties. For this assessment we have followed the Monte Carlo ensemble approach applied in the most recent skipjack (Teears et al., 2025) and south Pacific albacore assessments (Teears et al., 2024). The choice of uncertainties to included in the ensemble models is informed by the sensitivity analyses. For the uncertainty ensemble we include a combination of parameter distributions from which values are randomly drawn which implicitly weights those draws according to the shape of the distribution, and fixed values of alternative plausible options for other factors which are drawn at random with equal probability. 120 separate combinations of these parameters/factors were drawn and each combination formed a separate model run for the uncertainty ensemble. For the ensemble we only included models had a PDH or a gradient $\leq 1 \times 10^{-3}$ and fewer than 6 negative eigenvalues. This filter retained 48 models.

The parameters and fixed factors included in the uncertainty ensemble:

- Steepness [beta distribution] Figure 9
- Tag mixing period [1 quarter, 2 quarter, 3 quarter]
- Size data weighting divisor [10, 20, 40]
- M scale (M at age 40 quarters) [beta distribution] Figure 9
- Effort creep [5 x options -lower to higher, based on PAW discussions]
- Reporting rates during mixing period: 0-reporting rates activated, 1-reporting rates not activated.

Estimation uncertainty The inclusion of estimation uncertainty is recommended to form the basis for overall uncertainty in the management quantities. Hessian-based estimates of the standard deviations for the key management quantities are obtained using the variance-covariance matrix of the model parameters and the Delta Method. Estimation uncertainty is provided for the terminal management quantities: $SB_{\text{recent}}/SB_{F=0}$, $F_{\text{recent}}/F_{\text{MSY}}$, SB_{recent}/SB_{MSY} for the ensemble models included with positive definite Hessian. The estimation uncertainty is combined across all models in the ensemble using a parametric bootstrap similar to the approach used in the 2023. We generate 1000 random draws from lognormal distributions with mean and standard deviation specified for each ensemble model and transform the random deviates to normal space by taking their exponent. The management quantities with the full uncertainty are summarised with the mean, median, and 10 and 90 percentiles, minimum and maximum of the 48,000 values of each management quantity.

7 Results

7.1 Consequences of key model developments

The progression of model development from the 2023 diagnostic model to the 2026 diagnostic model is described below and results for depletion and spawning potential are displayed in Figure 11 and Figure 10. The stepwise analysis changes are discussed briefly below, each step compared to the previous step. The stepwise development should only be considered indicative of changes, because the models in each step do not undergo thorough diagnostics and tuning. The diagnostic model represents the culmination of a number of steps that incorporate changes from the previous assessment, many of which have been suggested and/or recommended through peer discussion at the Pre-assessment Workshop (PAW) or in response to comments from the SC on the previous assessment. Others relate to new data or biological information and changes the analysts decide were useful and could improve the model from their experimentation. The diagnostic model may or may not be the best possible model for the data, in fact it rarely if ever will be with the limited time to follow the many avenues of possible improvement. The diagnostic model is also not a model for management advice. It should be viewed as an experimental model for exploring model behaviour and performance, and it should pass the common performance diagnostics to provide a basis model for sensitivity analyses to inform the structure and components of the model grid or ensemble that is used to construct management advice.

The results of the stepwise model development can be explored in detail at this app.: **https://ofp-sam.shinyapps.io/yft-2026-stepwise/**

- **[01]** 2023 diagnostic-model refit: no change in $SB/SB_{F=0}$ or SB.
- **[02]** New MFCL executable and INI1007: no change in $SB/SB_{F=0}$ or SB.
- **[03]** Lorenzen natural-mortality scaling fixed to the 2023 diagnostic-model estimate, and bias-corrected length-weight parameters: no change in $SB/SB_{F=0}$ or SB.
- **[04]** Convert to the 36-fishery structure (31 extraction, 5 index): minor change in historical level during 2000s, slight increase in terminal $SB/SB_{F=0}$ (less depleted) and SB.
- **[05]** Size-composition inputs replaced with the reweighted weight data converted to lengths for longline fisheries, data until 2021: slight reduction of $SB/SB_{F=0}$ (more depeleted) and SB.
- **[06]** Weight-as-length plus more recent observed-length compositions for longline fisheries, data until 2021: slight reduction (more depleted) in $SB/SB_{F=0}$ and SB.
- **[07]** Update all data sets to 2024, except CAAL: some changes (lower levels) in late 1990s-2000s, but very minor change in terminal year $SB/SB_{F=0}$ and SB
- **[08]** Switch to separate sdmTMB regional CPUE indices and apply the MFCL regional scaling penalty feature based on a global sdmTMB model, time period 1952-2024 (all time period): **first major step change**, shift to notably lower $SB/SB_{F=0}$ (more depleted) and SB.
- **[09]** Update CAAL data, reweighted based the global CPUE sdmTMB relative mean density: very minor change in $SB/SB_{F=0}$ (less depleted) and SB.
- **[10]** Time-varying CPUE uncertainty: very minor change in $SB/SB_{F=0}$ (less depleted) or SB.
- **[11]** Growth curve: estimate L1 rather than fixed: minor increases in $SB/SB_{F=0}$ and SB.
- **[12]** Add effort creep adjustment to the regional CPUE abundance indices: very minor change in terminal $SB/SB_{F=0}$ and SB, but higher early SB with steeper decline.

- **[13]** Allow declining right limbs on all except one fishery (F5) selectivity at age: very minor change in $SB/SB_{F=0}$ and SB.

- **[14]** Implement the orthogonal polynomial recruitment : minor change, slightly increased terminal $SB/SB_{F=0}$ (less depleted) and very minor change in SB.

- **[15]** Deactivate the reporting rate adjustments during mixing: **second major step change**, notable scaling impact, large increase in historical and terminal $SB/SB_{F=0}$ (less depleted) and increased SB.

- **[16]** Tag data downweighted tau τ set to 2: (**third major change**) substantial reduction (more depleted) in $SB/SB_{F=0}$ and lower SB. **Model 16 = Diagnostic2026**

Overall, the biggest sequential step changes influencing depletion and spawning biomass scale were related to:

- the implementation of to the new CPUE methods (region specific sdmTMB for abundance and global sdmTMB for scaling) with the new MFCL regional scaling penalty, this had scaling being influenced by the full time series of the global sdmTMB rather than the shorter (1965-1969) maximum spatial coverage option which did not produce sensible results, noting the penalty term only applies for the time period of the input global CPUE data.

- the binary option to activate or deactivate the reporting rate adjustments in the tag mixing period, deactivating had the expected effect. Applying estimated reporting rate adjustments that reduce tag numbers remaining after the mixing period could significantly bias the fishing mortality higher and the population scale lower if these adjustment are biased high. Conversely not applying reporting rate adjustments, as in this stepwise change, will pass lower tag recaptures (more live tags remaining) to the post-mixing period, resulting in biasing fishing mortality lower and population scale higher. Both of these options (activate or deactivate) are not ideal, and hence they are included as a structural uncertainty for the uncertainty ensemble, somewhere in the middle is likely the reality. A different approach is required to allow different reporting rate adjustments within and after the mixing period, however, this would still remain an important uncertainty, but could be better dealt with compared to the current binary option.

- downweighting tag data by increase tau from 1 to 2, had the effect of estimating a more depleted stock and lower spawning biomass, as the influence of the tag data on scaling was downweighted. More work is required to explore tag data downweighting levels as this is not a binary option, unlike the mixing period reporting rate adjustments.

7.2 Fit of the diagnostic model to data sources

7.2.1 Standardised CPUE from index fisheries

The model fits to the index fishery standardised CPUE data were generally quite good for the main trends and dynamics for all model regions (Figure 12). An exceptions was the later period in model region 4 were the fit deteriorated in later 15 years, and the model tended to predict higher CPUE than the majority of the observations. The model over predicted the early period observation in region 3 (late 1960s-early 1970s) and region 4 (1950s) (Figure 13)

7.2.2 Size composition data

Longline fisheries: The aggregate model fits to the length compositions for the longline extraction fisheries Figure 14) were relatively good for all fisheries. The longline index fisheries in the 2026 assessment no longer share selectivity, and look good on aggregated fit. The fit for the region 3 and 5 indices tends to over predict the smaller size frequencies, whereas region 2 over predicts the larger size frequencies.

The time series plots of observed versus predicted mean lengths for the longline fisheries indicated the model struggled to follow the dynamics of the observations very closely, with tendency to fit through the middle. As such it did not do well in predicting the recent declines in the observed data, for example in F4, F9, F2, and LL index fisheries; F1,F2,and F4 (Figure 15)

Other fisheries: The aggregated length composition fits for the industrial purse seine fisheries, are variable, noting some fisheries have very low samples, and many are multimodal, they mostly overestimate the frequencies of the smaller fish, F25 and F29 (associated regions 3 and 4) also overestimated larger fish (Figure 16). For the domestic purse seine fisheries the aggregate fits were good, noting it will be difficult to ever get a good fit to the F12 (Japan) due to the trimodal size pattern (Figure 17).

Time series fits for purse seine fisheries are generally poor and tend to over predict the median length of the main industrial purse seine fisheries; F25, 26, 28, 29, 30 (Figure 18)

The fits to other gears domestic, handline and pole and line, were variable, the pole and line fits were the worst but they have very low data and low catches, so they are of limited importance. Fits to the important handline fisheries were good (Figure 19.

Time series fits for misc gears in Philippines and the handline fisheries in Indonesia and Philippines; F 22, 14, 15), are reasonable in that they don't have clear bias. For the other fisheries the data are poor and the model estimates are highly variable or fit through the middle of the observation (Figure 20).

7.2.3 Tagging data

When aggregated, the tag attrition estimates fit the tagging data relatively well (Figure 21), albeit slightly underestimating tag returns for quarters 4, 5 and 6 (Figure 22).

For the tag returns by year/quarter of recapture and groups, the model-predicted tag returns show relatively good agreement with the observed data, albeit with some spikes missed and some over- and under-prediction at various periods, but for the groups with very low data (Figure 22).

7.2.4 Tag reporting rates

The penalties on the reporting rates are relatively weak even for groups informed by tag seeding. Estimated tag reporting rates by fishery recapture groups are shown for RTTP and PTTP in Figure 23 and JTPT in Figure 24. As expected, the reporting rate estimates differ among fisheries groups and across tagging programs. For the RTTP and PTTP the reporting rate estimates for those groupings that received higher penalties (based on external estimated reporting rates from tag seeding i.e. 5, 9, 10, 11, 14, 17, 18, 19) did not estimate near the penalty mean, 5 were on the upper bound and the others wanted higher reporting rates. Of the RTTP and PTTP groups with relatively uniformed penalties, 4 were at the upper bounds, the other were at around 0.3. All JPTP were on the upper bounds. The large number of groups estimated at the upper bound is an issue that plagued previous assessments and has not been resolved in this assessment.

7.2.5 Conditional age-at-length

The estimated growth curve for the diagnostic model, which is based on both the CAAL data and the size data is shown in Figure 37, L1=11.9, L2=134.8, and K=0.168.

Aggregate fits to the conditional-age-at-length (CAAL) data are shown in Figure 25. For length bins up to around 130 cm, which constitute most of the observed CAAL data, the model's predicted median ages at lengths fit the observed data very well. At very small sizes ¡50 cm the model tends to slightly over predict the observed median age, and for the largest sizes, the model prediction is below the observed median. Overall the fit across the size range that constitutes most of the catch is reasonable.

At the model region level, region 1 has a reasonable amount of data, with the addition of the new samples from north of 10° N) and shows the best fit, with the model estimates tracking the observed median ages at length well. Region 2 has low data and the fit is poor for larger sizes. Region 3, has all small fish below about 130 cm, and the fit is good until around 80 cm above which the model slightly underpredicts the median age. Region 3 and 5 have more data, and similar to the aggregated data, predict the observed data reasonably well up until around 130 cm, after with the median age at length tends to be under predicted.

7.3 Diagnostic analysis for the diagnostic model

7.3.1 Convergence

The diagnostic model has a maximum gradient component of 9.96×10^{-5}. There were 49 parameters on bound: 29 movement coefficients, 5 selectivity spline nodes, and 15 tag reporting rates. There were 3 parameter correlations whose absolute value is higher than 0.9. These are between the three von Bertalanffy growth parameters.

7.3.2 Jitter

From the 40 jitter models, 33 jitter models ran to completion and achieved an objective function value that was within 200 units of log-likelihood from the diagnostic model. The remaining 7 models had a worse objective function value or stopped running. No jitter model achieved a better fit than the diagnostic model Figure 26. All 33 jitter models produced very similar SB and $SB/SB_{F=0}$, consistent with the diagnostic model Figure 27.

7.3.3 Retrospectives

The retrospective analysis of the yellowfin tuna diagnostic model produced Mohn's rho values of -0.196 for spawning potential, SB, and -0.057 for depletion, $SB/SB_{F=0}$, Figure 28. Within or close to the acceptable range of ± 0.2 to ± 0.3. The most notable feature was in the first peel when spawning potential scale and the $SB/SB_{F=0}$ dropped. This may relate to the large increase in the terminal year catch of the handline fishery, which is the result of a change in how Indonesia allocates yellow fin catch between small and large handline fisheries. In 2024 they started to allocate a much greater proportion of their yellowfin handline catch to the large fish handline fishery. This adjustment to the allocation method requires historical adjustments based on advice from Indonesia.

7.3.4 ASPM

7.4 Age-structured production model diagnostic

The age-structured production model (ASPM) diagnostic, run under the constant recruitment assumption, indicates a more pessimistic stock status than the diagnostic case (Figure 29). The ASPM with constant recruitment estimates a substantially smaller spawning biomass, especially historically, and lower depletion ($SB/SB_{F=0}$), with a sharp decline over the recent decade. In contrast The ASPM with recruitment variation is very similar to diagnostic model. Both the constant and variable recruitment ASPM provide similar fishing mortality levels and dynamics as the diagnostic model.

Overall, the ASPM diagnostics suggests that the fishing mortality trend from the diagnostic case is broadly reproducible from CPUE data alone, without length composition, CAAL or tag data, supporting general consistency between the abundance indices and the exploitation history. However, the magnitude of the estimated decline in spawning biomass, and the recent stock status is sensitive to the treatment of recruitment.

7.4.1 Likelihood profiles

The likelihood profile over the mean total biomass for the diagnostic model was an improvement from previous assessments in that conflict was reduced among the sized data, CPUE and tag data (Figure 30). The CAAL data had the strongest conflict with the other data.

Profiles on the length and CAAL data by fishery groupings and regions are shown in (Figure 31). These show that there was no strong conflict between the longline and purse seine fisheries, but the domestic mixed gear and handline fisheries in Indonesia and Philippines had some conflict, wanting lower biomass scale than the longline and purse seine. There was some conflict among model regions for the length data, region 1 and 3 preferring lower biomass scale and region 5 higher, region 2 and 4 in the middle. There was no strong conflict among the fishery groups or regions for CAAL data. Overall, data conflict looked to have improved compared to the 2023 assessment diagnostic model.

7.5 Population dynamics estimates

Plots of the diagnostic model time series of population dynamic quantities, $SB/SB_{F=0}$ and SB, are provided in Figure 41 and Figure 42.

7.5.1 Selectivity

A range of selectivity curves are estimated for the different fisheries in the model and can be largely classified by gear type. The age-specific selectivity patterns are shown in Figure 32 and length-specific selectivity patterns are shown in Figure 33.

7.5.2 Movement

The model estimated fairly low movement rates among model regions (Figure 34). There was very little movement from region 1 to any other regions and very minor movement from region 2 to other regions, despite the very high recruitment estimated for that region(Figure 39). Region 3 had movement mostly to regions 4 and 5, and region 4 had movement to all other regions, region 5 had movement mostly to regions 3 and 4.

7.5.3 Natural mortality

The Lorenzen form of natural mortality M is used in the 2026 assessment, with the scale of this curve fixed at 0.119 per quarter at age 40 quarters Figure 35. The shape is influenced by the estimated growth curve.

7.5.4 Maturity

Maturity-at-age is derived from the fixed maturity-at-length (fixed at the same values used in the 2020 and 2023 diagnostic models) and applying the estimated growth curve to this to get maturity-at-age for the 2026 diagnostic model. This maturity-at-age and length curves for the diagnostic model are shown in Figure 36.

7.5.5 Growth

The diagnostic model estimated growth curve is shown in Figure 37, and has the following estimated parameters: L1 11.9 cm, L2 134.5 cm , K 0.168.

7.5.6 Recruitment: diagnostic model

Total annual recruitment was higher early in the time series and declined to around 1990, after which it stabilised until a recent increase in the last 5-10 years Figure 38. There is high variation in Region 1, 3, and 5 on multiyear time scales. The key feature of the recruitment estimates is the domination by region 2, where large catches of small yellowfin are taken in the Indonesians and Philippines fisheries Figure 39.

The fit to the stock recruitment relationship for the 2026 diagnostic model is shown in Figure 40.

7.5.7 Spawning potential and depletion: diagnostic model

The 2026 diagnostic model (all regions) estimates an initial steep decline in the overall spawning potential from the 1950s through until the early 1990s, after which it stabilised (Figure 42). The pattern in the decline in spawning potential over time, is consistent with the CPUE indices (Figure 12) and mirrors the catch history (Figure 2), but there are differences among regions since the mid 1990s. Region 1, 3 and 5 show recent increasing trends, region 4 is stable but region 2 shows a recent decline.

The diagnostic model shows a decline in overall spawning potential depletion, $SB/SB_{F=0}$ through to around the 2010, which is consistent across regions (Figure 41). During the last 15-20 years the tends vary among regions. Region 2 declines steeply, regions 1 and 3 increase and region 4 is stable. The terminal depletion for diagnostic model ($SB/SB_{F=0}$) across all model regions is 0.352, and the $SB_recent/SB_{F=0}$ is 0.392.

7.5.8 Fished (SB) versus unfished ($SB_{F=0}$) spawning potential: diagnostic model

To interpret the trends in spawning depletion it is useful to compare the trends in spawning potential, SB with the predicted spawning potential that would have occurred in the absence of fishing, $SB_{F=0}$, also called the unfished biomass (Figure 43). Unfished biomass is the denominator in the depletion ratio.

The total unfished biomass follows the same decline as the spawning potential during the first two decades of the stock assessment period, indicating that the steep decline in the estimated spawning potential before 1970 can be primarily explained by the estimated recruitment trends rather than fishing. From 1970 onwards, the unfished biomass is estimated to have been relatively stable for the next decades, until a recent increase in the last few years near the end of the assessment period. This matches the estimated recent recruitment increase seen since around 2010 (Figure 38).

7.5.9 Fishing mortality: diagnostic model

The temporal trend in the adult fishing mortality has been a gradual increase until around 2010 after which it has stabilised (Figure 44). Juveniles, as defined by the maturity ogive (Figure 36), have fishing mortality rates that are higher than that of adults since 1970's, with strong fluctuations in regions 2 and 3, were the highest catches of small juveniles occur.

The juvenile fishing mortality is estimated to have increased rapidly in the last few years in region 3, however there is a drop in region 2 in the most recent year. Again this likely due to change in how Indonesia has allocated catch between small and large fish handline, with a substantial increase in allocation to the large fish handline in 2024. This allocation change needs historical correction based on advice from Indonesia (Figure 2).

7.6 One-off sensitivity analyses

Comparisons of the spawning depletion and spawning potential trajectories for the diagnostic model and the related one-off sensitivity models are provided in Figure 45, Figure 46, Figure 47.

These comparisons show that estimates of both spawning depletion and spawning potential were somewhat sensitive to the choice of steepness, tag mixing and the switch from tau 1 to 2 (Figure 45). Effort creep showed minor sensitivity between the diagnostic lower level and the maximum rate, with the sensitivity higher for the spawning potential at the start of the model period, higher effort creep resulting in much higher initial spawning potential and steeper decline. The model was highly sensitive to the natural mortality scale assumption (M at age 40) (Figure 46). The diagnostic model showed minor sensitivity to the choice of activating or deactivating reporting rate adjustments in the mixing period. This was somewhat surprising given the influence of this change in the stepwise and and low sensitivity may related to interaction with the tau of 2 downweighting the influence of tag data. The size data weighting divisors had minor sensitivity (Figure 47).

7.7 Stock assessment results

7.7.1 Uncertainty ensemble

Across the 48 models with suitable convergence in the uncertainty ensemble, the most important factors when evaluating stock status uncertainty for depletion ($SB_{\text{recent}}/SB_{F=0}$) were the assumed values for the natural mortality scale (most influential), the choice of whether to activate or deactivate the estimated reporting rate adjustments within the mixing periods (second most influential) and the tag mixing period assumption (third most influential). The data weighting divisors for the length data also had some influence on $F_{\text{recent}}/F_{\text{MSY}}$ (Figure 52).

The trajectories of $SB_{\text{recent}}/SB_{F=0}$, spawning potential ($SB$), fishing mortality and recruitment summaries across the ensemble are displayed in Figure 48, Figure 49, Figure 51 and Figure 50.

The depletion shows a declining trend overall and in all regions until around 2010. The overall trend then stablises after 2010, but there are different recent trends across the model regions, for example, region 2 declines sharply from around 2015, regions 1, 3 and 5 show increasing trend and region 4 is stable (Figure 48).

The spawning potential has a large spread early in the time series which is influenced by the inclusion of the alternative effort creep scenarios, that strong influence on the spread on the initial biomass. The uncertainty contracts and the trend stabilises from the early 1990s, overall, and in each model region (Figure 49).

The fishing mortality has very large uncertainty for region 1, and highly variable trends and variability across regions. The overall trend is increasing from the 1970s until 2010 and then highly variable, with peaks in the 2000s and 2020s (Figure 51).

The recruitment also has very large uncertainty for region 1, and is highly variable for regions 3, 4 and 5. The overall trend is, however, driven by region 2 where the highest recruitment is estimated (Figure 50).

Summary plots of the distributions of the key management quantities with respect to different components of the uncertainty ensemble are provided for $SB_{\text{recent}}/SB_{F=0}$, $F_{\text{recent}}/F_{\text{MSY}}$ and SB_{recent}/SB_{MSY} (Figure 53, Figure 54, Figure 55).

7.7.2 Analyses of stock status

There are several ancillary analyses related to stock status that are typically undertaken on the diagnostic model including dynamic Majuro and Kobe analyses and fisheries impacts analyses.

We do not present the results of all analysis for all models in this report. In this section, we rely largely on the tabular results (definitions Table 2) of the structural uncertainty grid (Table 3, Table 4, Table 5) and the dynamic spawning depletion and spawning potential plots for the models in the structural uncertainty grid (Figure 48, Figure 49, Figure 51), and the dynamic Majuro and Kobe plots (Figure 57).

Majuro and Kobe plots and comparisons with Limit and Target Reference Points: Majuro and Kobe plots show the spread of the ensemble in the green area, that would be considered not overfished and not undergoing over fishing. These plots are model point estimates, addition of estimation uncertainty would expand the spread, this can be added in rev.(Figure 56).

The dynamic Majuro and Kobe plots for the diagnostic model show the steady increase in depletion (declining stock status) and fishing mortality of the stock since the 1950s with an increase in fishing mortality from the late 1960s. There more recent trend indicates that the stock status has been fluctuating around a similar condition over the last decade or more (Figure 57).

Fishing impact and time dynamic MSY: In addition to the above analysis, it is possible to attribute the fishery impact with respect to depletion levels to specific fishery components (i.e., grouped by gear-type), in order to estimate which types of fishing activity have the most impact on the spawning potential (Figure 58). Fishing impacts were estimated to be very minor in all regions before about 1970, resulting primarily due to the longline fishery. The impact of longline and pole and line gears has increased very slightly over the time series. In the early 1970s, catch information from the miscellaneous fisheries (mostly from Region 2) leads to an increase in impact, particularly in that region. Subsequently, the onset of notable impacts due to purse seine fishing occurs from the 1980s in all tropical regions. Examining the overall impact, the purse seine associated and unassociated fisheries have similar impact, and the largest impact in recent years is due to the miscellaneous fisheries, which in this analysis include the large fish handline fisheries in ID and PH, future versions of this analysis should separated out the large handline fisheries from the miscellaneous gears category.

Yield analysis: The yield analyses (Figure 59) conducted in this assessment incorporates the spawner recruitment relationship (Figure 40) 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 a range of steepness values. The yield distributions under different values of fishing effort relative to the current effort are shown in Figure 59 for the diagnostic model. For the diagnostic model, it is estimated that MSY would be achieved by increasing fishing mortality by a factor of 1.4, although the resulting increase in yield would be relatively small.

The yield analysis also enables an assessment of the MSY level that would be theoretically achievable under the different patterns of age-specific fishing mortality observed through the history of the fishery (Figure 60). Prior to 1970, the WCPO yellowfin fishery was almost exclusively conducted using longline gear, with a low exploitation of small yellowfin. The associated age-specific selectivity pattern resulted in a much higher MSY in the early period compared to the recent estimates. This pronounced decline occurred after the expansion of the small-fish (other) fisheries in region 2 and, soon after, the rapid expansion of the purse seine fishery which shifted the age composition of the catch towards younger fish. This lower MSY is due to a combination of fish being removed from the system at smaller sizes and also before they have the chance to reproduce. The final year uptick in the MSY is due to ID change its approach to allocation their yellowfin tuna catch between large fish handline and small fish handline, whereby their new approach shifts substantial catch from the small fish handline to the large fish handline (implemented in the last year of this assessment), resulting in a change in the size and age selectivity to larger/older fish. Historical correction of this catch allocation to large fish handline would be expected to alter (increase) the historical MSY.

Catch based projections: These will be added in a rev. and likely not available for SC22.

8 Discussion and conclusions

8.1 General discussion

The overarching objectives for the 2026 assessment were to develop a model with good convergence properties, to simplify the model as much as possible, and to incorporate innovations recommended by experts, including the Scientific Committee.

8.1.1 Model convergence

This was known to be a considerable challenge, based on the experience from the yellowfin 2023 assessment. The diagnostic model in that assessment had a positive definite Hessian (PDH), but only 17% of the grid models. There was no guarantee that the 2026 assessment model could achieve a PDH.

Early in the stepwise development of the current assessment, after adding three years of new data, the model no longer achieved a PDH. This continued to be the case for every subsequent stepwise change until at the very end, after a determined search across a variety of model configuration options, a PDH model was found. From the full set of 120 ensemble models, 35% had a PDH, which increased to 88% after filtering the ensemble based on gradients and negative eigenvalues.

8.1.2 Simplifying the model

At the start of the assessment work, the fisheries definitions were revisited. The 2023 assessment adopted a simplified regional structure near the end of the model development, but the fisheries definitions remained unchanged. With fewer regions, it was clear that some fisheries could be combined, especially fisheries that use the same gear in the same region.

An attempt was made by the stock assessor of the yellowfin assessment this year to combine many fisheries. For example, to absorb some of the smaller fisheries that have negligible annual catches, and also to combine fisheries that use different gears but have similar length composition. A series of proposals were evaluated by domain experts with detailed knowledge of the fisheries and the sampling process, and the final conclusion was that it would not be advisable to go very far in combining different fisheries. Some fisheries were in fact split into separate fisheries in this year's assessment work, to distinguish between the fisheries of Indonesia, Philippines, and Vietnam that used to be combined. In the end, the total number of extraction fisheries changed from 32 to 31 between the assessments.

Near the end of the stepwise model development, as the search for a PDH model was becoming more urgent, the feature of orthogonal polynomial recruitment (OPR) was found to be the only way to produce a PDH model for the yellowfin assessment. Some of the negative eigenvalues of stepwise models were related to specific recruitment in a given region, year, and quarter, where there is little data to guide the model estimation of such specific recruitment events. These negative eigenvalues are a model artifact coming from the estimation of too many parameters where there is too little information. OPR is a very substantial simplification of the model in terms of estimated parameters, as this change in the stepwise development caused the number of parameters to decrease from 1963 to only 452 parameters. OPR is a polynomial smoother that requires the modeller to explore and specify different degrees of smoothing over years, seasons, and regions, while keeping an eye on and penalizing unusual recruitment predictions near the end of the time series. In other words, OPR can be seen as a simplification in terms of parameter estimation, but a slightly increased level of complexity for the stock assessment work.

8.1.3 Innovations

Converting all weight composition data to length was a structural change that is likely to yield several benefits in the long term. Among other things, it facilitates the analysis of data conflict, comparing fisheries using the same currency, and tracking progress of reducing data conflict. Using only length composition data can also lead to simple and streamlined design and implementation of the next-generation tuna assessment software.

Regional scaling was a highly deterministic factor in previous assessments, where the CPUE indices effectively dictated how the population biomass was divided between regions. In the current assessment, this approach is replaced with a multivariate normal prior that gives the model more freedom regarding the distribution of biomass between the regions.

Effort creep has been a discussion point for the Scientific Committee in the past, and there is a wide consensus that the level of effort creep should be modelled as greater than zero. At this year's Pre-Assessment Workshop, a decision was made on how technological effort creep should be implemented in the yellowfin assessment. The incorporation of this new feature is an improvement that should result in a more accurate estimation of historical stock trends, relevant for the sustainable management of the fishery.

8.2 Recommendations for further work

The assessment identifies the following priorities:

8.2.1 Data conflict

The intention of the yellowfin stock assessor for this year's analysis was to address the data conflict between different likelihood components in greater detail than in the previous assessment. Due to a continuous stream of other issues and challenges in this year's assessment work, the analysis of data conflict had to be put aside, but it remains an important aspect to analyze further. It appears that the various structural changes introduced in the current assessment may have reduced the data conflict to some degree, but future work is required to quantify and analyze the data conflict between different fisheries and perhaps apply expert judgement to reduce the level of data conflict between specific fisheries.

In some cases, it might be as simple as being one fishery disagreeing with many fisheries, within the same region. In such cases, one might choose to put more faith on the many fisheries that are independently sampled and are in agreement, while downweighting the fishery that is at odds with the rest. One might conclude that there is empirical evidence that something is off in the sampling, data processing, and analytical assumptions for that particular fishery, making it less reliable about informing about population trends.

In other cases, it might be one fishery in disagreement with another fishery in the same region. In such cases, it is sometimes known that fishery A has a much better sampling coverage and is more likely to represent the underlying population trends than fishery B.

8.2.2 Regional aspects of the assessment

With the new regional scaling approach, it would be worthwhile to track the regional distribution of biomass and compare the estimated proportion of biomass in each region. The comparison could focus on the current assessment and the previous assessment, and also on stepwise models to evaluate whether the new regional scaling approach appears to give sensible estimates, how much the estimated proportions vary from the prior, and how stable are the estimated proportions, e.g., between different stepwise models.

The previous assessment simplified the regional structure from 9 to 5 regions, but SC discussions have reflected an interest in possible further simplifications. One option could be a single-region model with fleets as areas. This year, single-region models were developed both for yellowfin and bigeye. This analysis could be followed up and taken further, both analytically and as an important discussion at the interface between science and management. Another option that has been mentioned in the past would be to consider excluding the higher latitudes, which may cause model estimation difficulties and where less than 5% of the yellowfin catch comes from.

Analyses of alternative regional structures are best conducted outside of the main assessment season. The ongoing design and development of next-generation tuna assessment software has committed some time and focus on this topic. It remains an open question that has an important effect both from the standpoint of statistical estimation and for the management of tuna resources across widely differing regions.

8.2.3 Growth and natural mortality

The hard linking between the Lorenzen natural mortality curve and the growth curve may be a concern. For example, the estimated value of the L_1 parameter, the length at the youngest age of

1 quarter, in the diagnostic model is relatively low compared to the observed lengths at age 1 in the otolith data. The likelihood component that benefits from this low value of L_1 is the tagging data, because this low value of M produces a steep Lorenzen natural mortality curve that enables the model to fit the tagging data better. Overall, the estimation of L_1 seems somewhat unreliable, as its estimated value was observed to change substantially between stepwise models in this year's assessment.

8.2.4 Tag modeling options

This year's bigeye and yellowfin assessment work made it clear that different tag modeling options are an important issue as a source of model instability, sensitivity, local minima, and interactions between the various different model settings related to tags. Some tag related modeling issues were clarified and resolved this year, but many issues remain unclear and are highly influential for the estimated stock status.

Work is already underway to analyze tagging data externally, in a fine scale spatio-temporal model, and there is interest in designing the next-generation tuna assessment software with this pathway in mind. The external tagging analysis produces abundance indices, a simple data type that any assessment model can easily incorporate. At this point, it's not clear whether and how the complexities and decision points encountered in the current pathway of analyzing the tags will be carried over into the external tagging analysis.

If tagging data will be analyzed externally in future tuna assessments, it may provide several important benefits. For assessment software development, it should reduce cost and time for development and maintenance. For statistical estimation, simple abundance indices should reduce the tension between growth, M, and tags described above. For organizational purposes, it decouples the tagging analysis work from the stock assessment work, a combination that became quite complex and intractable in this year's assessments, causing substantial delays. If the external tagging analysis is conducted and completed before the main assessment work, it could make the overall process more efficient and reliable.

Overall, the above list of tag modeling options and issues might be best addressed with a dedicated workshop. Currently, the MULTIFAN-CL specific questions, caveats, and workarounds seem quite separate from the ongoing DTU development of the external tagging analysis. A tagging workshop could consolidate the discussion and make sure that the current stock assessment modeling challenges related to tags are acknowledged and addressed in some manner in the upcoming external tagging analysis.

8.3 Main assessment conclusions

The general conclusions of this assessment are as follows:

- The spawning potential of the stock has become more depleted across all model regions until around 2010. After 2010 there is several years of increase, followed by a decrease until the end of the assessment.

- The recent depletion trends are estimated to differ among model regions, with region 2 showing a sharp declining trend, and the others more stable.
- Average fishing mortality rates for juvenile and adult age classes have increased throughout the period of the assessment, although more so for juveniles which have experienced considerably higher and more variable fishing mortality than adults.
- Overall, median depletion from the uncertainty ensemble (with estimation uncertainty) for the recent period (2021–2024; $SB_{\text{recent}}/SB_{F=0}$) is estimated at 0.44 (80 percentile range 0.38–0.53, full range 0.28–0.64).
- No models from the uncertainty ensemble, including estimation uncertainty, estimate the stock to be below the LRP of $SB_{\text{recent}}/SB_{F=0} = 0.2$.
- CMM 2025-02 contains an objective to maintain the spawning biomass depletion ratio at or above the average $SB_t/SB_{t,F=0}$ for 2012–2015. Across the 48 ensemble models this average was 0.424. The corresponding average for 2021–2024 was 0.411, giving a recent-to-2012–2015 spawning depletion ratio of 0.97; stock status was therefore slightly below the objective.
- Recent (2020–2023) median fishing mortality relative to fishing mortality to achieve MSY $F_{\text{recent}}/F_{\text{MSY}}$ was 0.70 (80 percentile range, including estimation uncertainty 0.61–0.80, min 0.52, max 0.87)
- Assessment results suggest that the yellowfin stock in the WCPO is not overfished, nor undergoing overfishing according the WCPFC stock status reference points.

9 Acknowledgements

We thank WCPFC members, participating territories and other providers of the fisheries data, and the observers, port samplers, tagging teams and fishers who collected the biological and tagging observations. We also acknowledge the 2026 Pre-assessment Workshop participants and the Pacific Community data, assessment and computing support teams for their contributions, and the SPC schlerochronology lab team who aged many additional otoliths for this assessment. Thanks also to Japan FRA for providing the additional age-length data from the north Pacific. This assessment was supported by funding under the WCPFC regular 'scientific services' and 'additional resourcing SPC' budget lines to the scientific services provider, the budget line for the additional stock assessment scientist who also contributed to this assessment. The data preparation (CPUE) work was also supported by funding from European Union's "Pacific-European Union Marine Partnership Programme - 2" project.

10 Tables

Table 1: Definition of fisheries for the 2026 yellowfin stock assessment in the WCPO.

Fishery	Fishery label	Flag	Region	Sel group	% Catch last 10 yrs	% Catch all yrs
F1	1.LL.WEST.1	ALL	1	1	1.4	0.8
F2	2.LL.EAST.1	ALL	1	2	0.7	0.8
F3	3.LL.US.1	US	1	3	0.1	0.2
F4	4.LL.ALL.2	ALL	2	4	1.1	0.1
F5	5.LL.OS.2	OS	2	5	8.3	5.8
F6	6.LL.ALL.3	ALL	3	6	0.9	0.7
F7	7.LL.WEST.4	ALL	4	7	3.6	1.0
F8	8.LL.EAST.4	ALL	4	8	5.1	3.7
F9	9.LL.OS.4	OS	4	9	0.9	1.7
F10	10.LL.ALL.5	ALL	5	10	2.8	2.7
F11	11.LL.AU.5	AU	5	11	0.3	0.3
F12	12.PS.JP.1	JP	1	12	1.3	0.7
F13	13.PL.JP.1	JP	1	13	1.2	0.5
F14	14.HL.ID.2	ID	2	14	2.1	4.4
F15	15.HL.PH.2	PH	2	15	2.3	3.1
F16	16.PL.ALL.2	ALL	2	16	2.9	2.8
F17	17.PS.ID.2	ID	2	17	0.5	0.9
F18	18.PS.PH.2	PH	2	18	1.6	0.8
F19	19.PS.ASS.2	ALL	2	19	0.6	0.1
F20	20.PS.UNA.2	ALL	2	20	0.7	1.0
F21	21.DOM.ID.2	ID	2	21	9.4	16.7
F22	22.DOM.PH.2	PH	2	22	10.0	5.3
F23	23.DOM.VN.2	VN	2	23	0.5	0.9
F24	24.PL.ALL.3	ALL	3	24	0.3	0.0
F25	25.PS.ASS.3	ALL	3	25	5.7	3.2
F26	26.PS.UNA.3	ALL	3	26	7.4	12.6
F27	27.PL.ALL.4	ALL	4	27	0.1	0.0
F28	28.PS.ASS.WEST.4	ALL	4	19	10.1	6.0
F29	29.PS.ASS.EAST.4	ALL	4	28	4.7	6.4
F30	30.PS.UNA.WEST.4	ALL	4	20	8.9	10.5
F31	31.PS.UNA.EAST.4	ALL	4	29	4.4	6.4
Index fisheries						
F32	32.LL.IDX.1	IDX	1	30		
F33	33.LL.IDX.2	IDX	2	31		
F34	34.LL.IDX.3	IDX	3	32		
F35	35.LL.IDX.4	IDX	4	33		
F36	36.LL.IDX.5	IDX	5	34		

Table 2: Description of symbols used in the yield and stock status analyses.

Symbol	Description
C_{latest}	Catch in the last year of the assessment (2024)
F_{recent}	Average fishing mortality-at-age for a recent period (2020–2023)
$Y_{F_{\text{recent}}}$	Equilibrium yield at average fishing mortality for a recent period (2020–2023)
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}
$F_{\text{recent}}/F_{\text{MSY}}$	Average fishing mortality-at-age for a recent period (2020–2023) relative to F_{MSY}
SB_{latest}	Spawning biomass in the latest time period (2021)
SB_{recent}	Spawning biomass for a recent period (2021–2024)
$SB_{F=0}$	Average spawning biomass predicted in the absence of fishing for the period 2014–2023
SB_{MSY}	Spawning biomass that will produce the maximum sustainable yield (MSY)
$SB_{\text{MSY}}/SB_{F=0}$	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 2014–2023
$SB_{\text{latest}}/SB_{F=0}$	Spawning biomass in the latest time period (2021) relative to the average spawning biomass predicted to occur in the absence of fishing for the period 2014–2023
$SB_{\text{latest}}/SB_{\text{MSY}}$	Spawning biomass in the latest time period (2024) relative to that which will produce the maximum sustainable yield (MSY)
$SB_{\text{recent}}/SB_{F=0}$	Spawning biomass for a recent period (2021–2024) relative to the average spawning biomass predicted to occur in the absence of fishing for the period 2014–2023
$SB_{\text{recent}}/SB_{\text{MSY}}$	Spawning biomass for a recent period (2021–2024) 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 $t-1$ (2014–2023)

Table 3: Summary of reference points over the 48 individual models in the structural uncertainty grid, along with results incorporating estimation uncertainty (bottom)

Quantity	Mean	Median	Min	10%	90%	Max
F_{msy}	0.05	0.05	0.03	0.03	0.07	0.09
f_{mult}	1.44	1.44	1.14	1.24	1.64	1.91
F_{recent}/F_{msy}	0.70	0.69	0.52	0.61	0.80	0.87
MSY	709,492	712,800	610,000	657,080	758,040	838,800
SB_{latest}	4,225,501	3,601,588	1,449,490	1,781,367	7,185,084	12,284,825
SB_{recent}	4,239,072	3,595,731	1,546,934	1,845,127	7,163,028	12,139,288
SB_{F0}	9,116,300	8,293,368	4,138,987	5,214,403	13,617,615	20,203,280
SB_{latest}/SB_{F0}	0.44	0.44	0.28	0.37	0.53	0.61
SB_{latest}/SB_{msy}	1.72	1.73	1.11	1.49	1.89	2.10
SB_{msy}	2,409,656	2,076,000	758,100	1,246,700	3,799,600	6,074,000
SB_{msy}/SB_{F0}	0.26	0.25	0.18	0.23	0.29	0.30
$SB_{\text{recent}}/SB_{F=0}$	0.44	0.44	0.30	0.38	0.53	0.60
SB_{recent}/SB_{msy}	1.74	1.75	1.18	1.51	1.91	2.22
$YF_{current}$	167,600	166,700	150,500	158,050	178,030	192,400

Table 4: Ratio of 2021-2024 to 2012-2015

Quantity	Mean	Median	Min	10%	90%	Max
Ratio of SB/SB_{F0} 2021–2024 to 2012–2015	0.97	0.97	0.90	0.92	1.01	1.04

Table 5: Ensemble quantiles (with estimation uncertainty)

Quantity	Mean	Median	Min	10%	90%	Max
F_{recent}/F_{msy}	0.70	0.70	0.50	0.61	0.81	0.95
$SB_{\text{recent}}/SB_{F=0}$	0.44	0.44	0.28	0.38	0.53	0.64
SB_{recent}/SB_{msy}	1.74	1.75	0.95	1.48	2.00	2.73

11 Figures

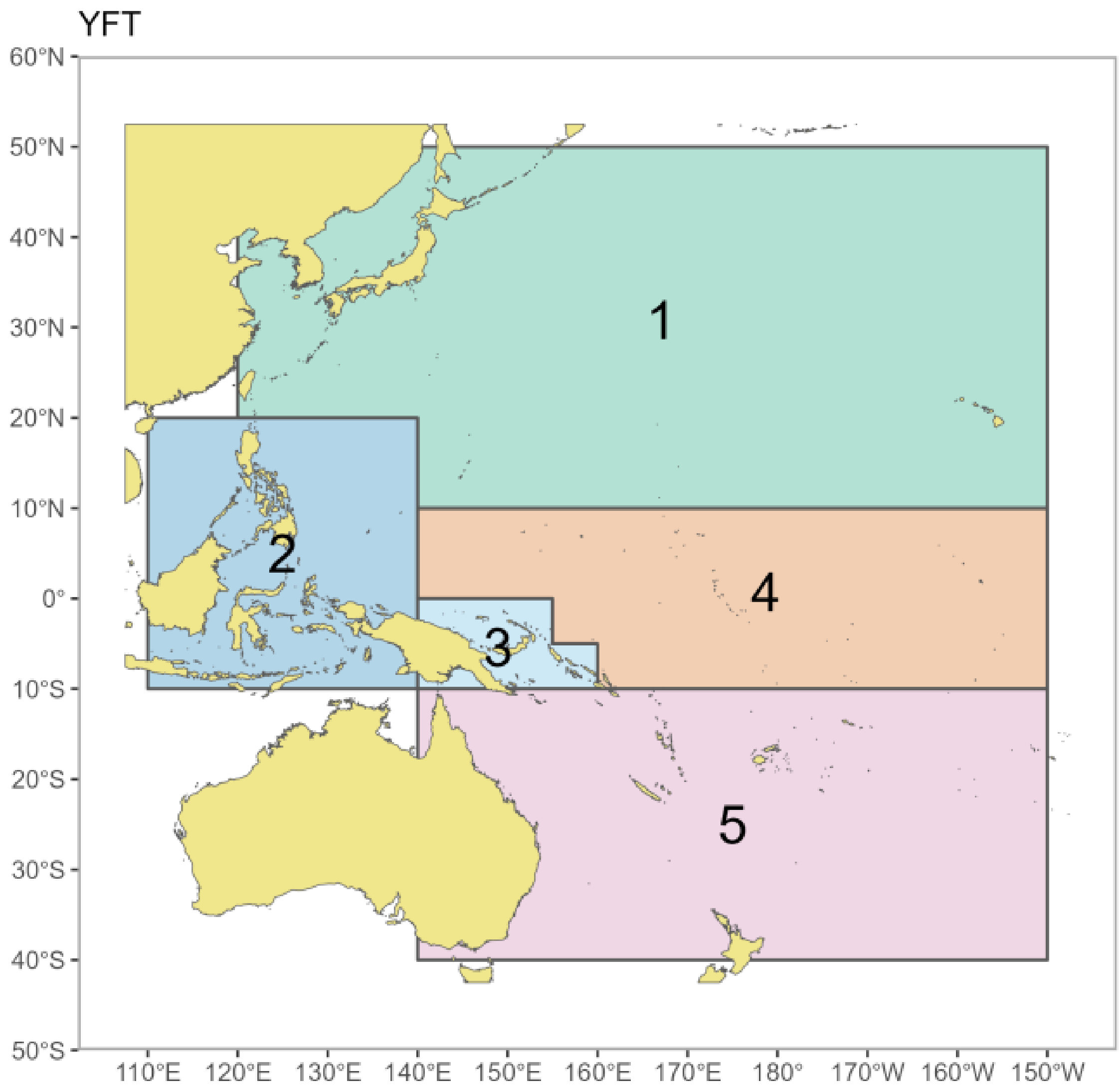

Figure 1: The geographical area covered by the stock assessment and the boundaries of the five model regions used for 2026 WCPO yellowfin tuna assessment.

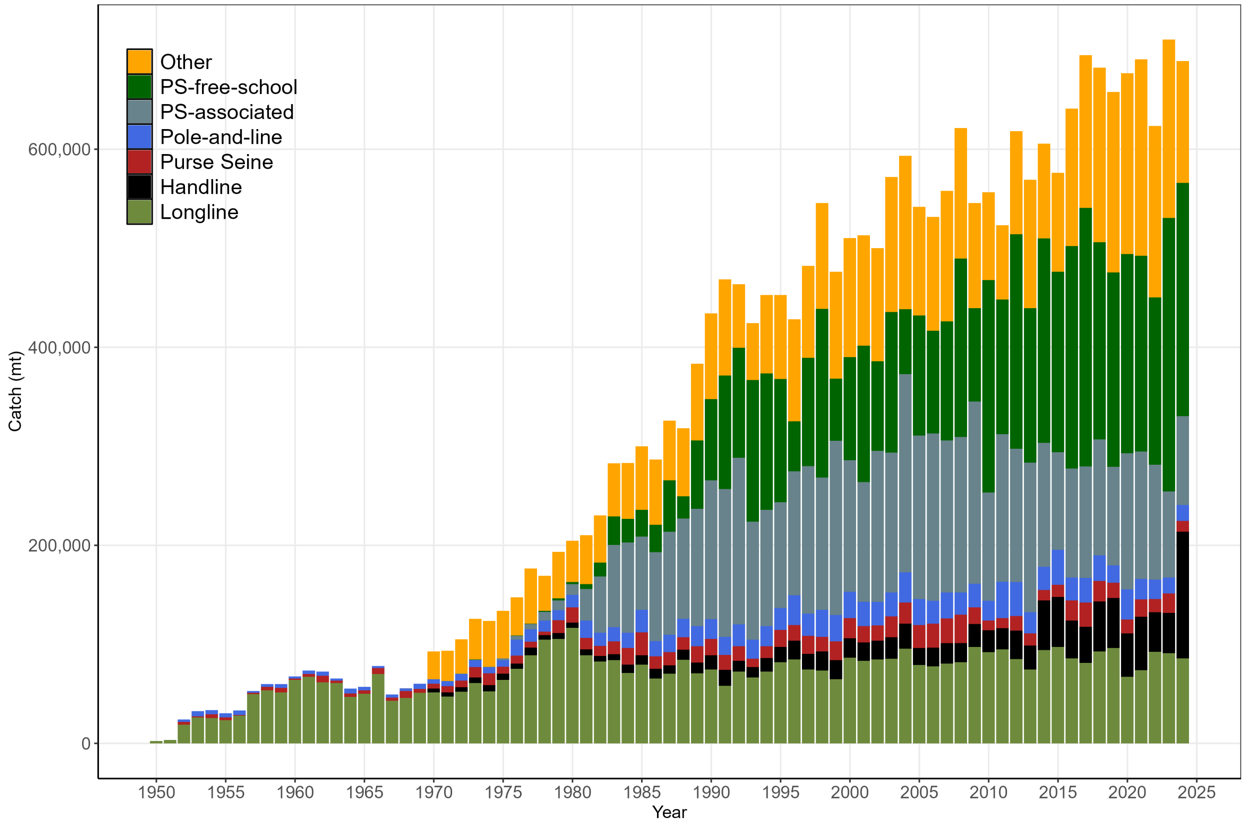

Figure 2: Annual catches of yellowfin by gear type in the WCPO area covered by the assessment.

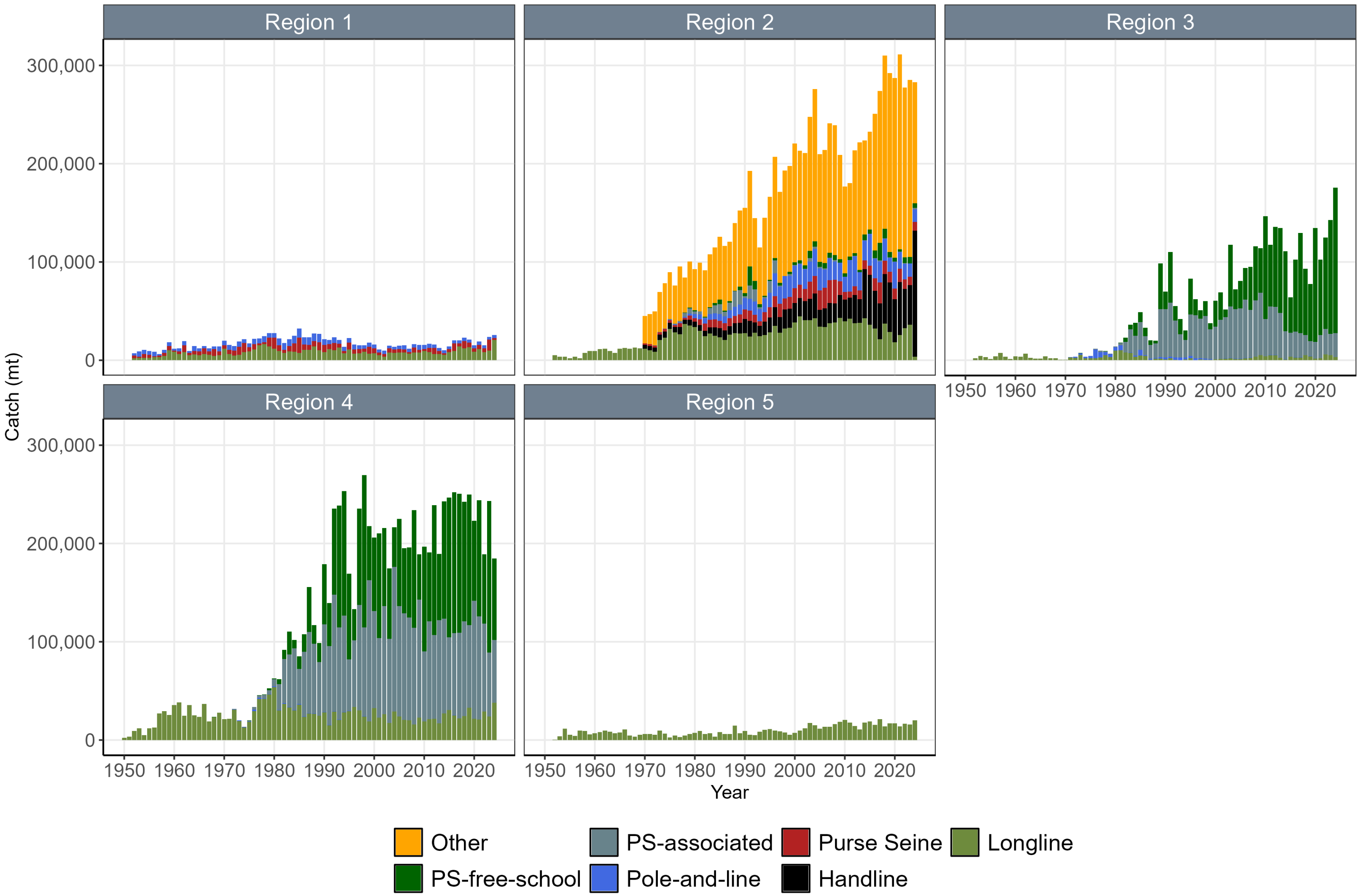

Figure 3: Annual catches of yellowfin by gear type for each of the five model regions.

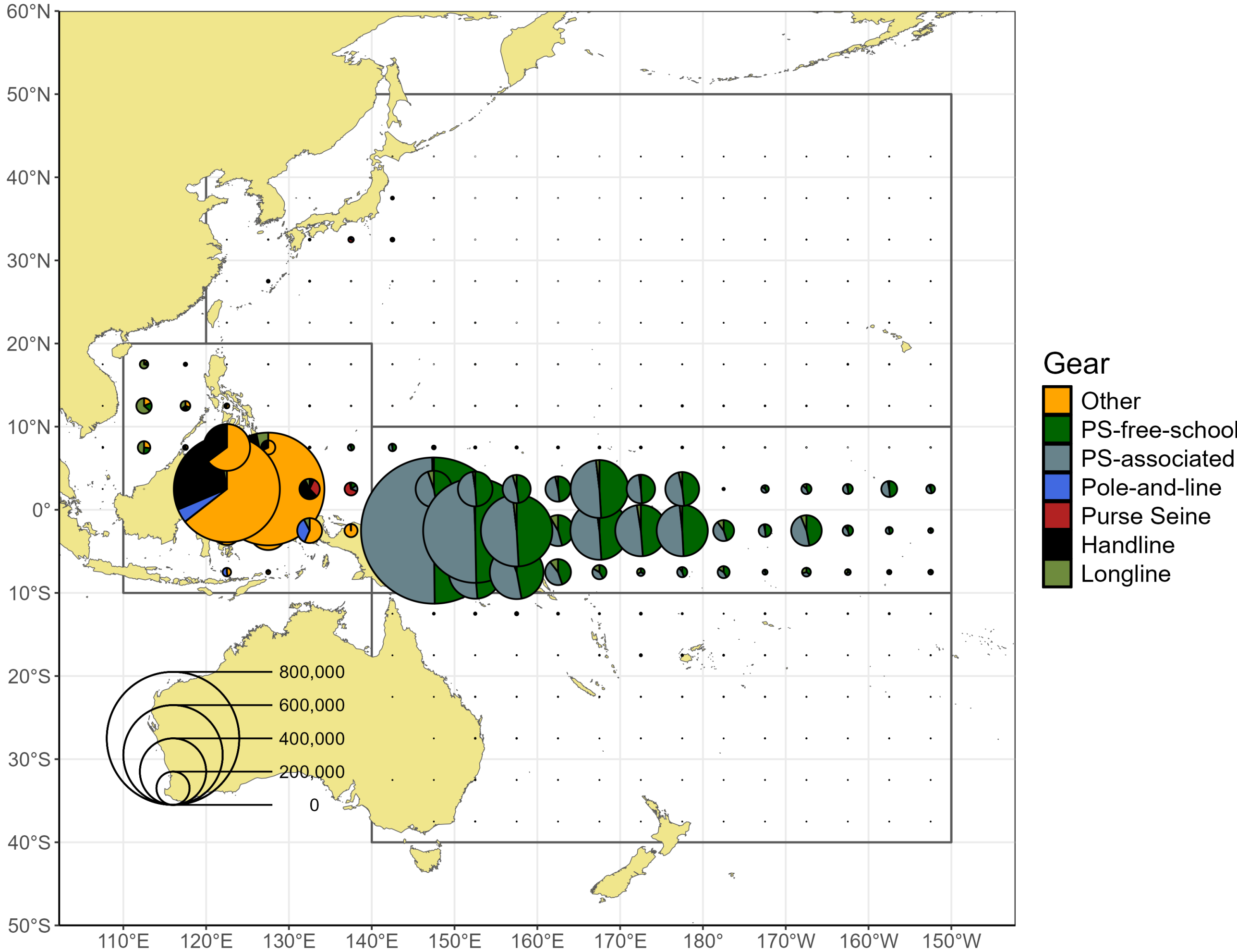

Figure 4: Distribution and magnitude of yellowfin catches (mt) by gear type summed over the last 10 years (2015-2024) for 5 x 5 degree cells.

Figure 5: Summary of data coverage by fishery for the WCPO 2026 yellowfin assessment.

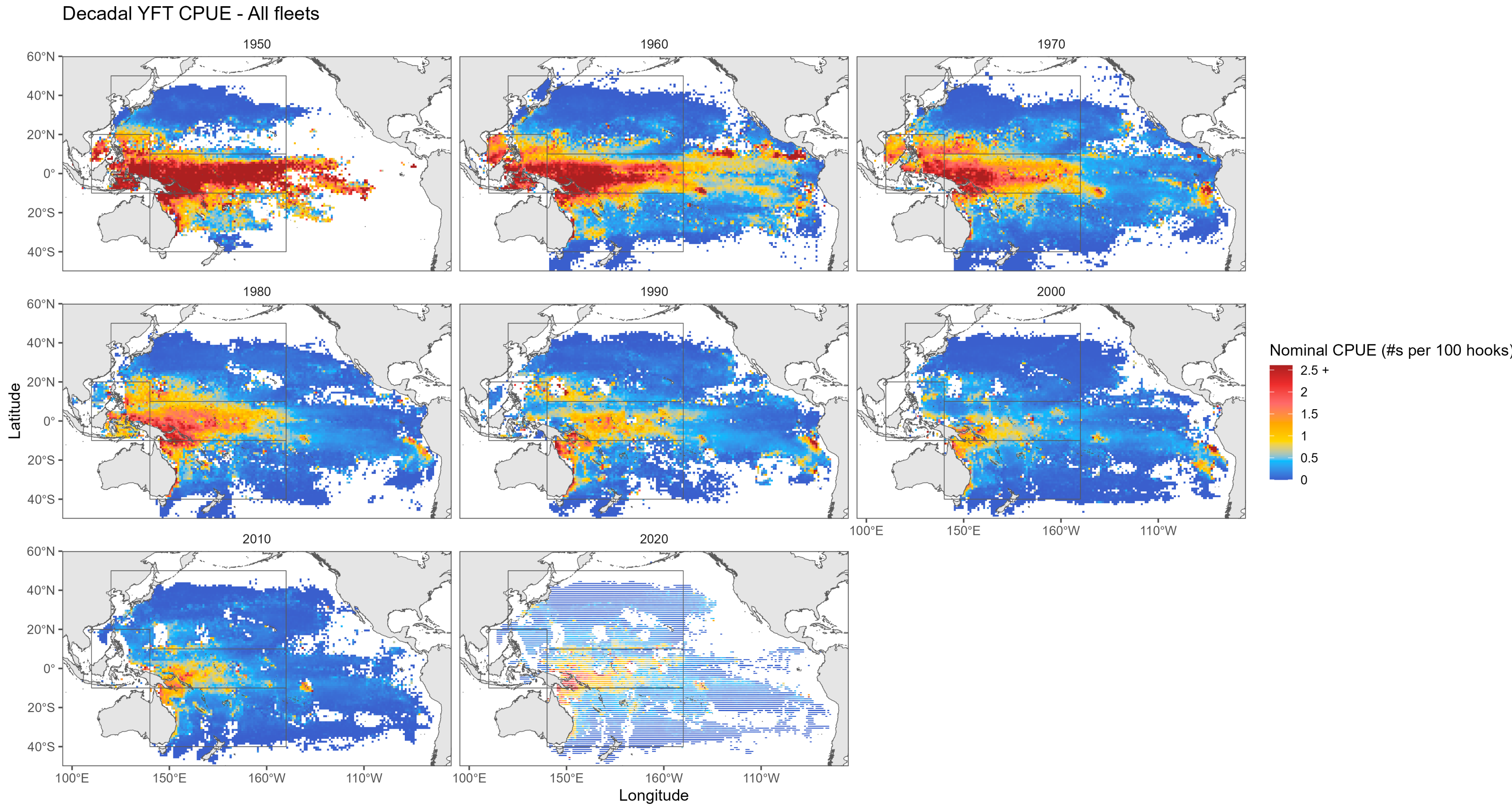

Figure 6: Spatial distribution of nominal longline CPUE (all fleets) for yellowfin in the Pacific.

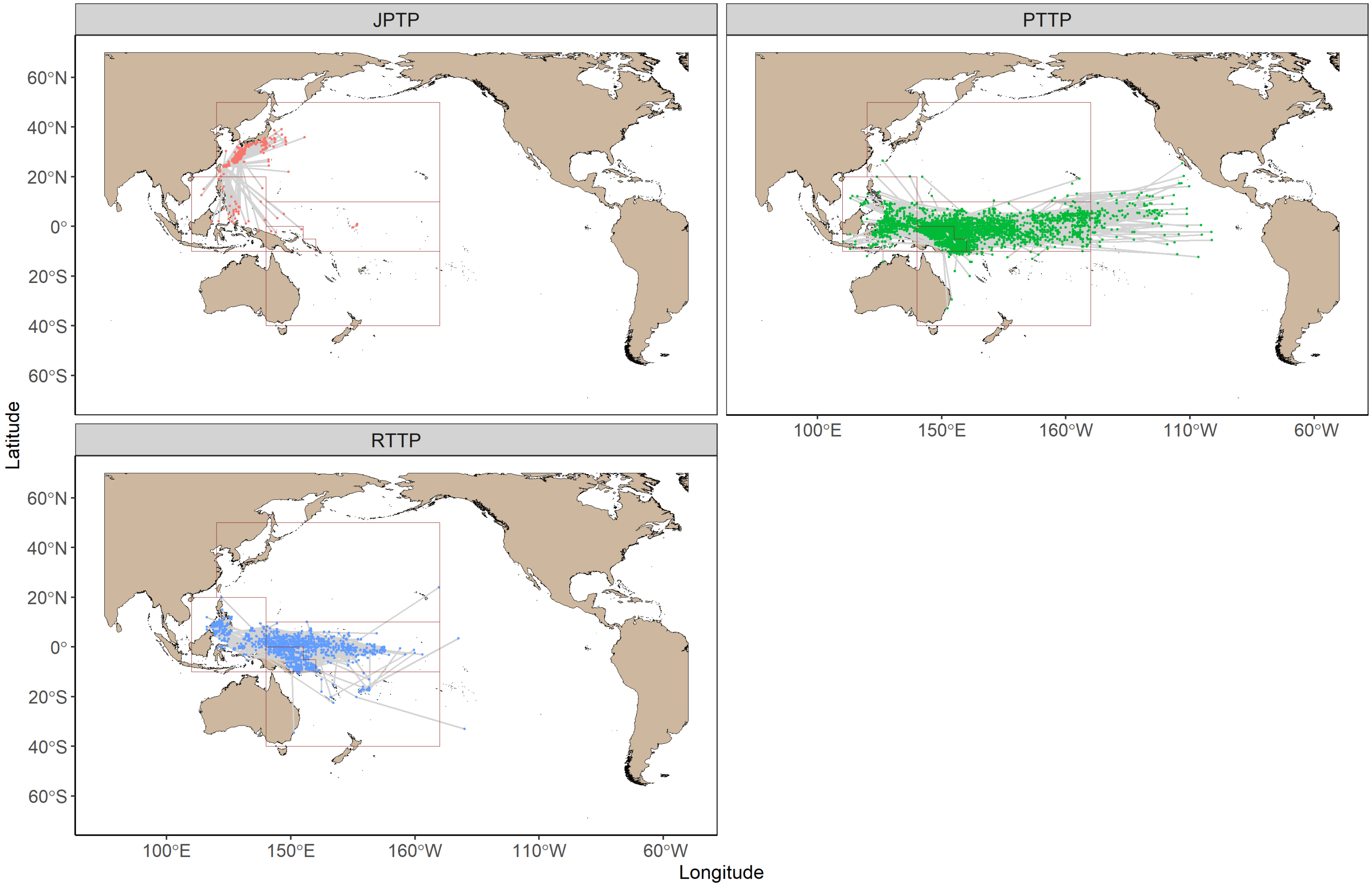

Figure 7: Map of tag recaptures. The panel shows the distributions of release and recapture displacements for the different tagging programs. Pacific Tuna Tagging Program (PTTP), Regional Tuna Tagging Program (excluding Coral Sea tags) (RTTP) and the Japanese Tagging Program (JPTP). Dots indicate recapture locations.

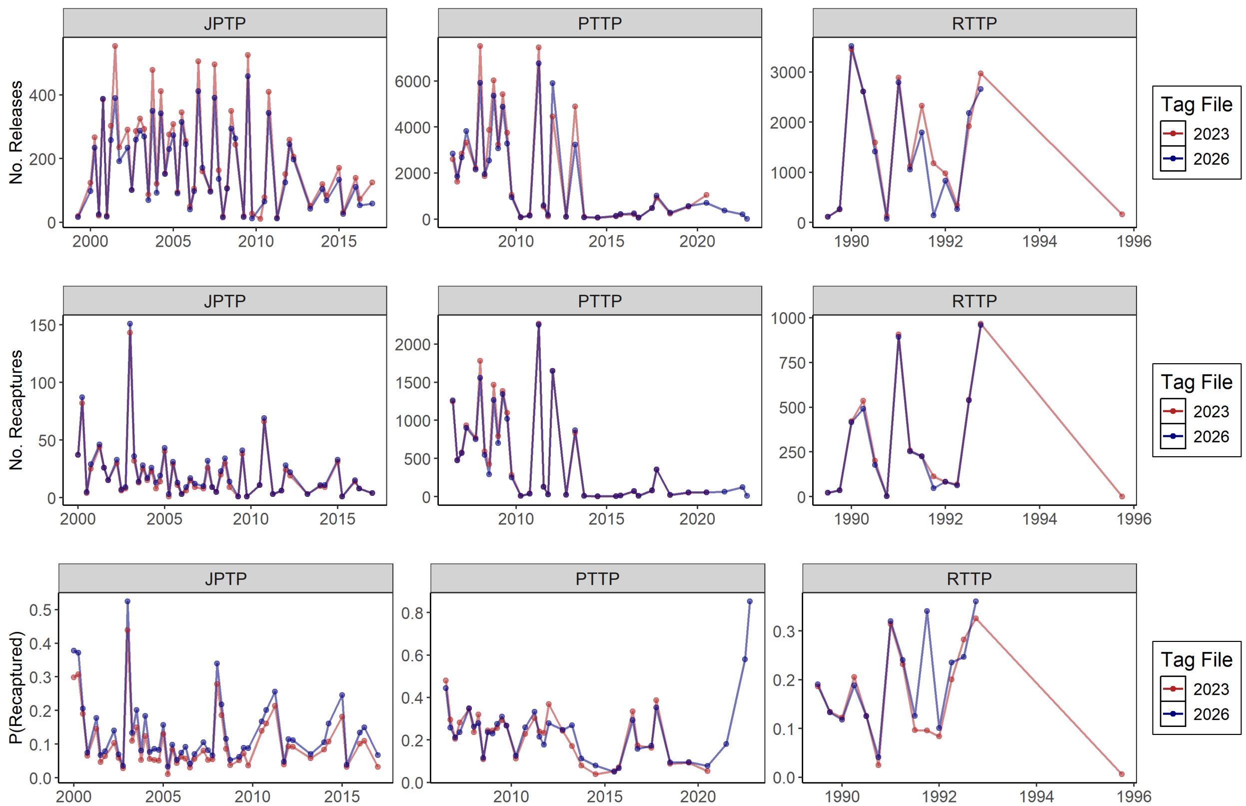

Figure 8: Summary plots of the number of releases, recaptures, and recapture rate of tags, by tagging program and region.

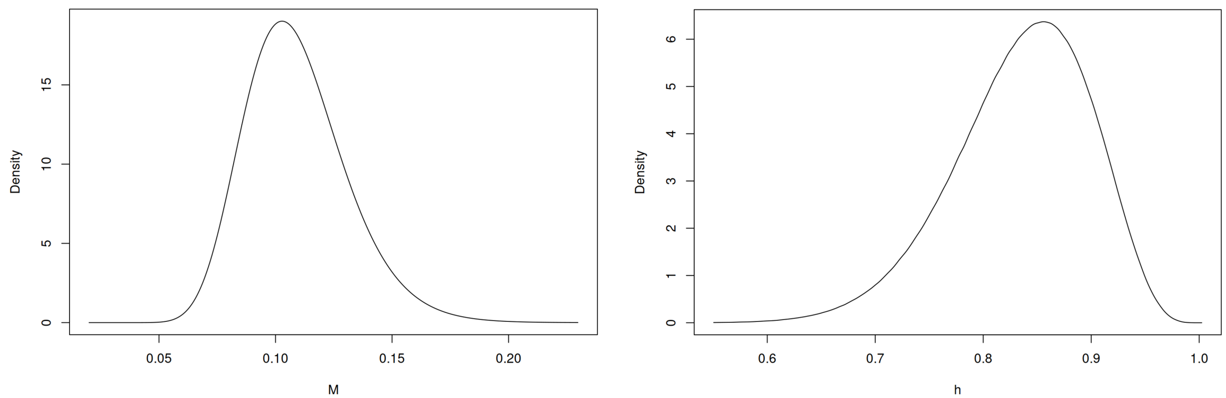

Figure 9: Beta distribution of (left) M at age 40 quarters, and (right) steepness, used for the uncertainty ensemble.

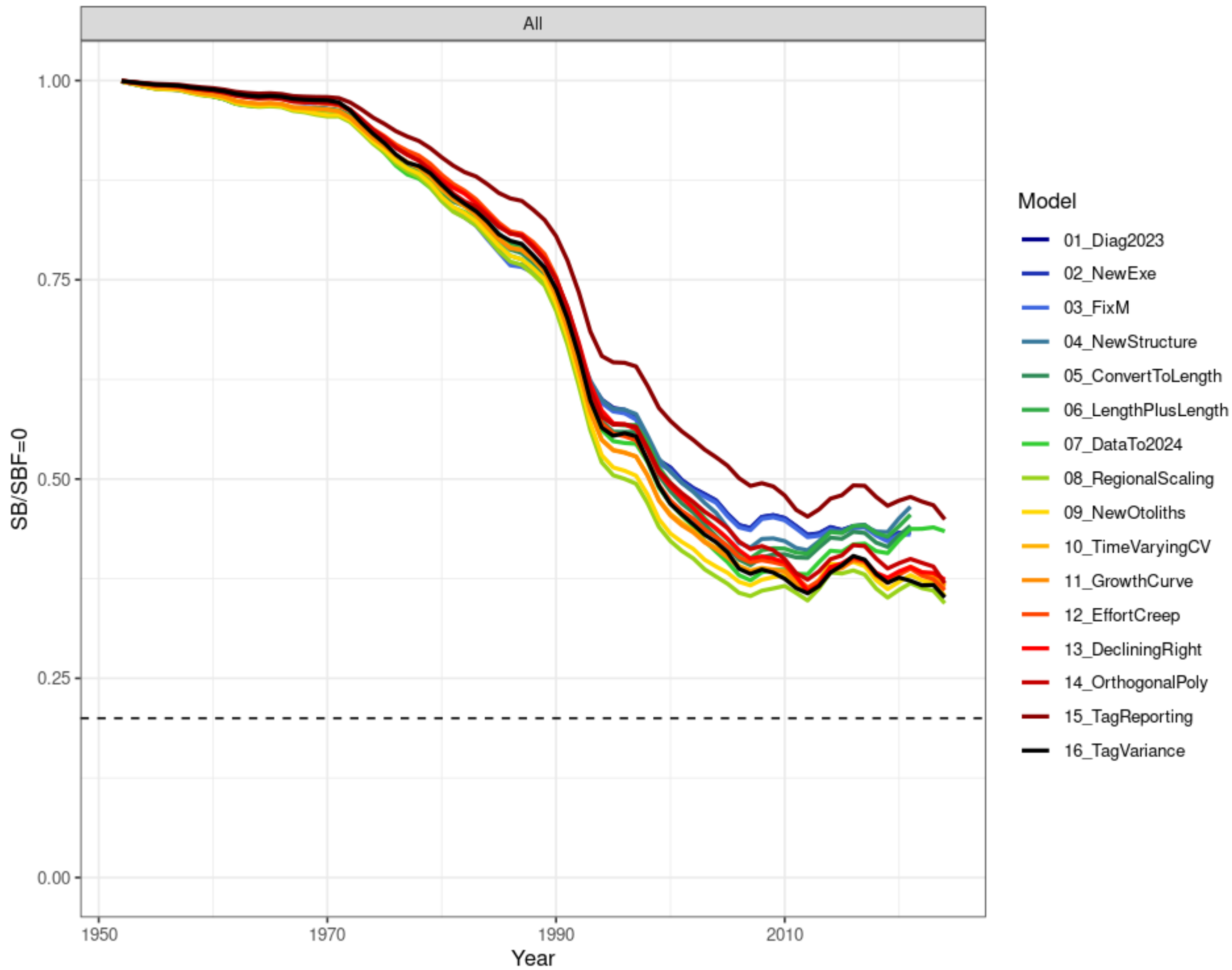

Figure 10: Stepwise model changes for depletion ($SB/SB_{F=0}$), model 16 black line is the diagnostic model

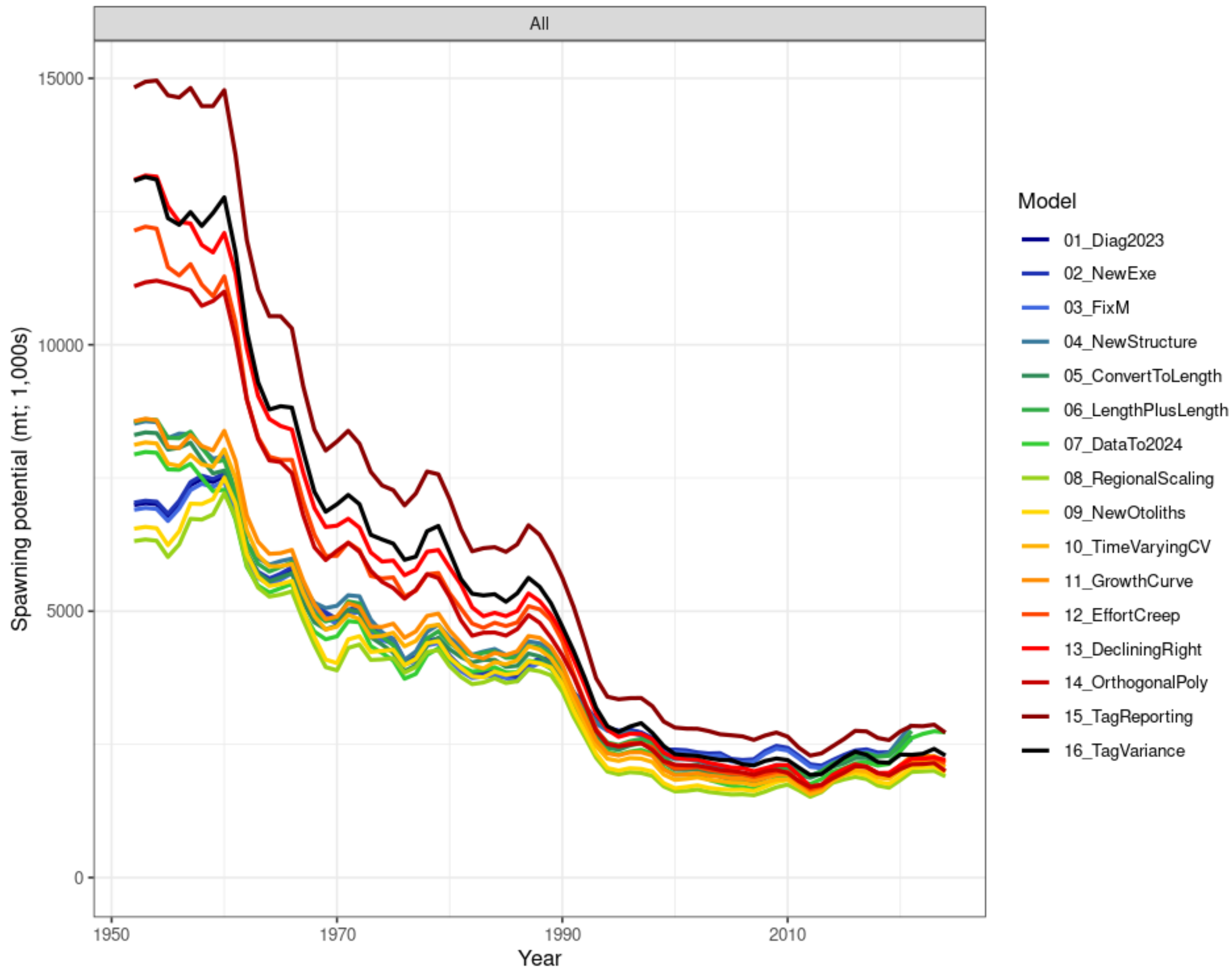

Figure 11: Stepwise model changes for spawning potential (SB), model 16 black line is the diagnostic model

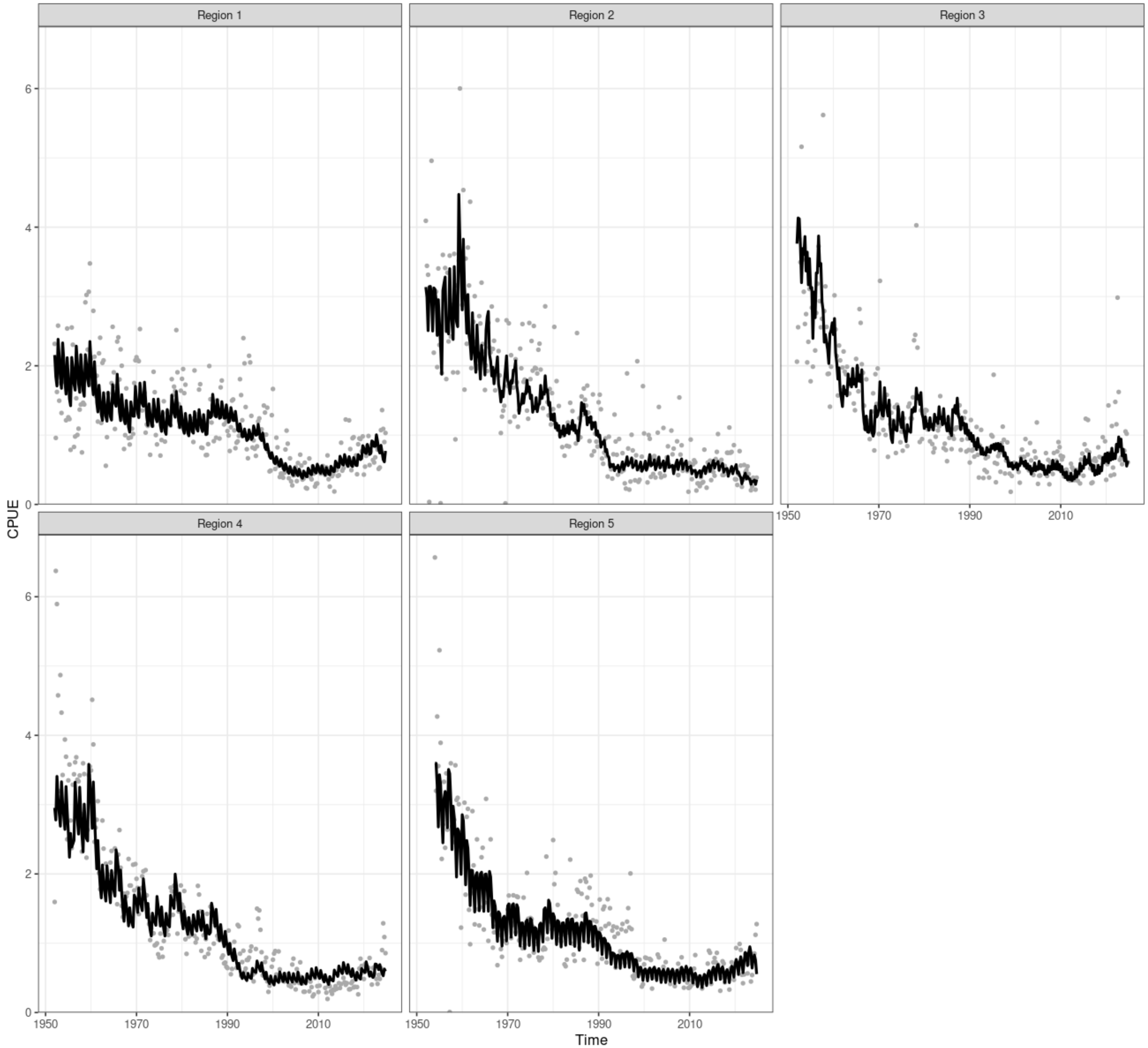

Figure 12: Fits (black line) to the standardised CPUE (gray dots) for the longline index fisheries in region 1-5.

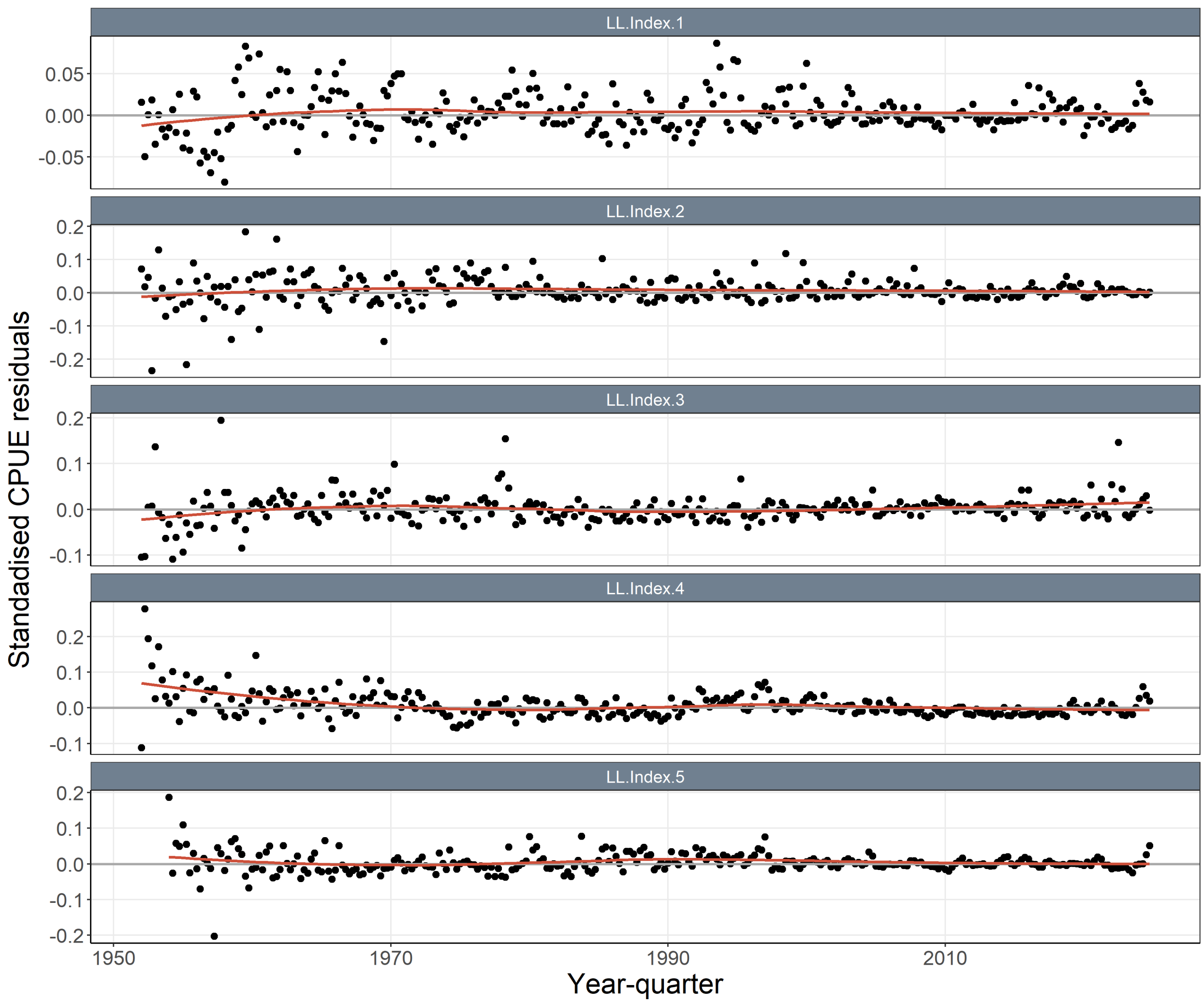

Figure 13: Residuals for the standardised CPUE fit for the longline index fisheries in region 1-5, gray points are the quarterly residuals, red line is a LOWESS smoother.

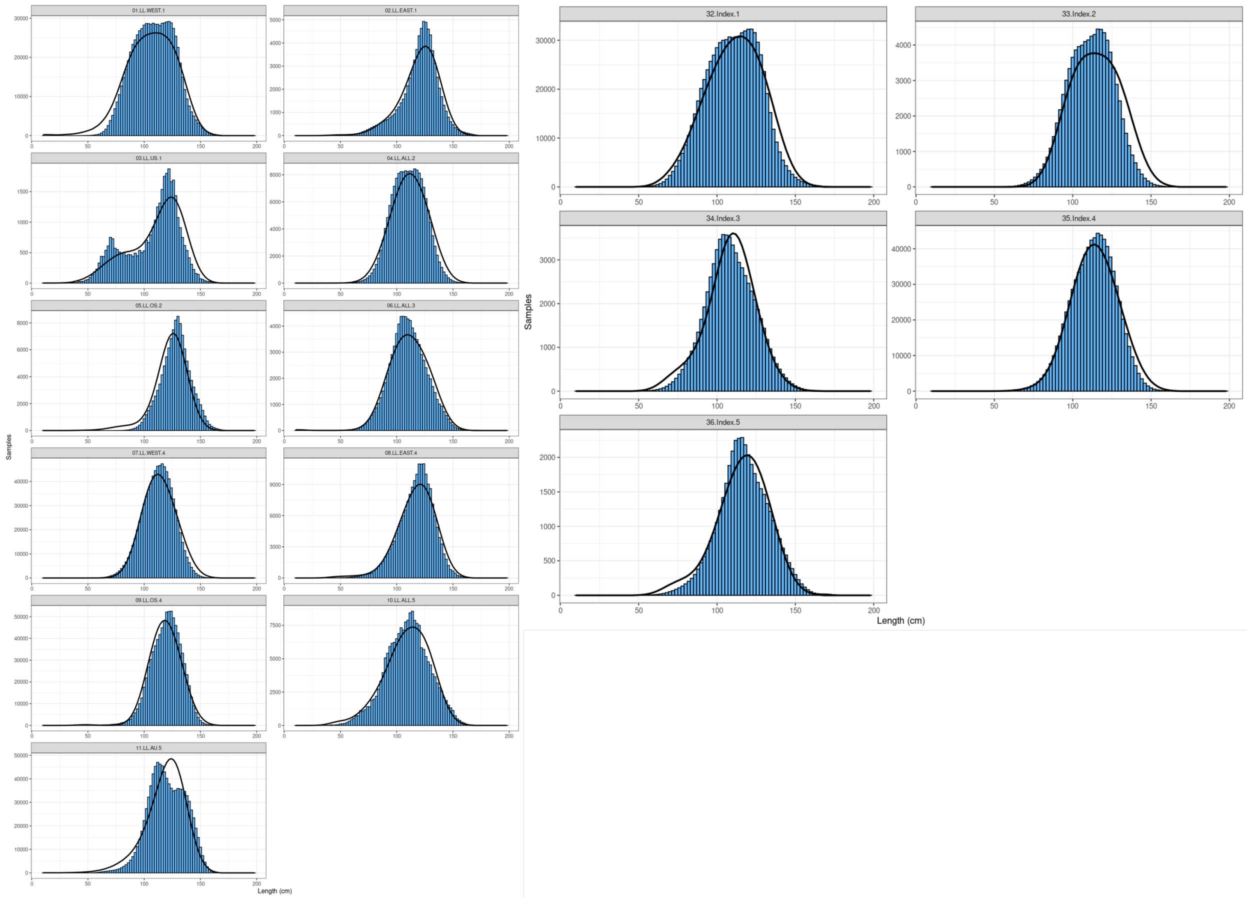

Figure 14: Composite (all time periods combined) observed (blue histograms) and predicted (black line) length frequency for fisheries with length frequency data for longline extraction and index fisheries in the 2026 diagnostic model.

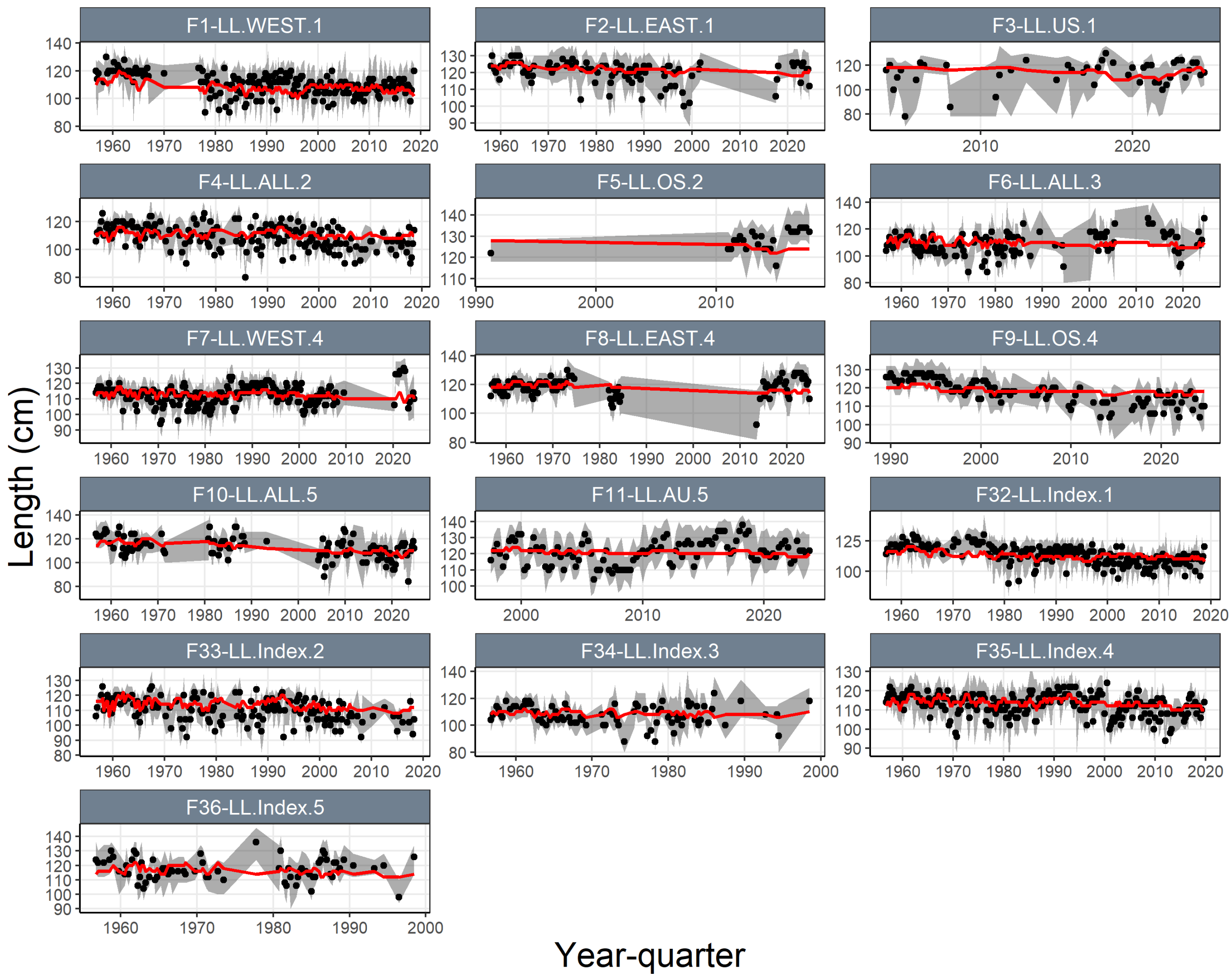

Figure 15: Observed (black points) and predicted (red line) median fish lengths (FL, cm) for the longline fisheries for the 2026 diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples with more than 250 fish per quarter/fishery strata are plotted.

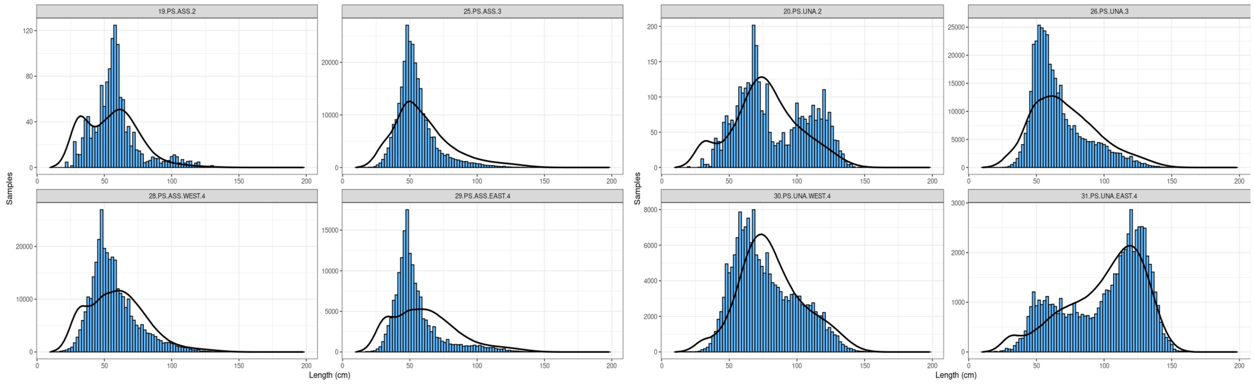

Figure 16: Composite (all time periods combined) observed (blue histograms) and predicted (black line) length frequency for industrial purse seine fisheries in the 2026 diagnostic model.

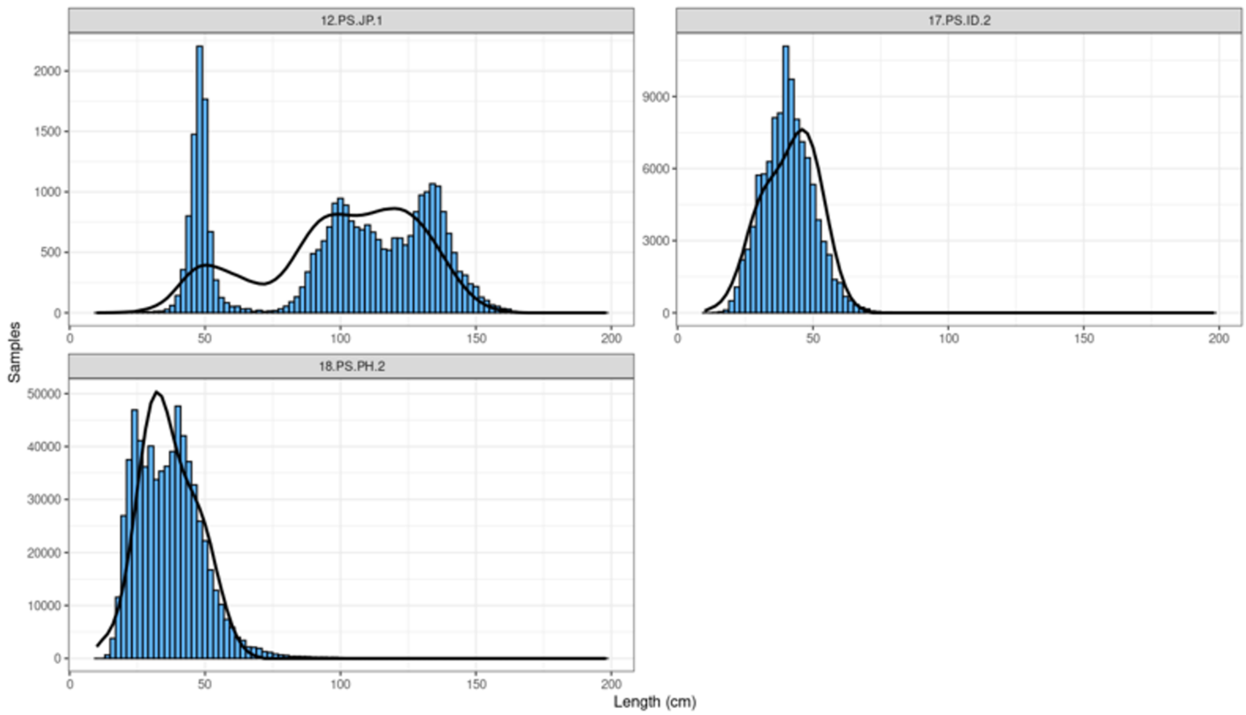

Figure 17: Composite (all time periods combined) observed (blue histograms) and predicted (black line) length frequency for domestic purse seine fisheries in the 2026 diagnostic model.

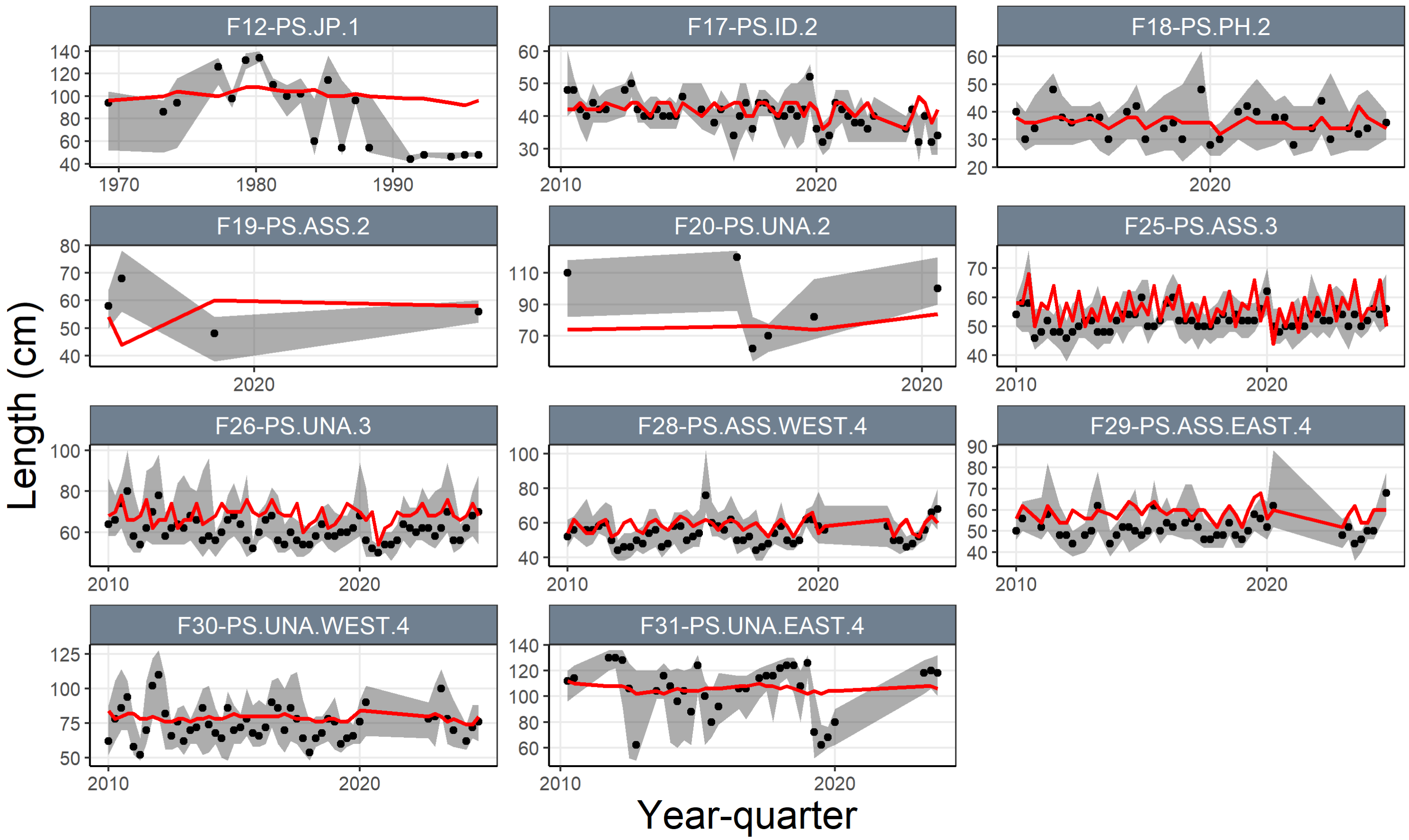

Figure 18: Observed (black points) and predicted (red line) median fish lengths (FL, cm) for all purse seine fisheries for the 2026 diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples with more than 100 fish per quarter/fishery strata are plotted.

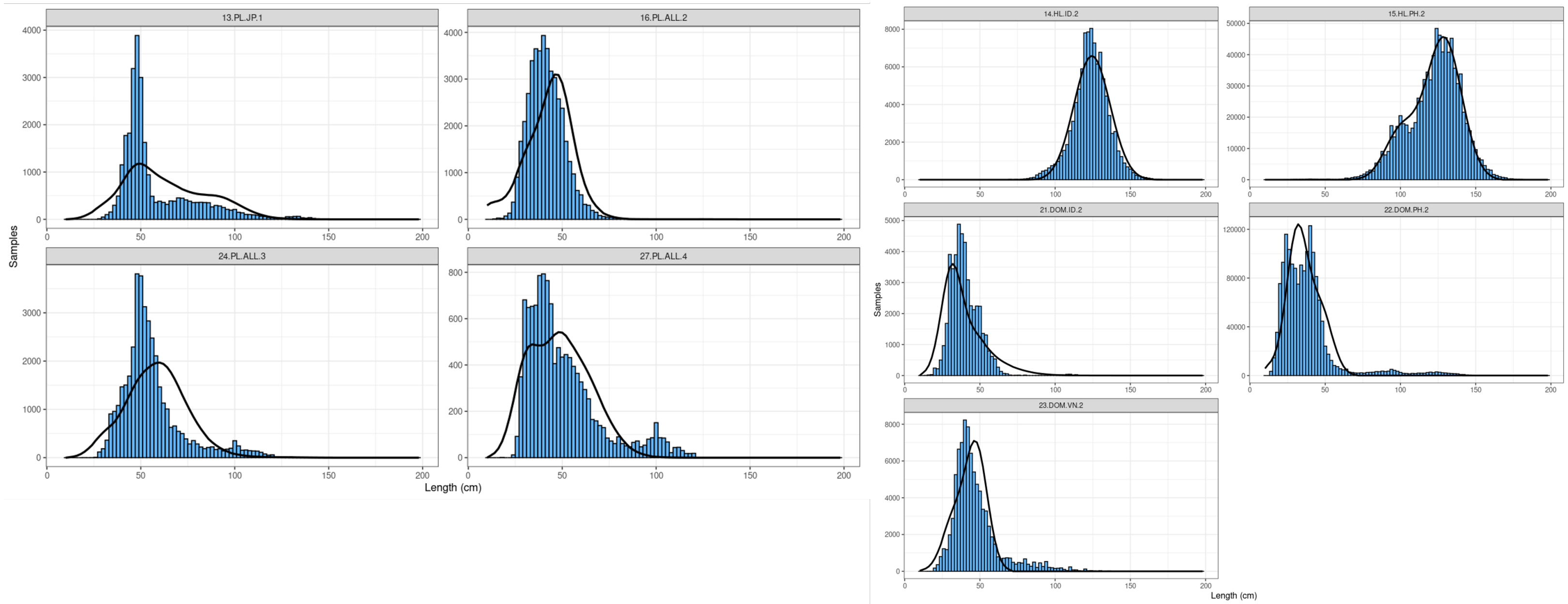

Figure 19: Composite (all time periods combined) observed (blue histograms) and predicted black line) length frequency for domestic other gears, handline and pole and line fisheries in the 2026 diagnostic model.

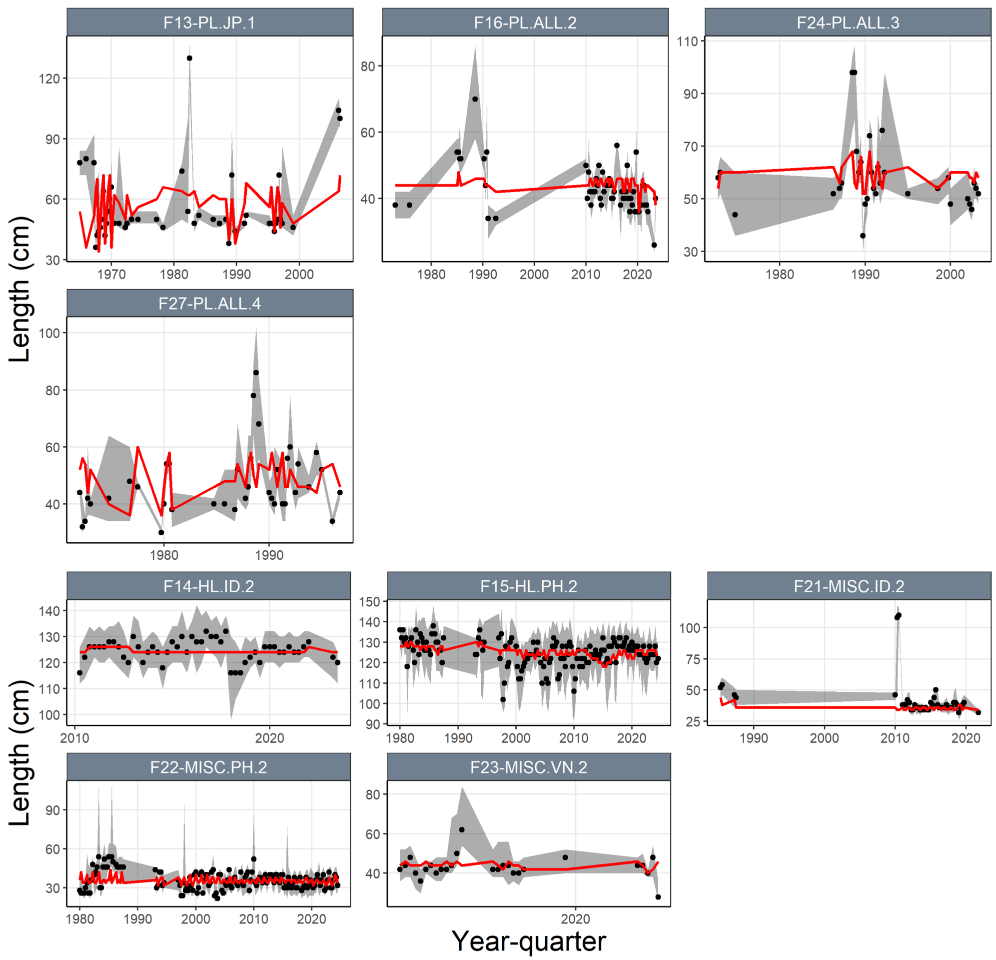

Figure 20: Observed (black points) and predicted (red line) median fish lengths (FL, cm) for domestic other gears, handline and pole and line fisheries for the 2026 diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples with more than 100 fish per quarter/fishery strata are plotted.

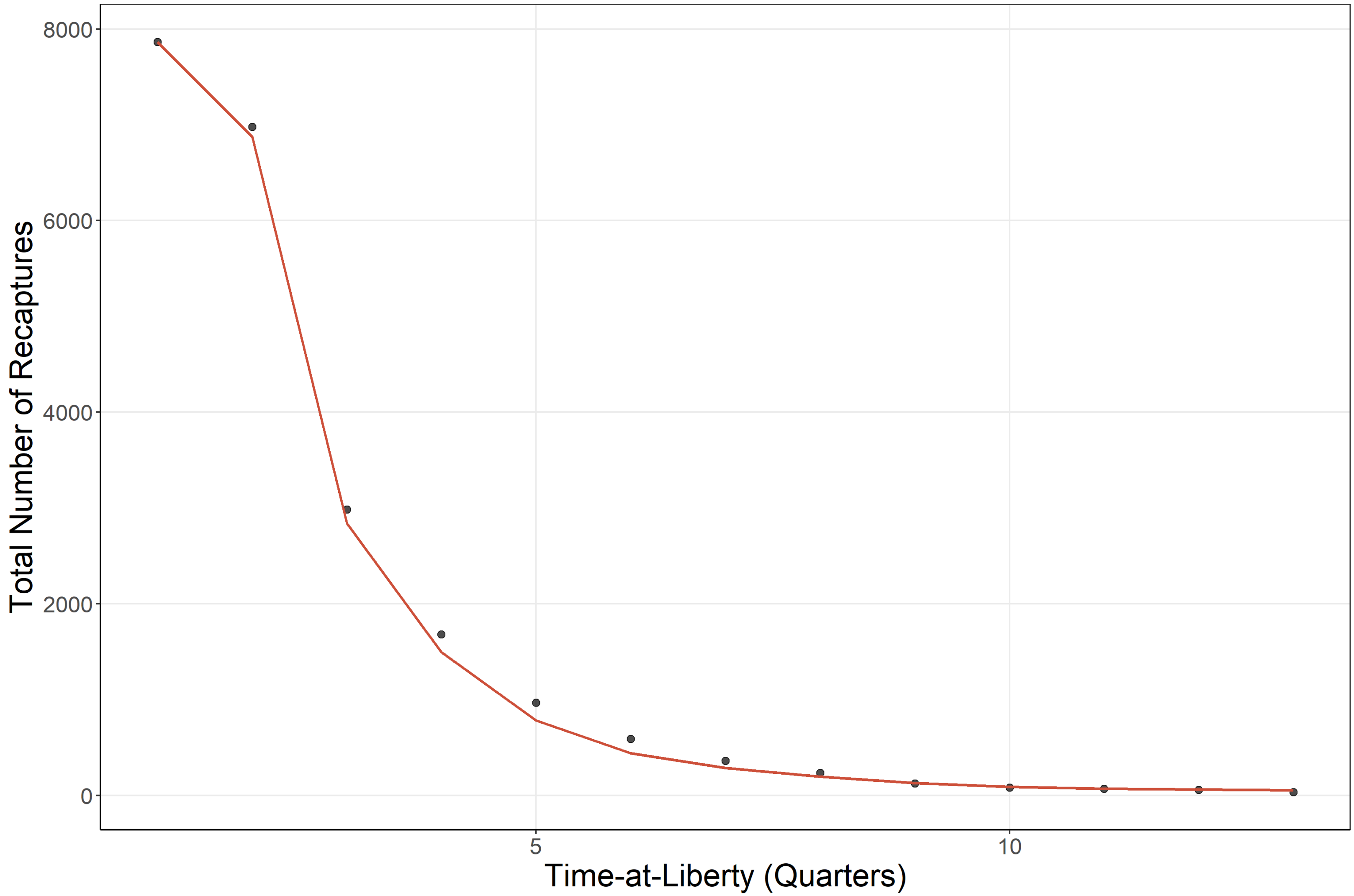

Figure 21: Observed (black points) and model-predicted (red line) tag attrition across all tag release events for the 2026 diagnostic model.

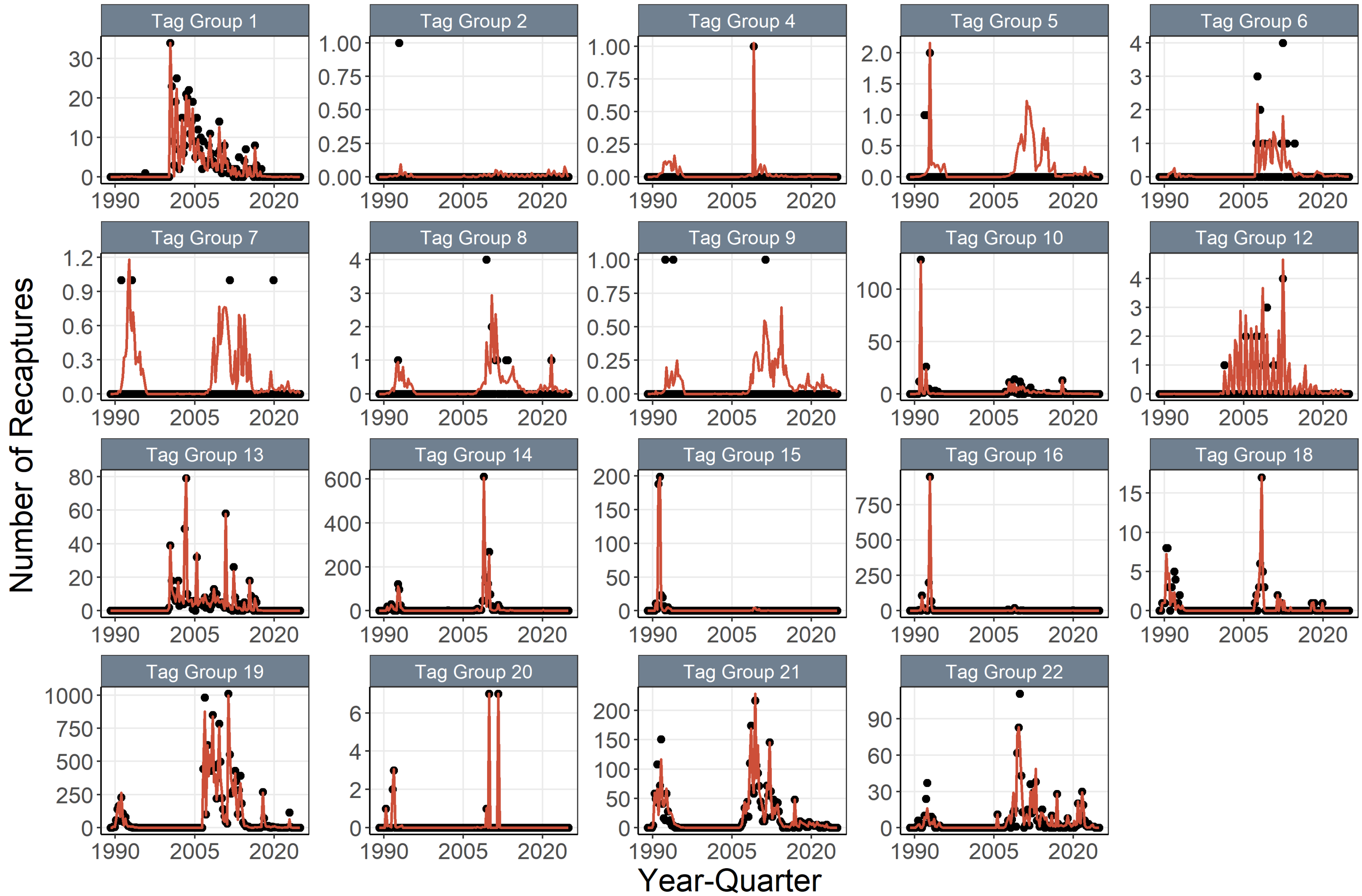

Figure 22: Observed (black points) and model-predicted (red lines) tag attrition by tagging programme for the 2026 diagnostic model.

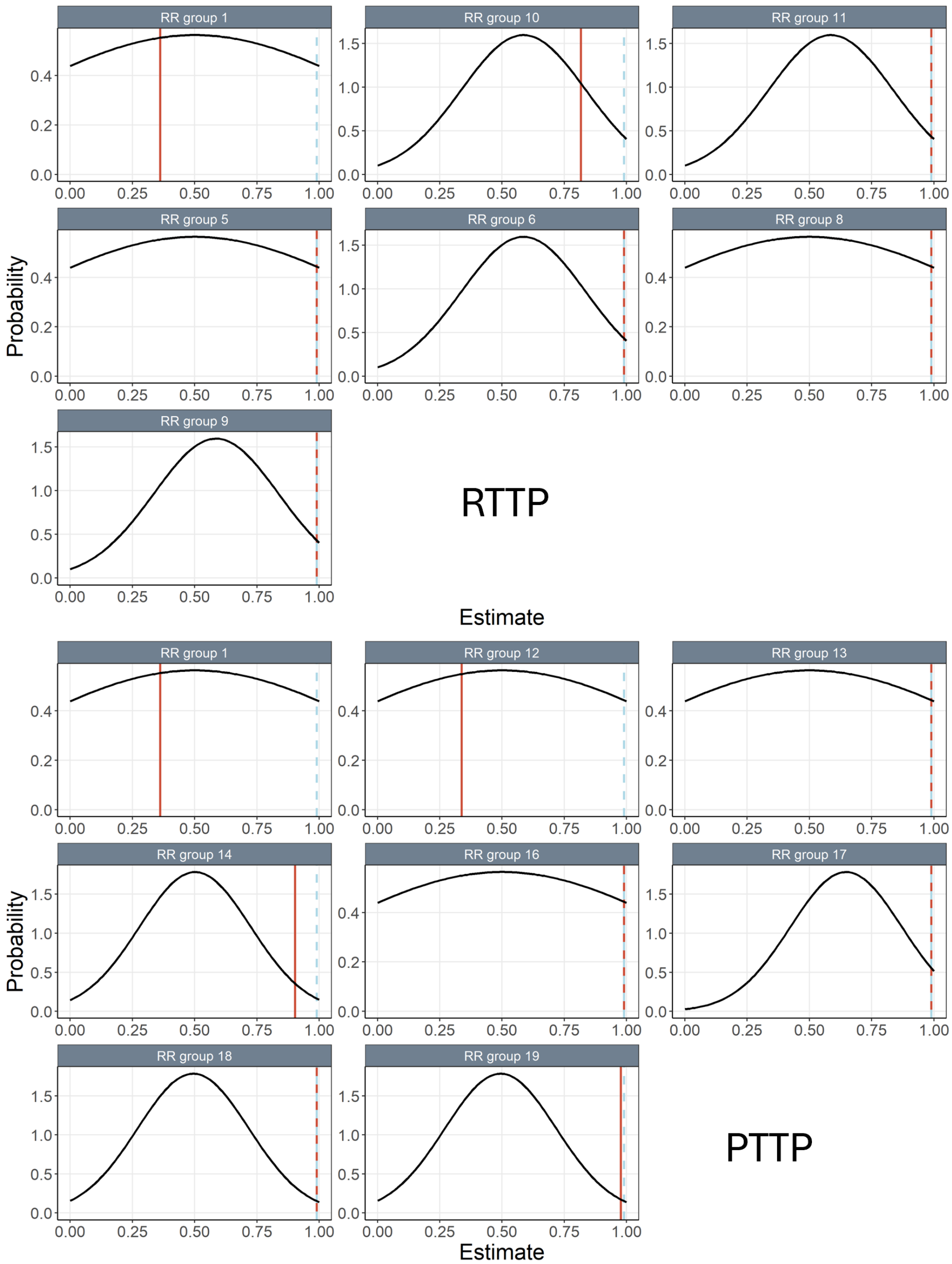

Figure 23: Estimated reporting rates for RTTP and PTTP for the diagnostic model (red lines) and the prior distribution for each tag reporting rate group (black lines). The upper bound (0.99) on the reporting rate parameters is shown as a blue dashed line. Reporting rates can be estimated separately for each release program and recapture fishery group but in practice are aggregated over some recapture groups to reduce dimensionality.

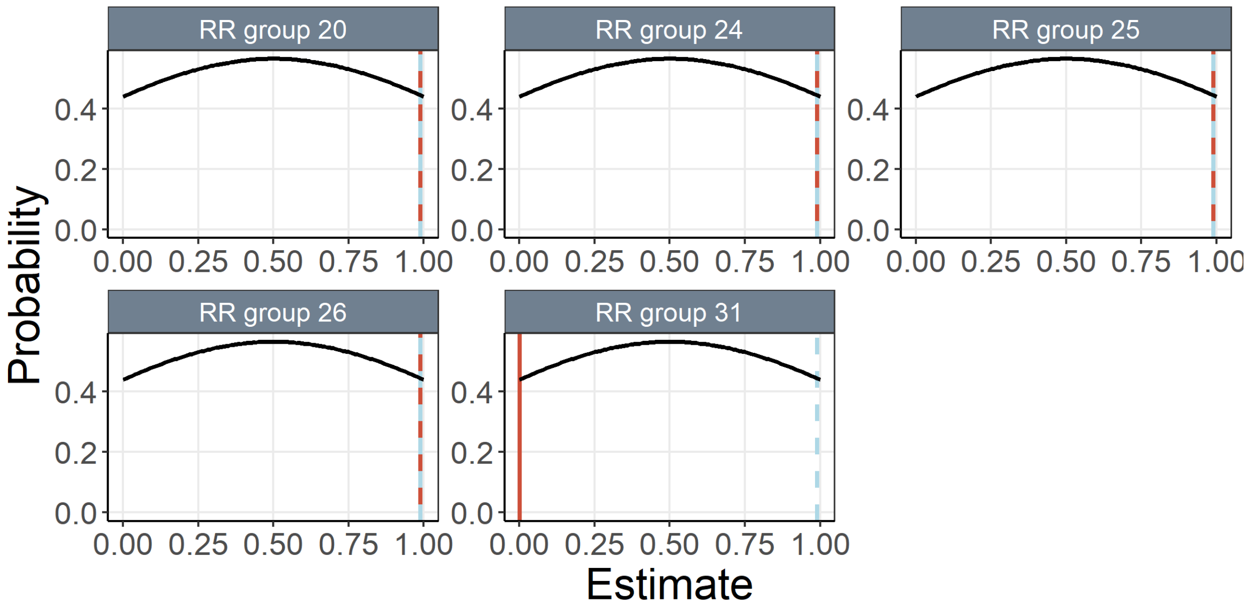

Figure 24: Estimated reporting rates for JPTP for the diagnostic model (red lines) and the prior distribution for each tag reporting rate group (black lines). The upper bound (0.99) on the reporting rate parameters is shown as a blue dashed line. Reporting rates can be estimated separately for each release program and recapture fishery group but in practice are aggregated over some recapture groups to reduce dimensionality.

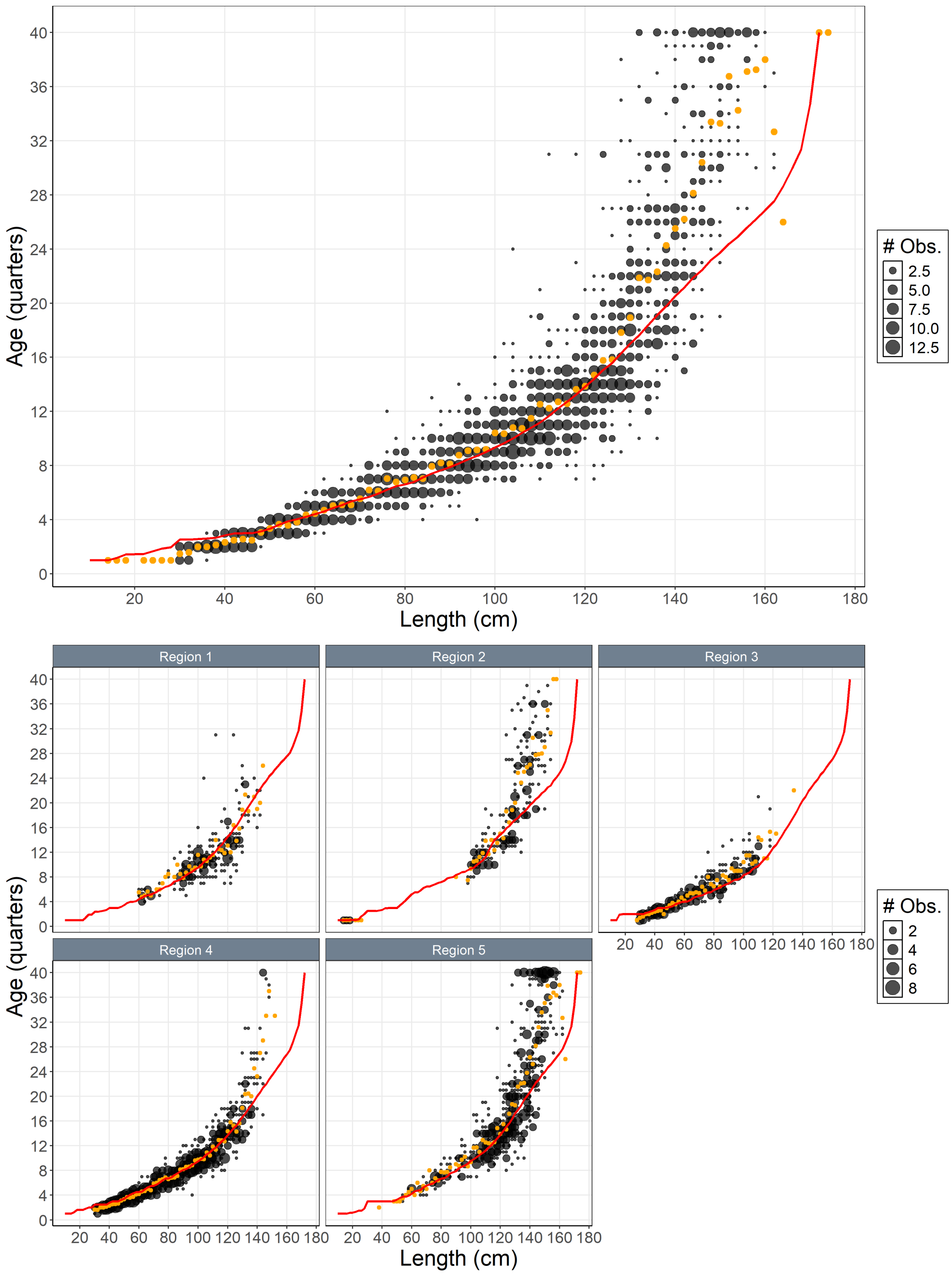

Figure 25: Model estimated median age at length (orange line), versus observed age at length (black circles), and median observed age at lengths (orange circles), aggregate across all samples (top) and by model regions (bottom), The distribution of observed age at length is viewed by the vertical distribution of the black dots.

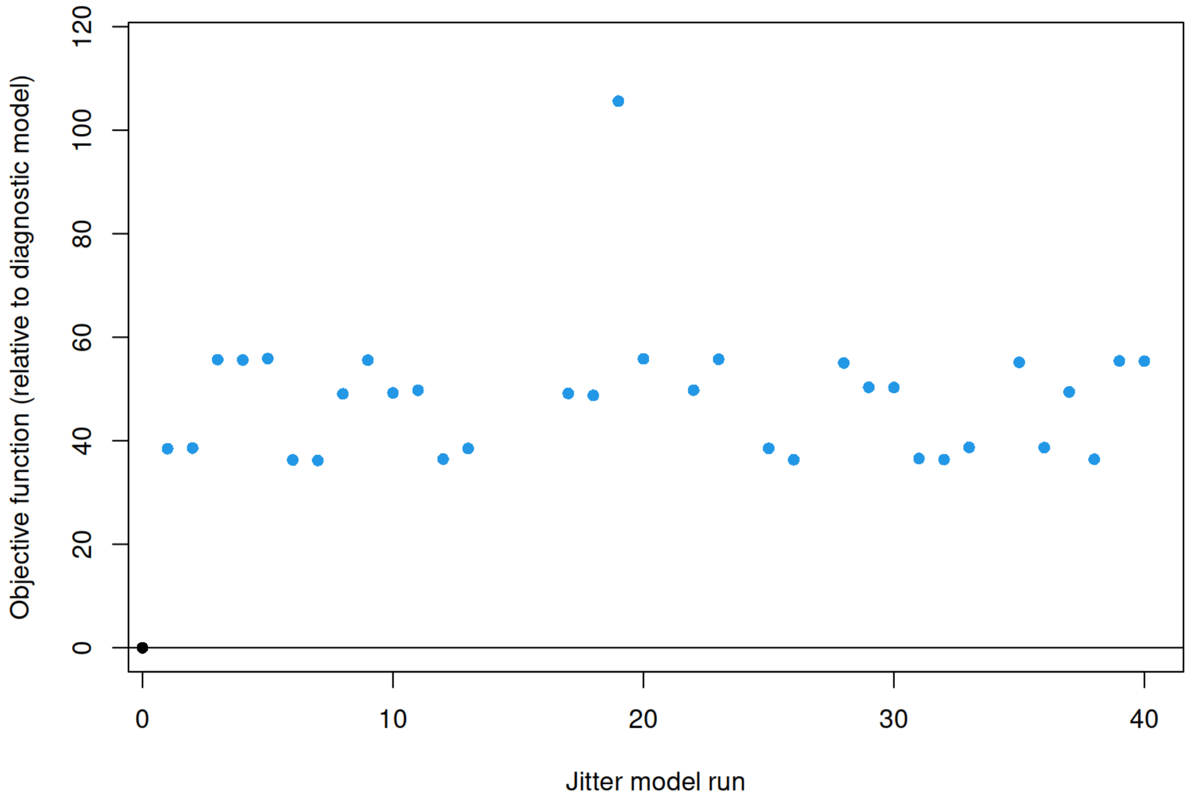

Figure 26: Jitter convergence diagnostics for the diagnostic model. Each blue point represents one jitter run. The black dot identifies the Diagnostic model without jitter.

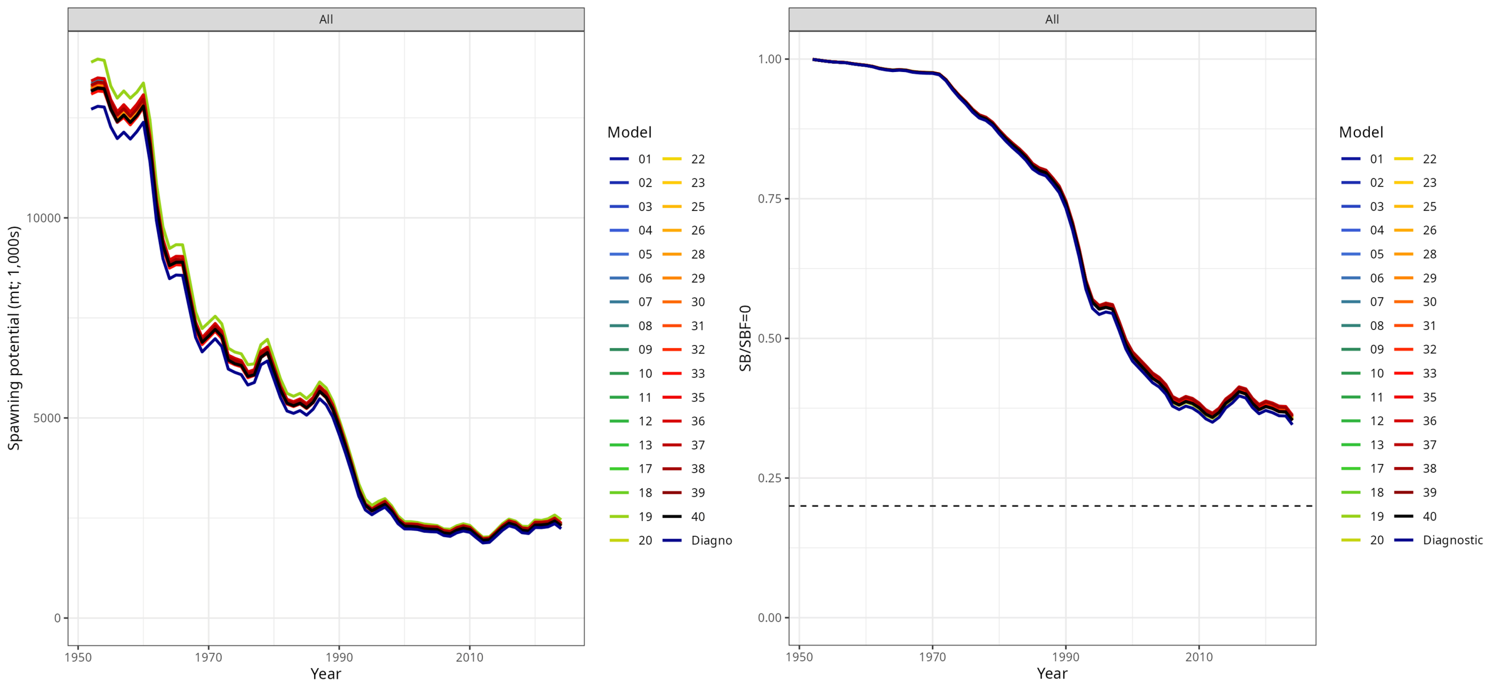

Figure 27: Annual derived quantities across the 33 retained jitter fits and the diagnostic are shown for (left) spawning potential, SB, and (right) depletion, $(SB/SB_{F=0})$.

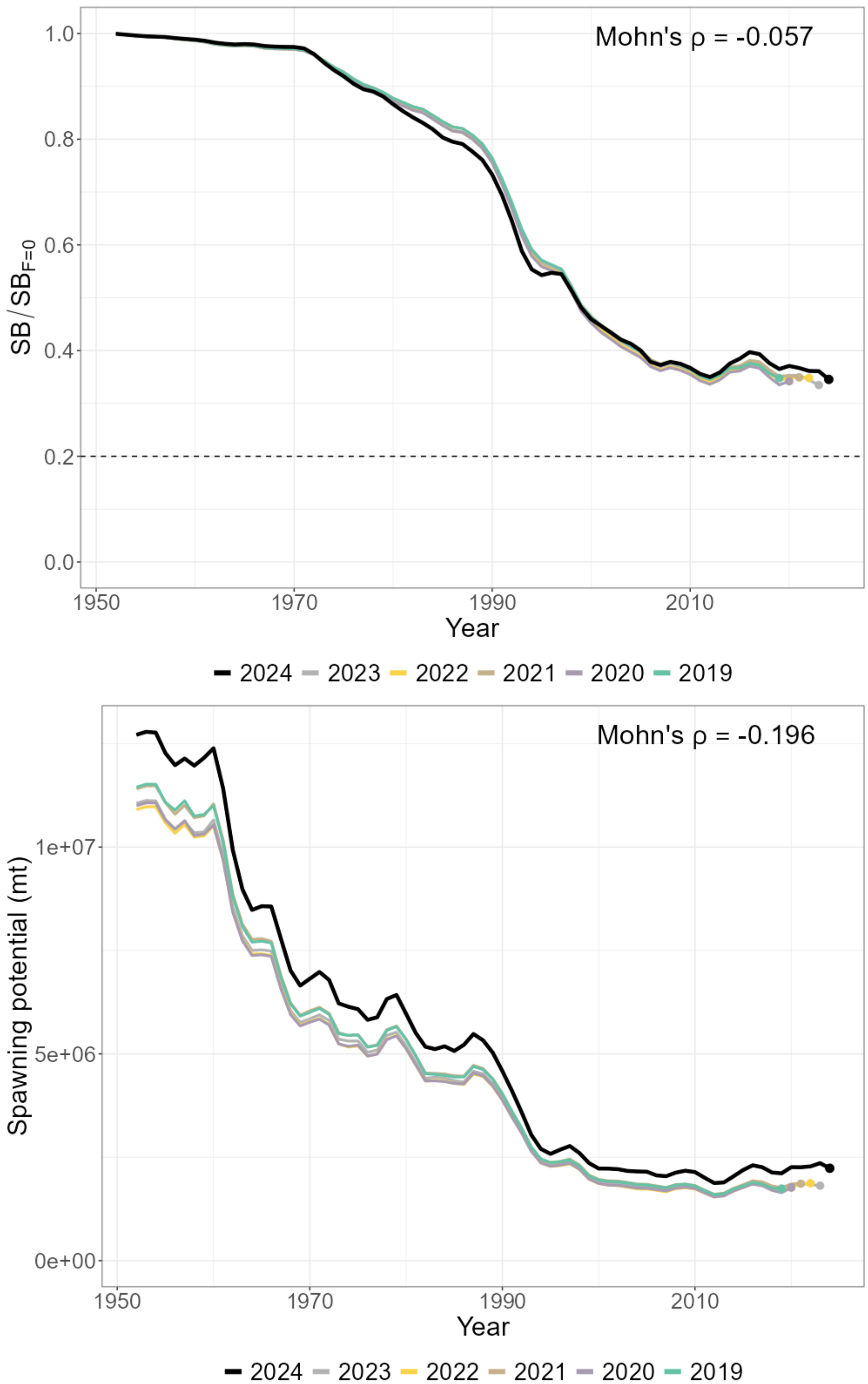

Figure 28: Retrospective estimates of (top) depletion and (bottom) spawning potential for the Diagnostic model. The black line is the full-data fit; the 5 converged retrospective peels are coloured by terminal year, with points marking terminal estimates. The dashed line in the depletion panel marks the limit reference point ($SB/SB_{F=0} = 0.2$).

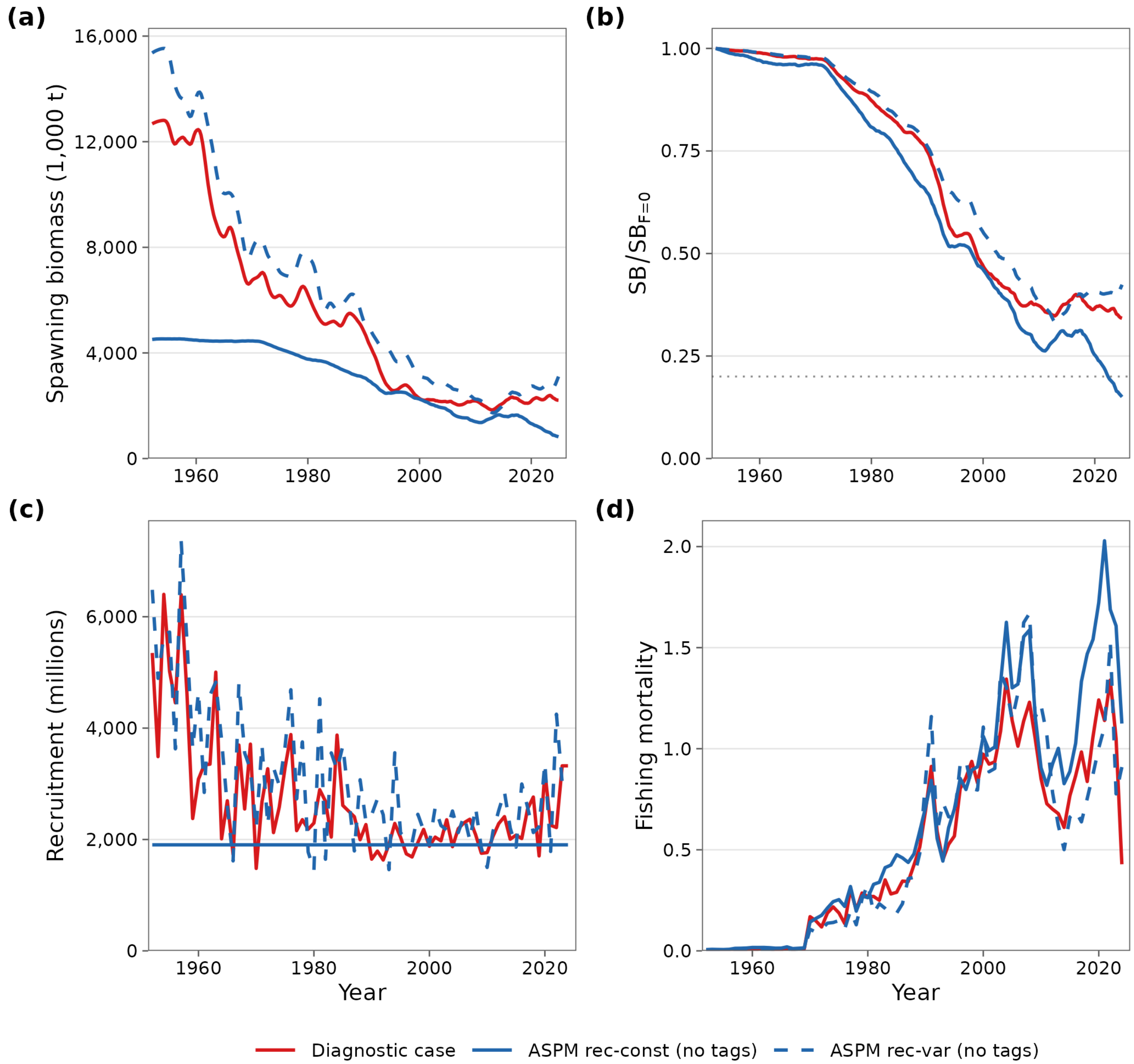

Figure 29: Results of the age structured production model (ASPM) diagnostic using average recruitment and recruitment deviations.

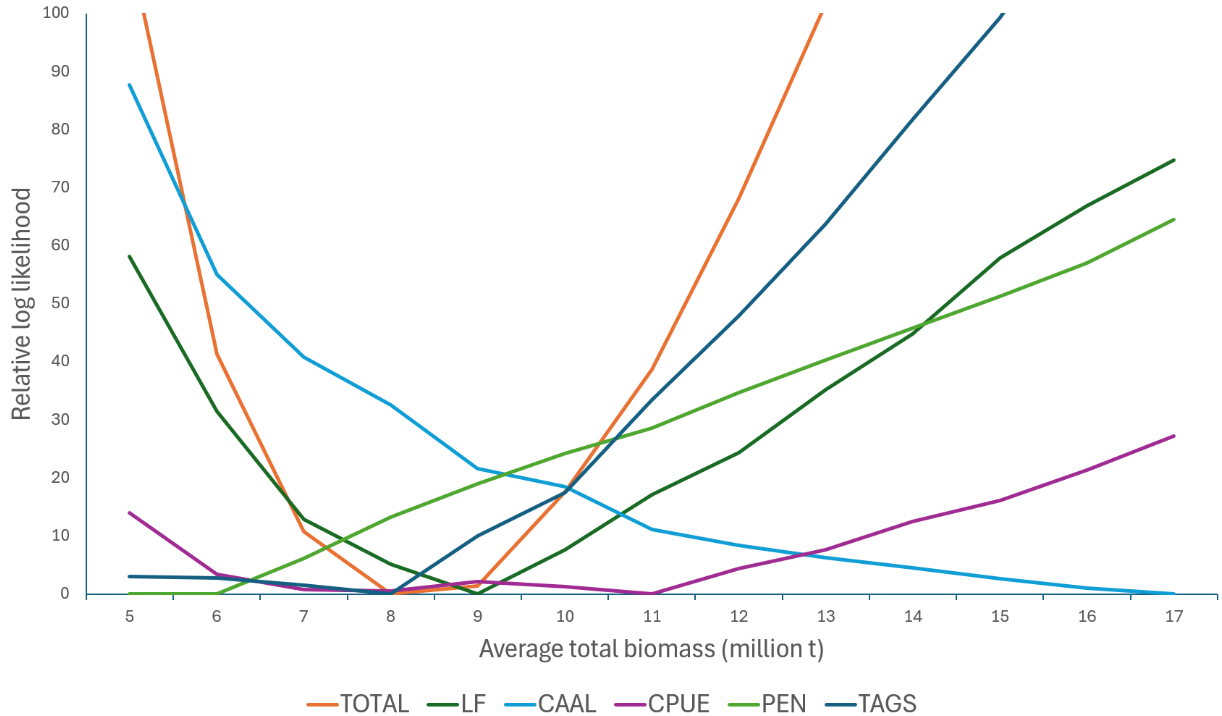

Figure 30: Diagnostic model likelihood profile of average total biomass in million mt. The orange line indicates the total likelihood with the other colours representing component of the total likelihood coming from different data sources as indicated on the figure.

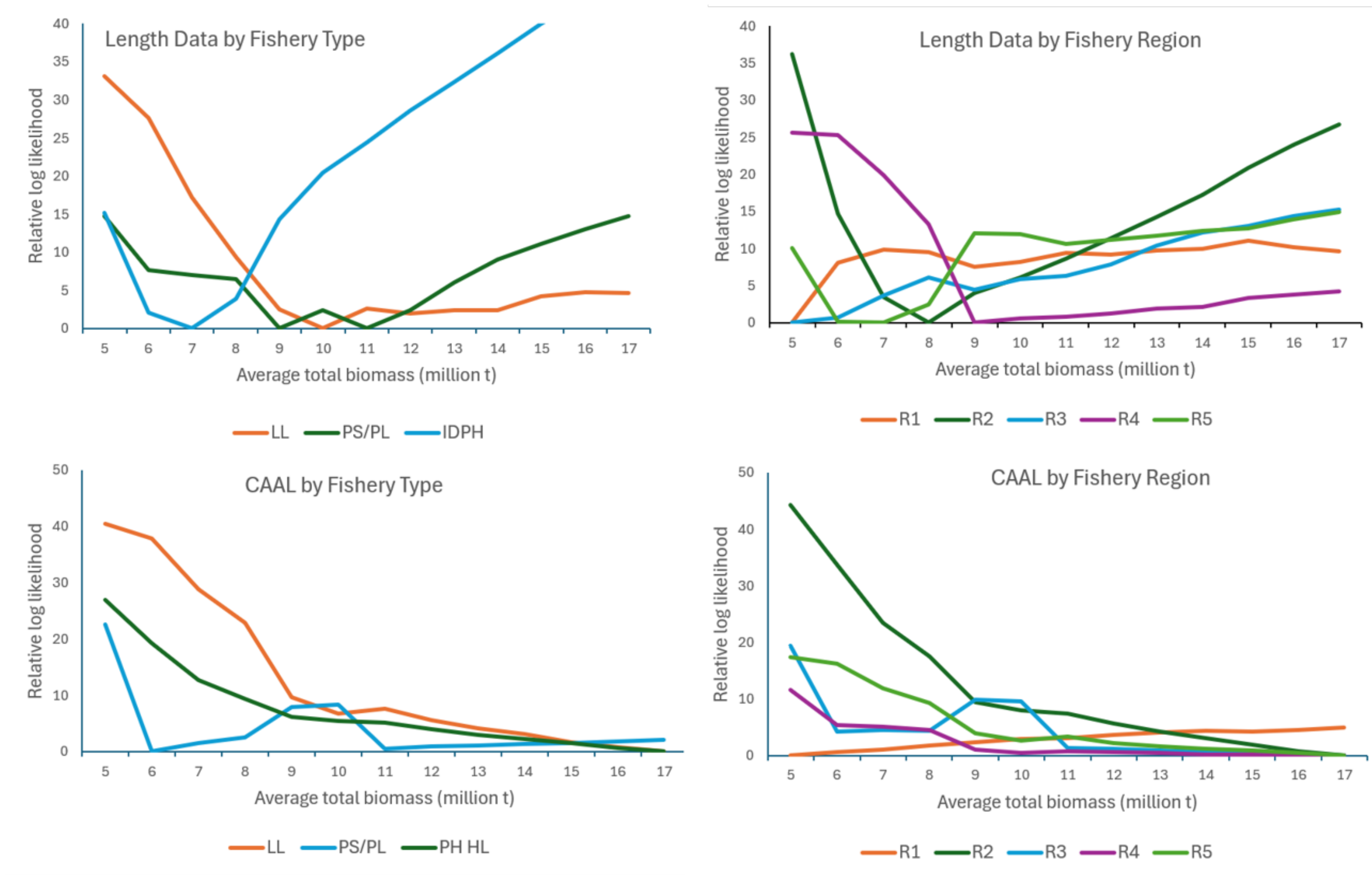

Figure 31: Diagnostic model likelihood profile of average total biomass in million mt. The orange line indicates the total likelihood with the colours representing component of the total likelihood coming from different fisheries groups or regions as per the label on each plot, for the length and CAAL data.

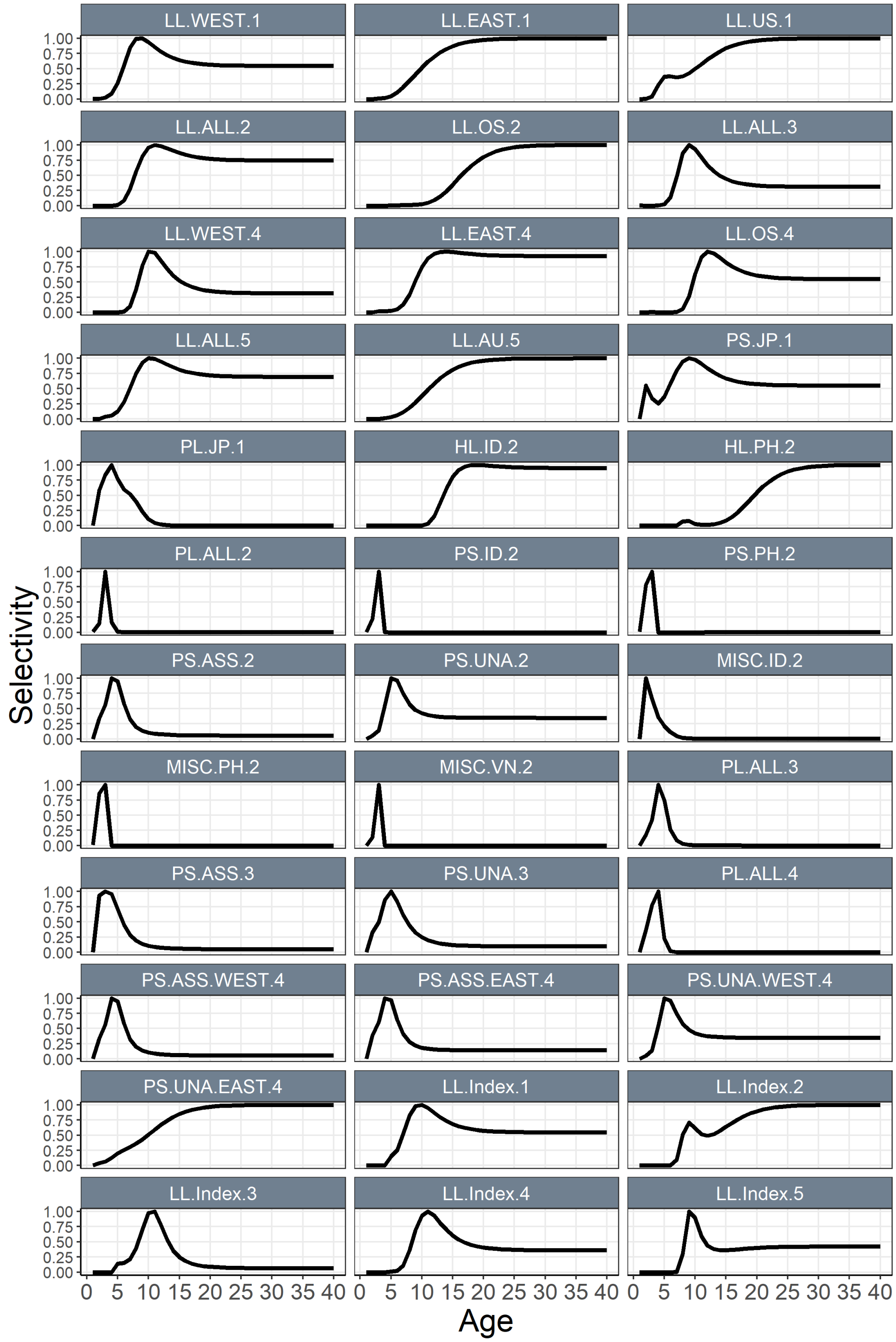

Figure 32: Age-specific (quarters) selectivity curves for all extraction and index fisheries.

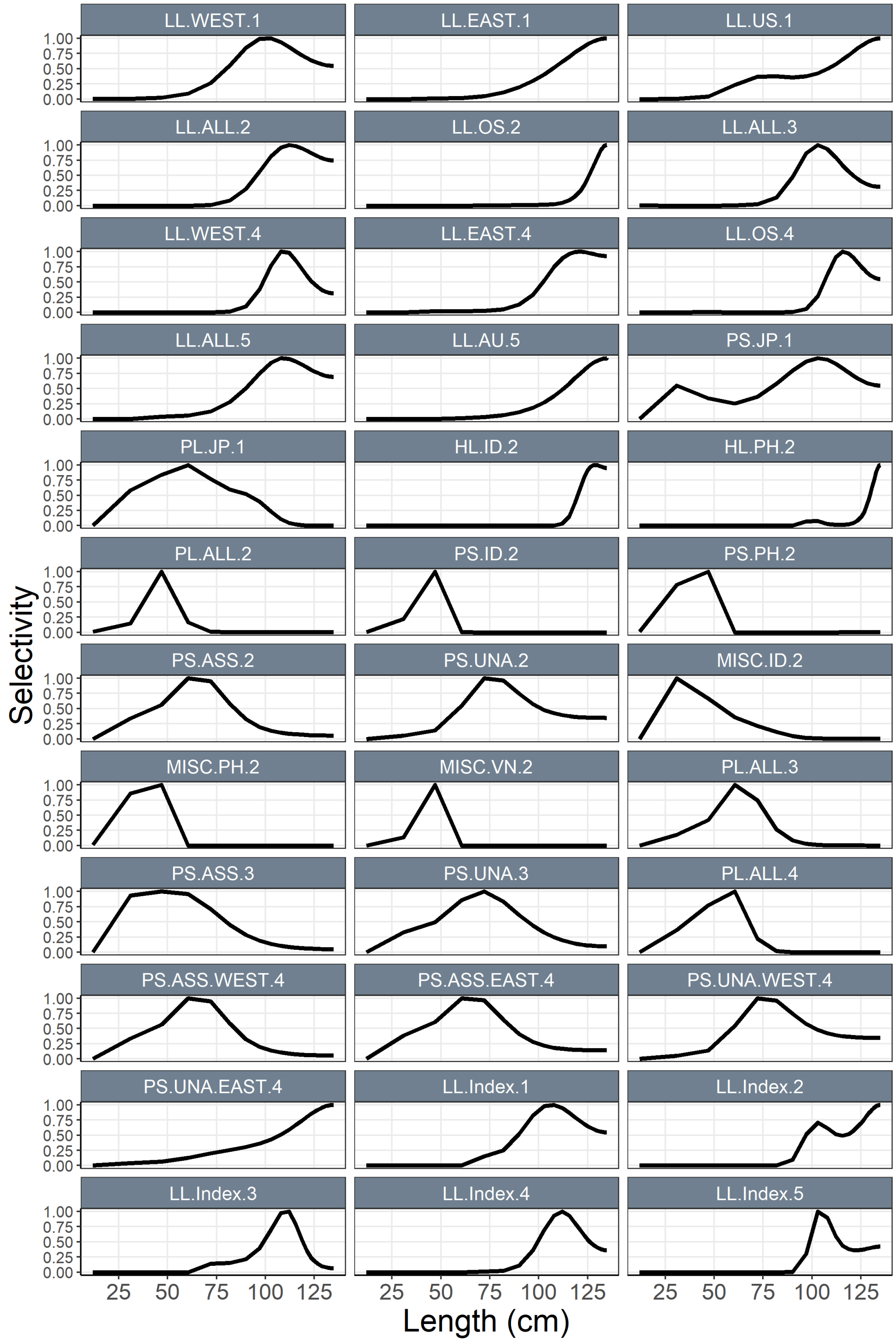

Figure 33: Length (cm) selectivity curves for all extraction and index fisheries.

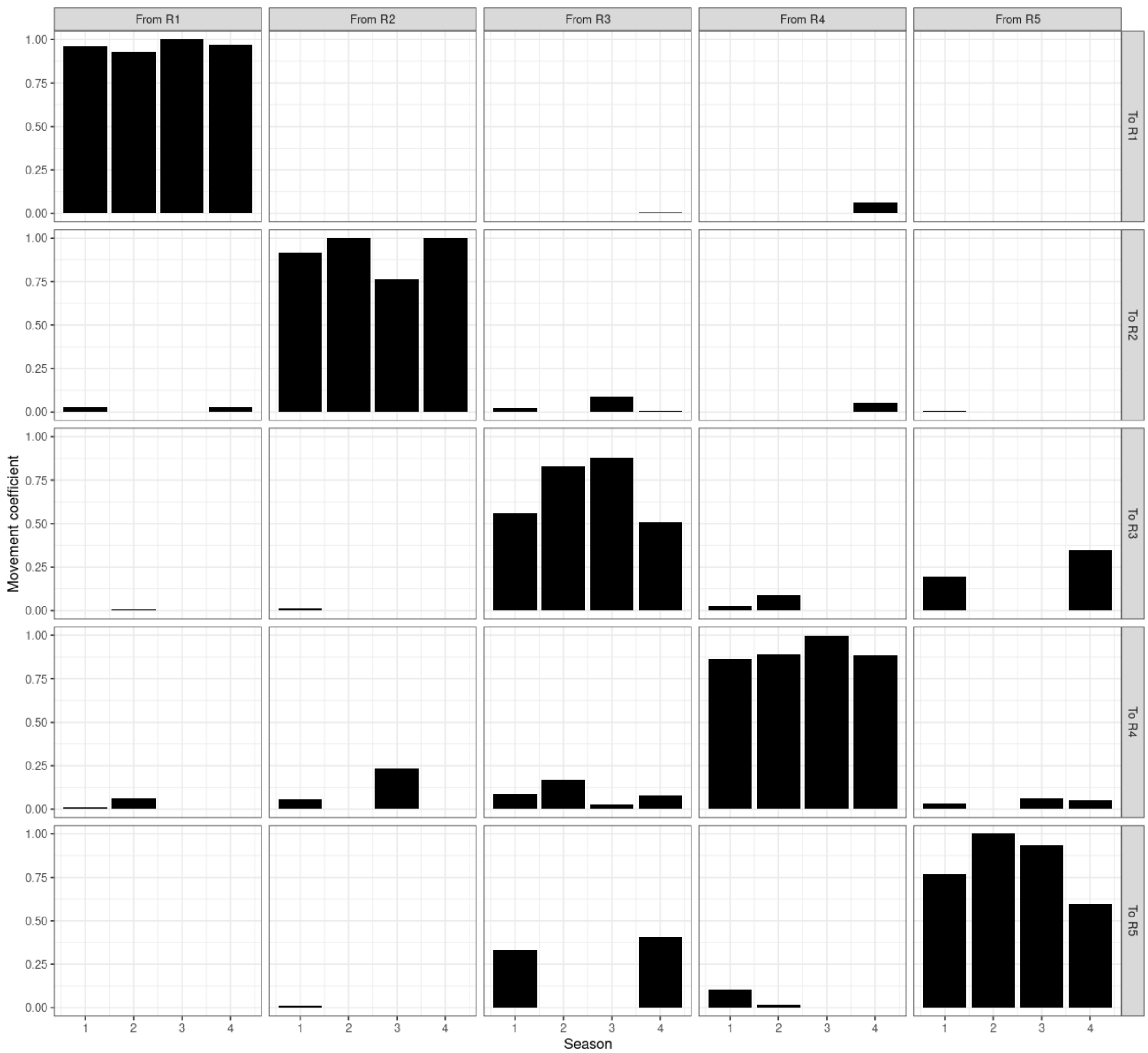

Figure 34: Season-specific movement probabilities among model regions estimated by the 2026 diagnostic model.

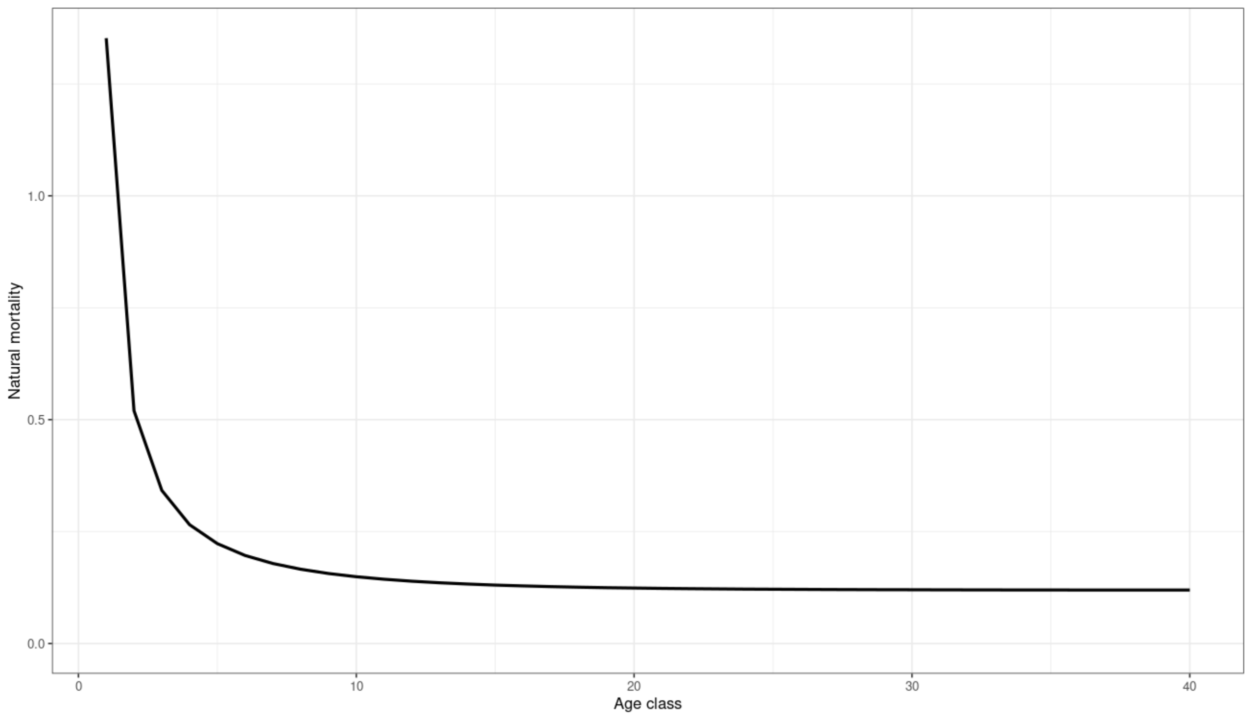

Figure 35: Lorenzen natural mortality-at-age (quarters) (top) and at-length (bottom) for the 2026 diagnostic model, based on M at age 40 quarters of 0.119.

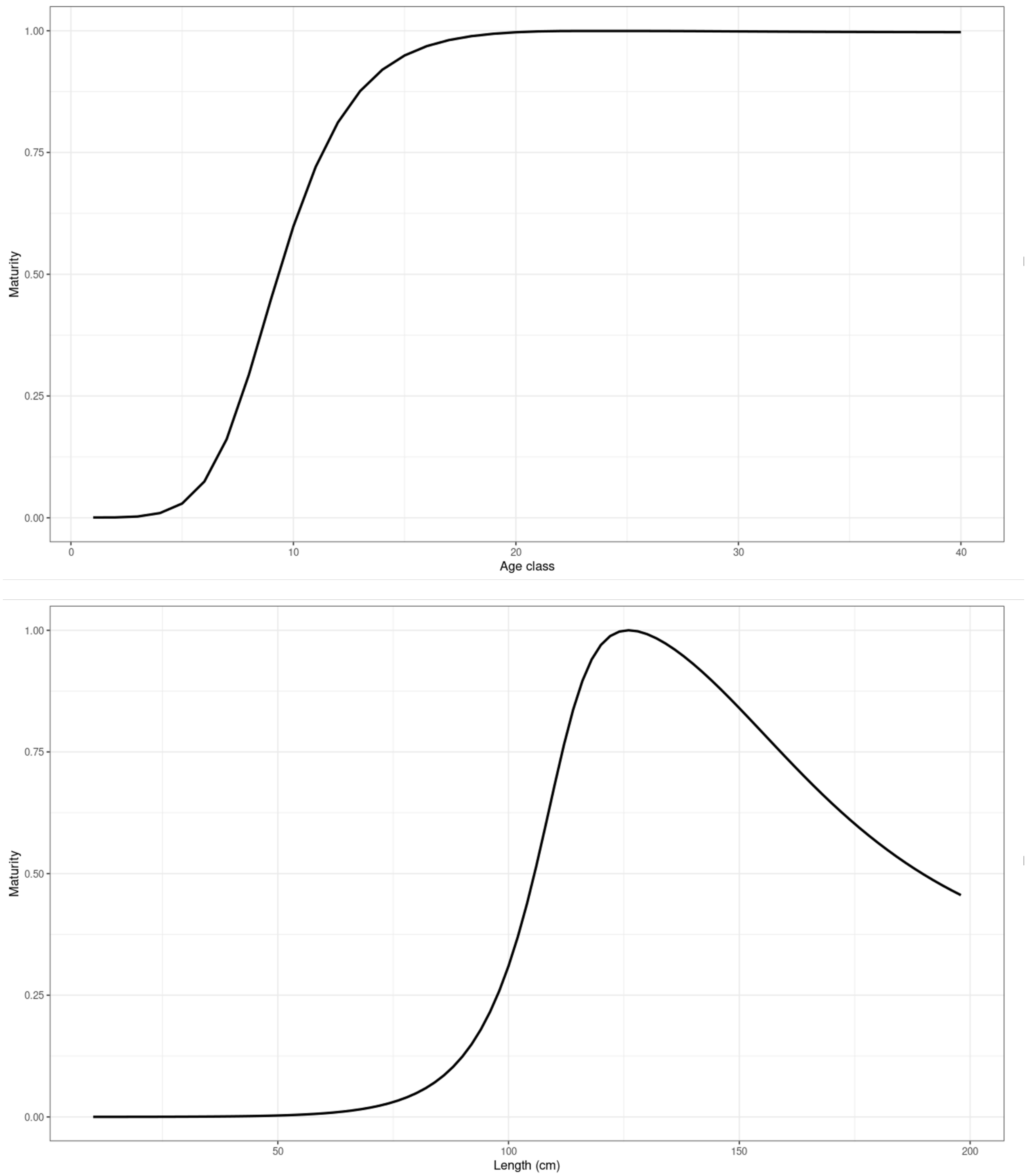

Figure 36: Maturity-at-age (top) and at length (bottom) ogives for the 2026 diagnostic model.

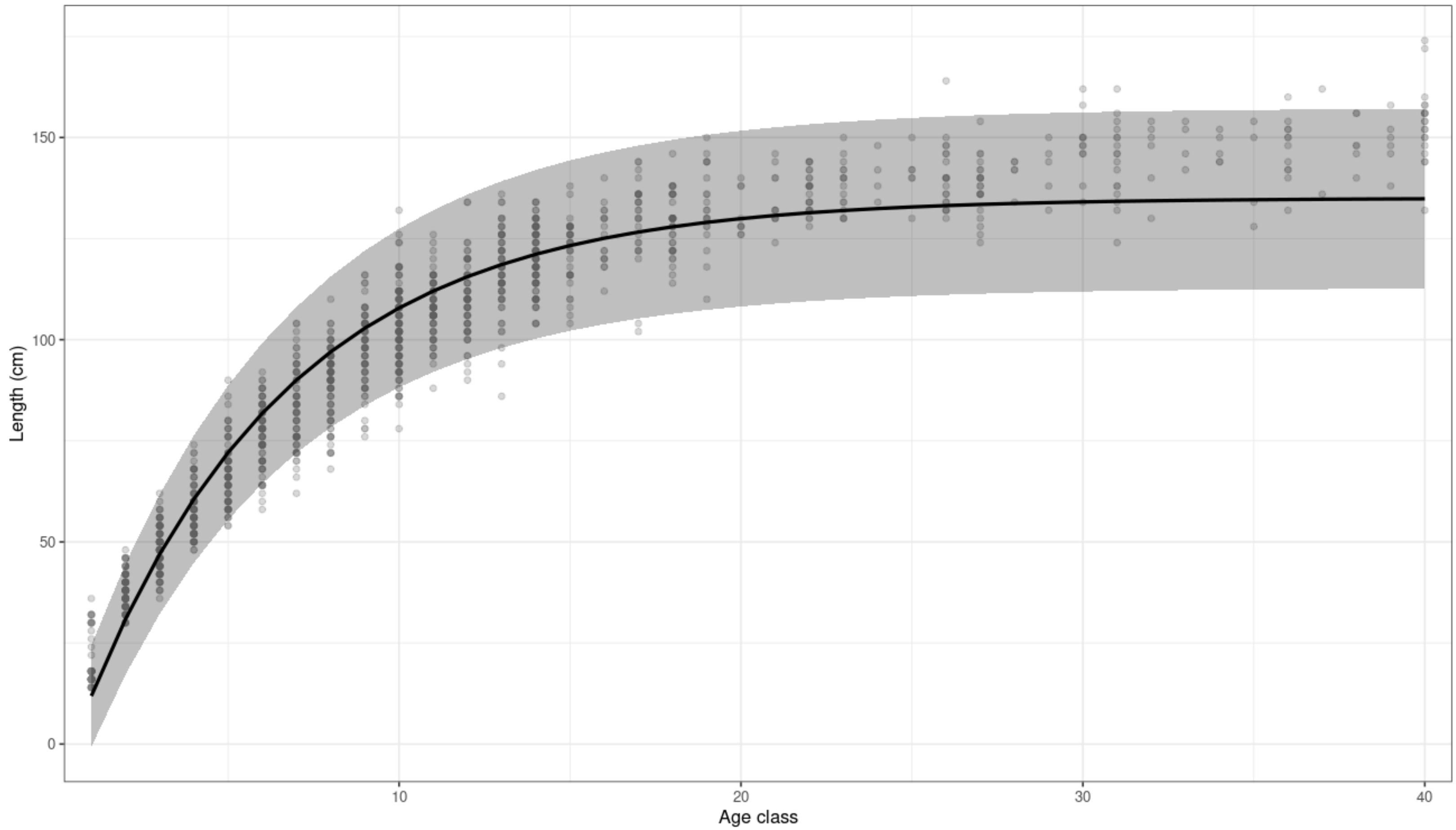

Figure 37: Estimated growth curve for the 2026 diagnostic model.

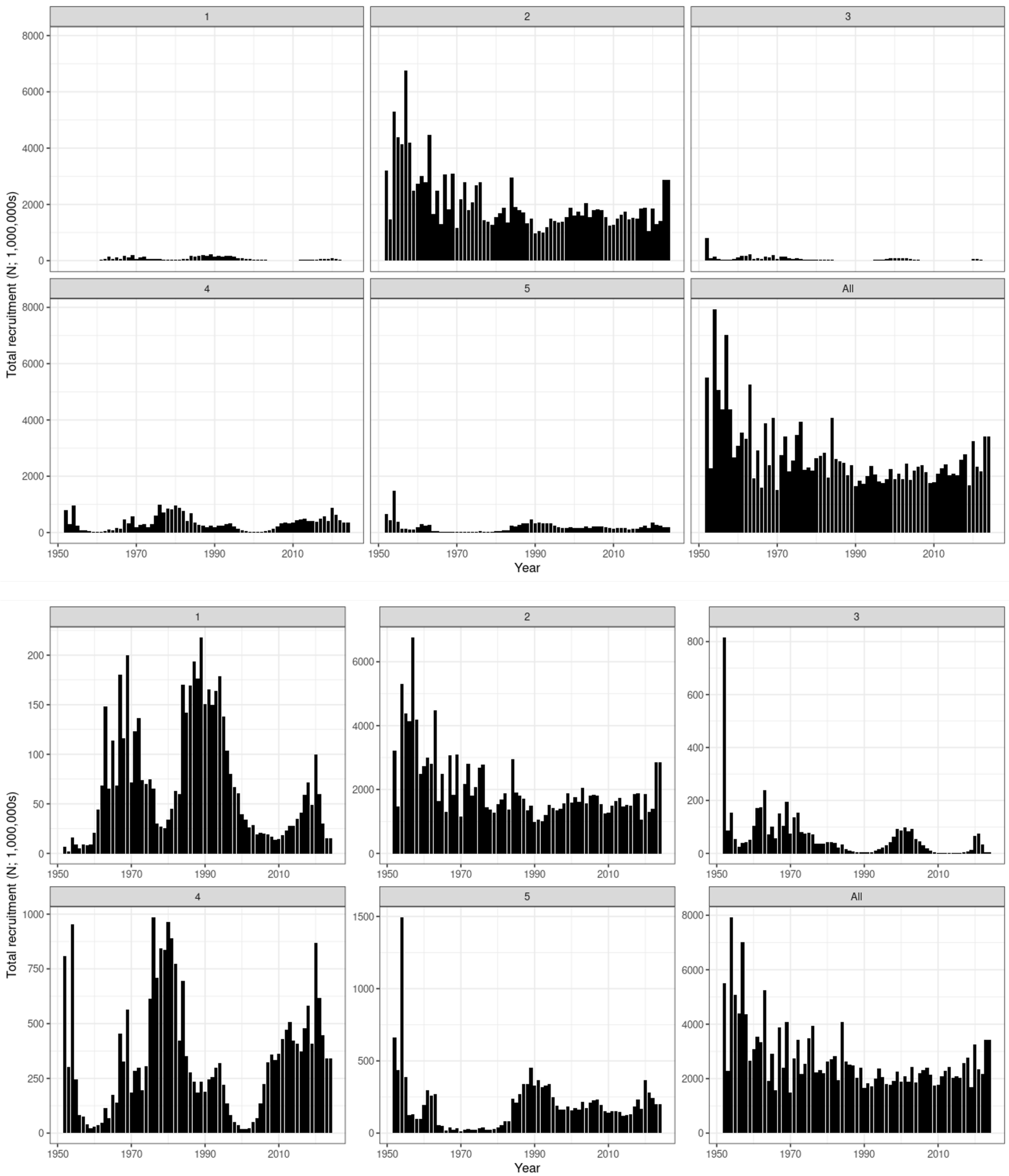

Figure 38: Time series of estimated total annual recruitment by regions and summed over all regions for the 2026 diagnostic model, top plots have same Y-axis scale, bottom plots have variable Y-axise scale.

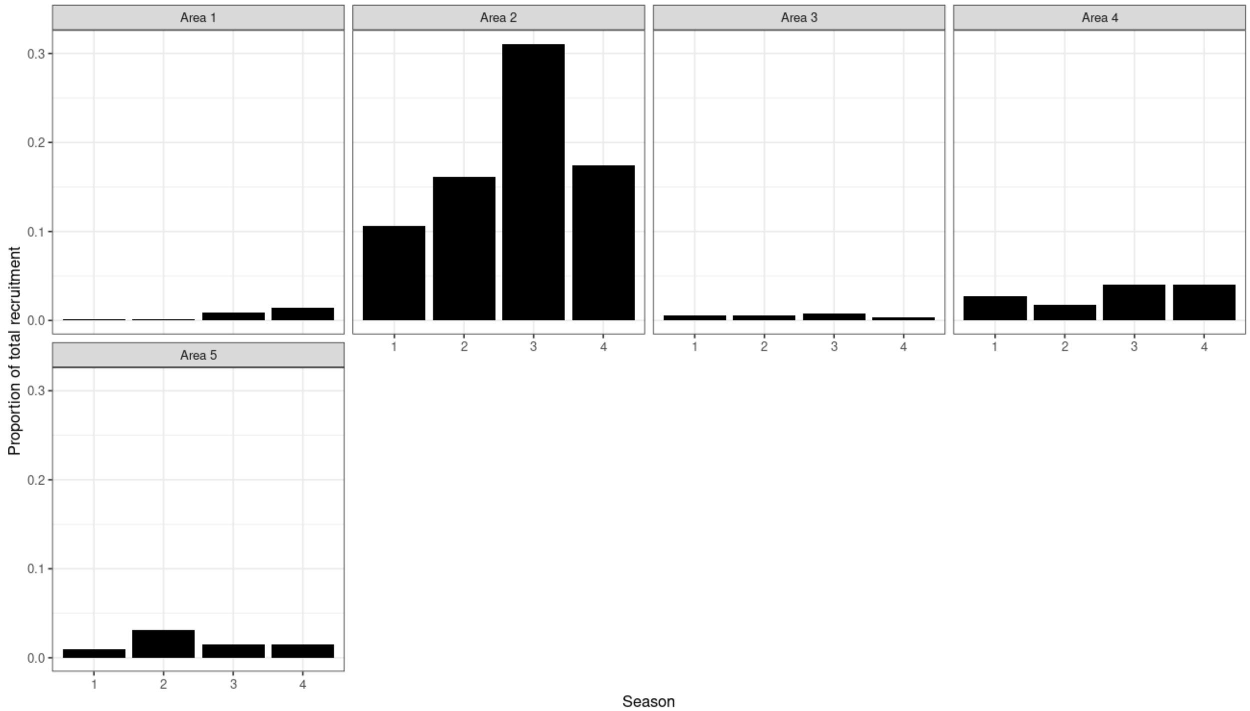

Figure 39: Estimated recruitment distribution by region and quarter.

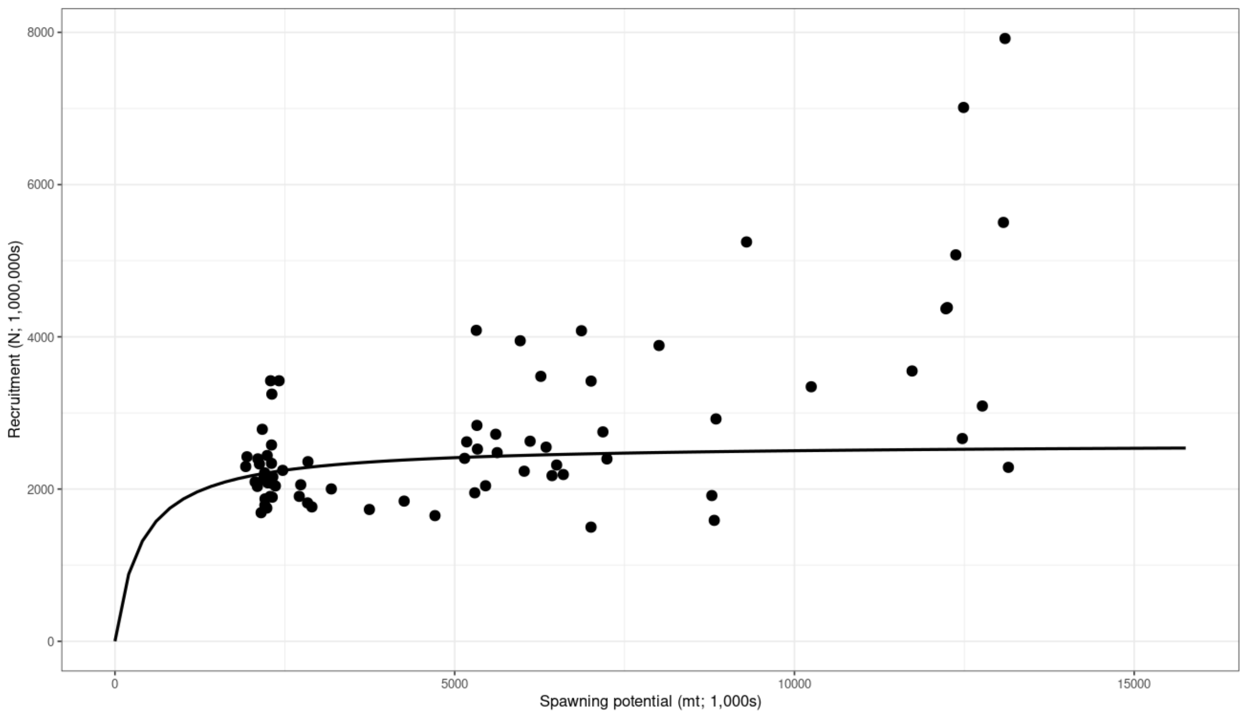

Figure 40: Estimated stock-recruitment relationship between recruitment and spawning potential based on annual estimates of recruitment for the 2026 diagnostic model. The colour of the circles transitions from purple (early in the time series) to yellow (more recent) through time. Points shown prior to 1968 are not used in the fit to the stock recruitment relationship.

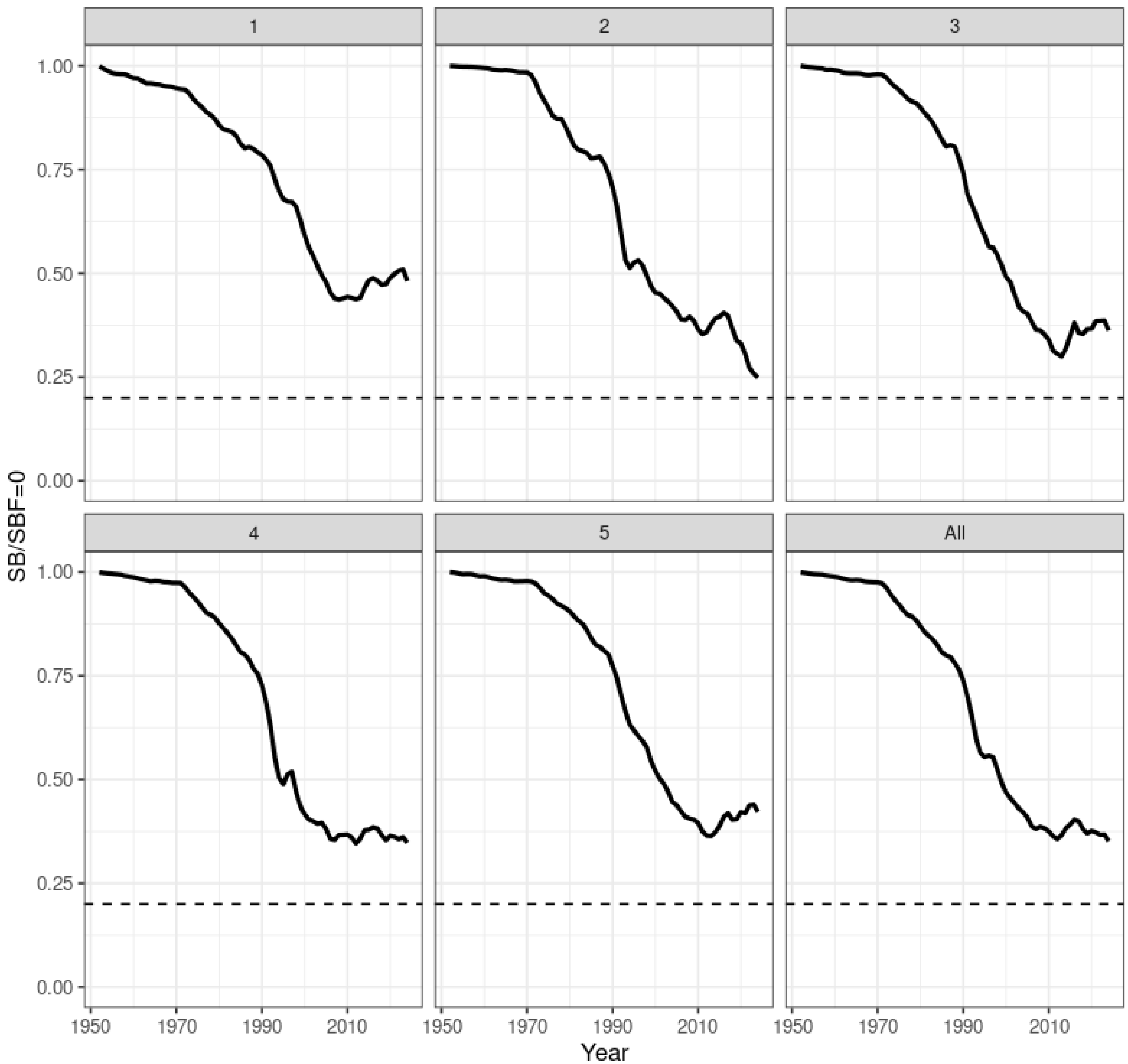

Figure 41: Population dynamics summary for depletion ($SB/SB_{F=0}$) for the diagnostic model for each region and combined all regions.

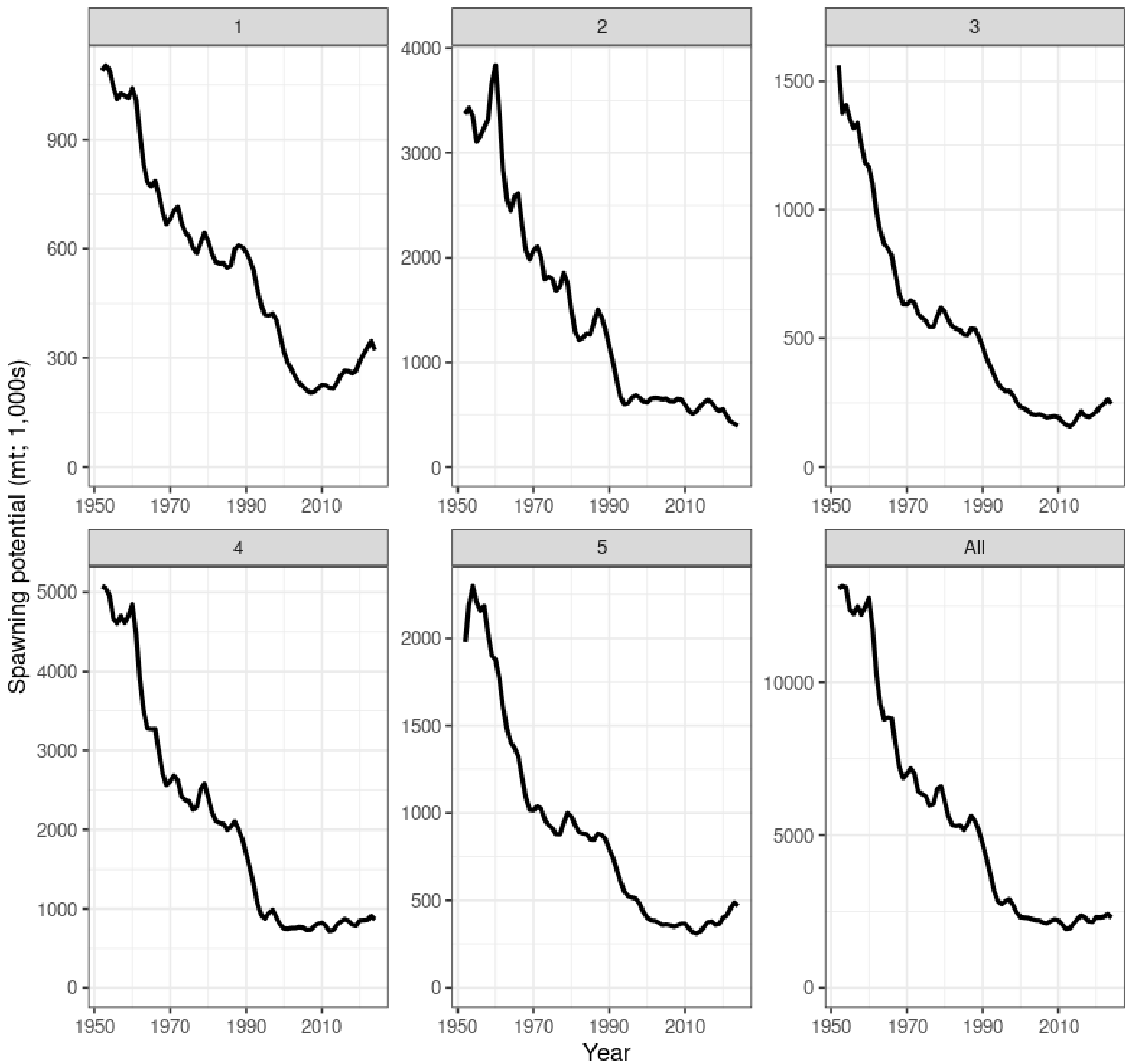

Figure 42: Population dynamics summary for spawning potential (SB) for the diagnostic model for each region and combined all regions.

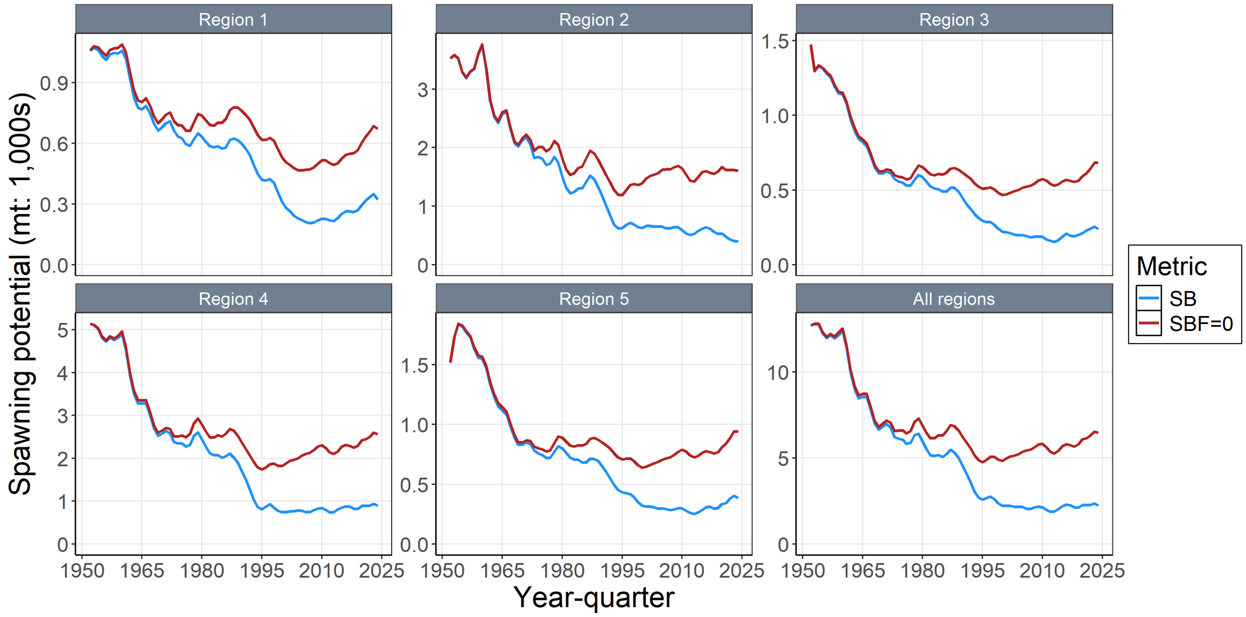

Figure 43: Comparison of the estimated annual spawning potential trajectories (lower, blue lines) with the spawning potential trajectories predicted in the absence of fishing (upper, red lines) for each region and overall, for the 2026 diagnostic model. Note the scales of the y-axis varies for each region.

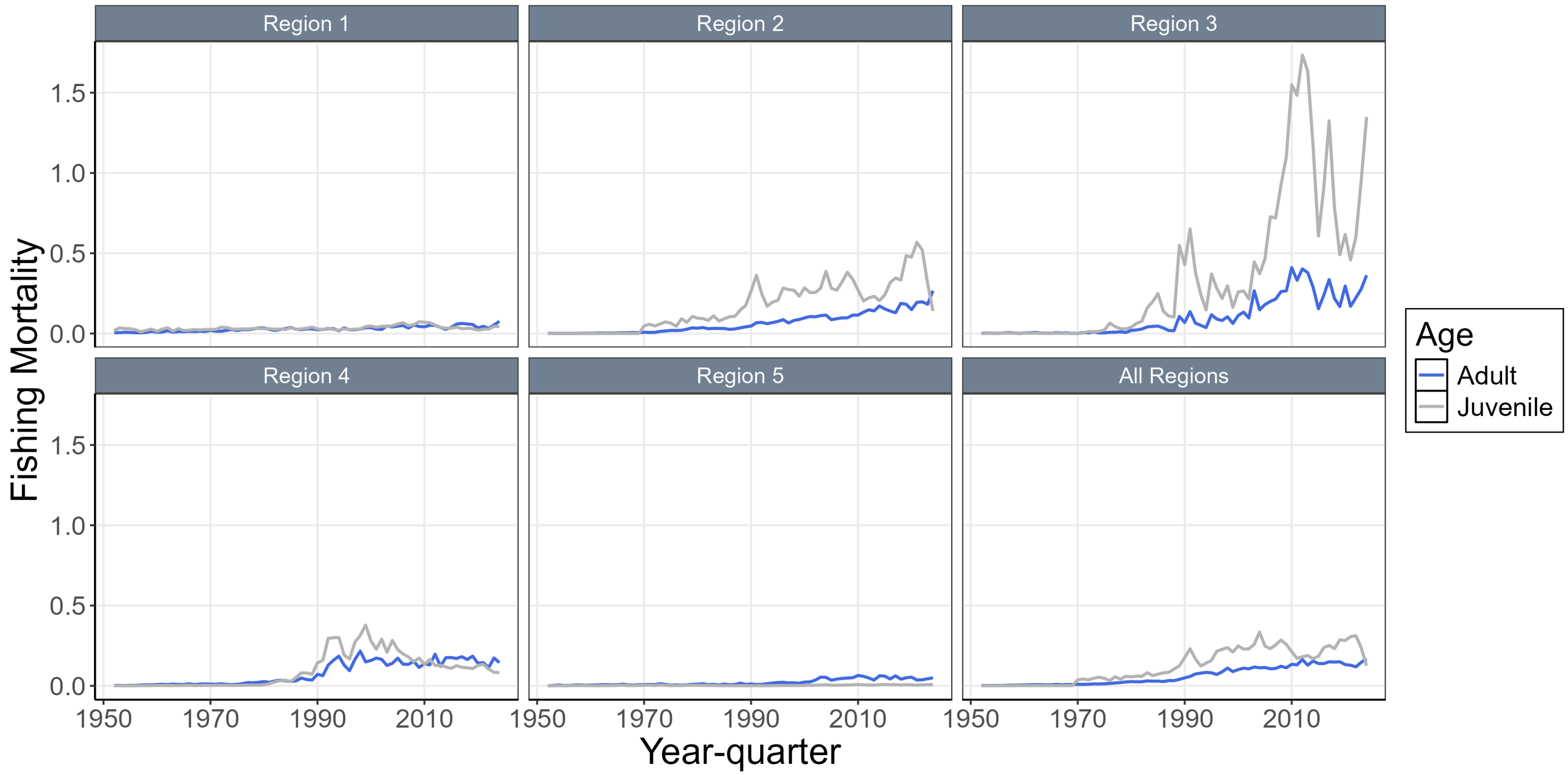

Figure 44: Estimated annual average adult (blue line) and juvenile (gray line) fishing mortality for the 2026 diagnostic model by model region and across all regions.

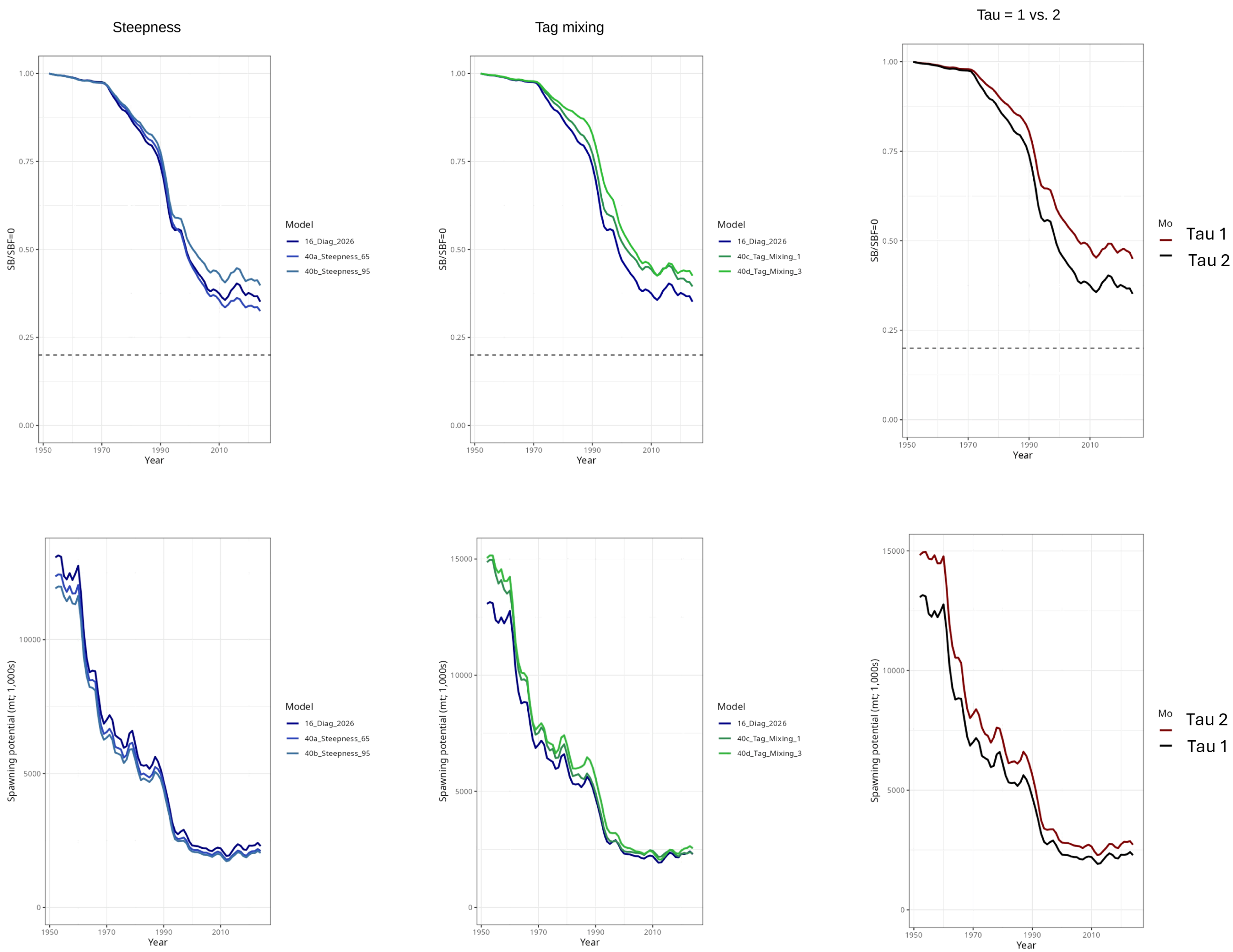

Figure 45: Estimated dynamic spawning depletion $SB/SB_{F=0}$ (Top) and spawning potential (SB) (Bottom), for the one-off sensitivities from the 2026 diagnostic model for steepness (left), tag mixing (middle) and tau (right) .

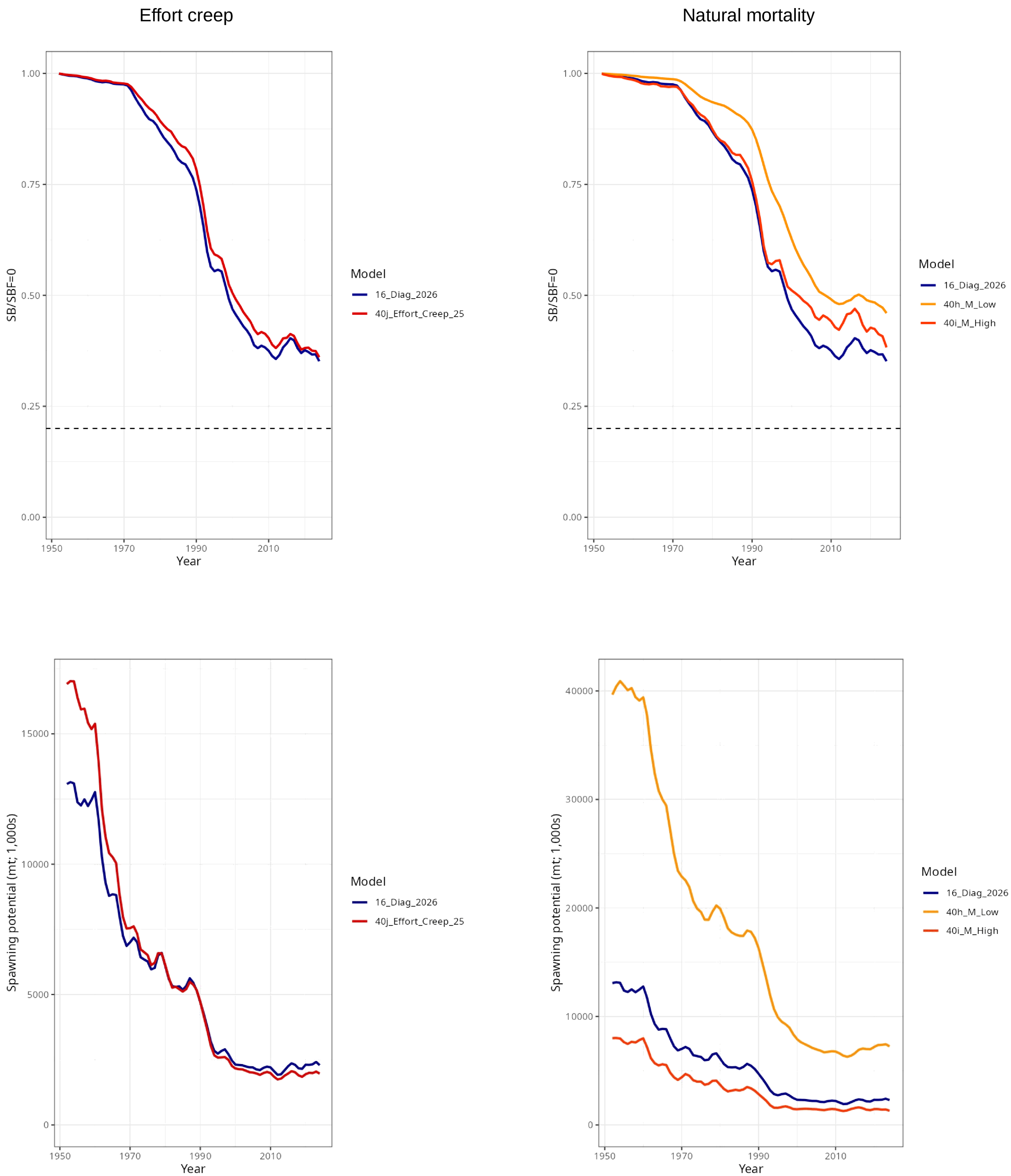

Figure 46: Estimated dynamic spawning depletion $SB/SB_{F=0}$ (Top) and spawning potential (SB) (Bottom), for the one-off sensitivities from the 2026 diagnostic model for effort creep (left), and natural mortality (right).

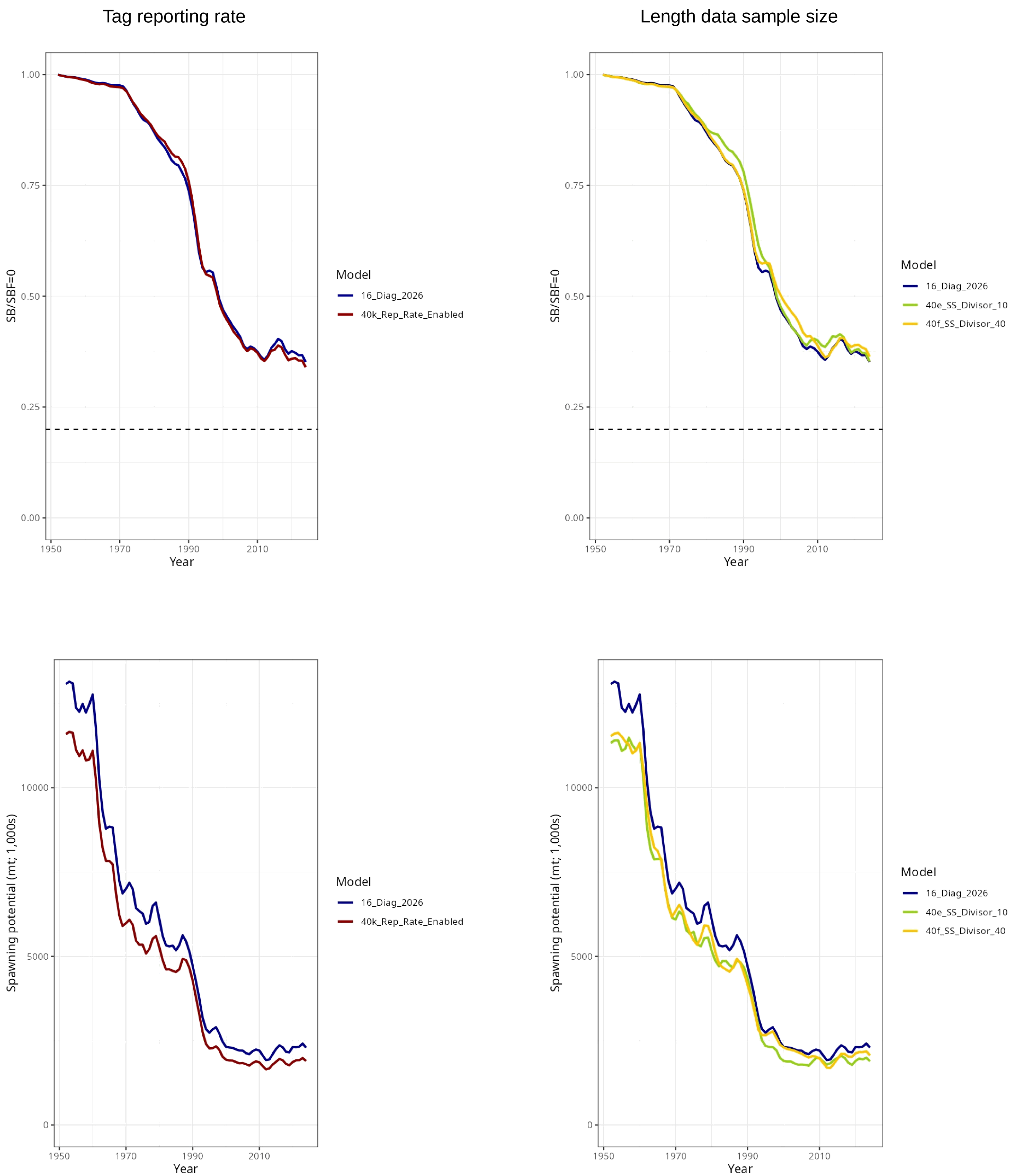

Figure 47: Estimated dynamic spawning depletion $SB/SB_{F=0}$ (Top) and spawning potential (SB) (Bottom), for the one-off sensitivities from the 2026 diagnostic model for activation or deactivation (i.e. diag 2026 model) of tag reporting rate adjustments in mixing period (left), and length samples data weighting (right).

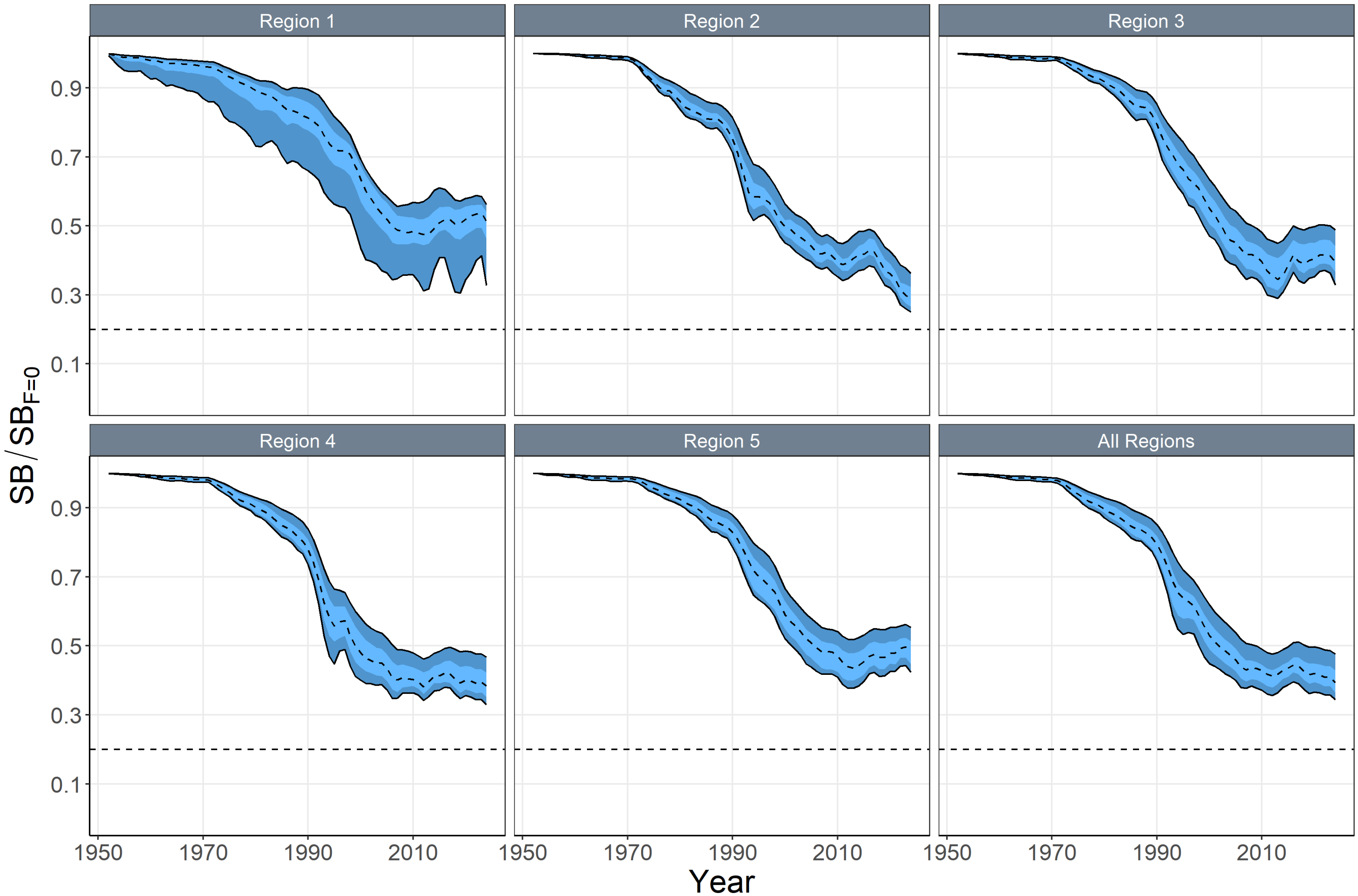

Figure 48: Ribbon plots of estimated dynamic spawning depletion, $SB_t/SB_{t,F=0}$, for the 48 models retained for the uncertainty ensemble including the estimation uncertainty. 90% (dark blue) and 75% (light blue) quantiles, black dashed line is median.

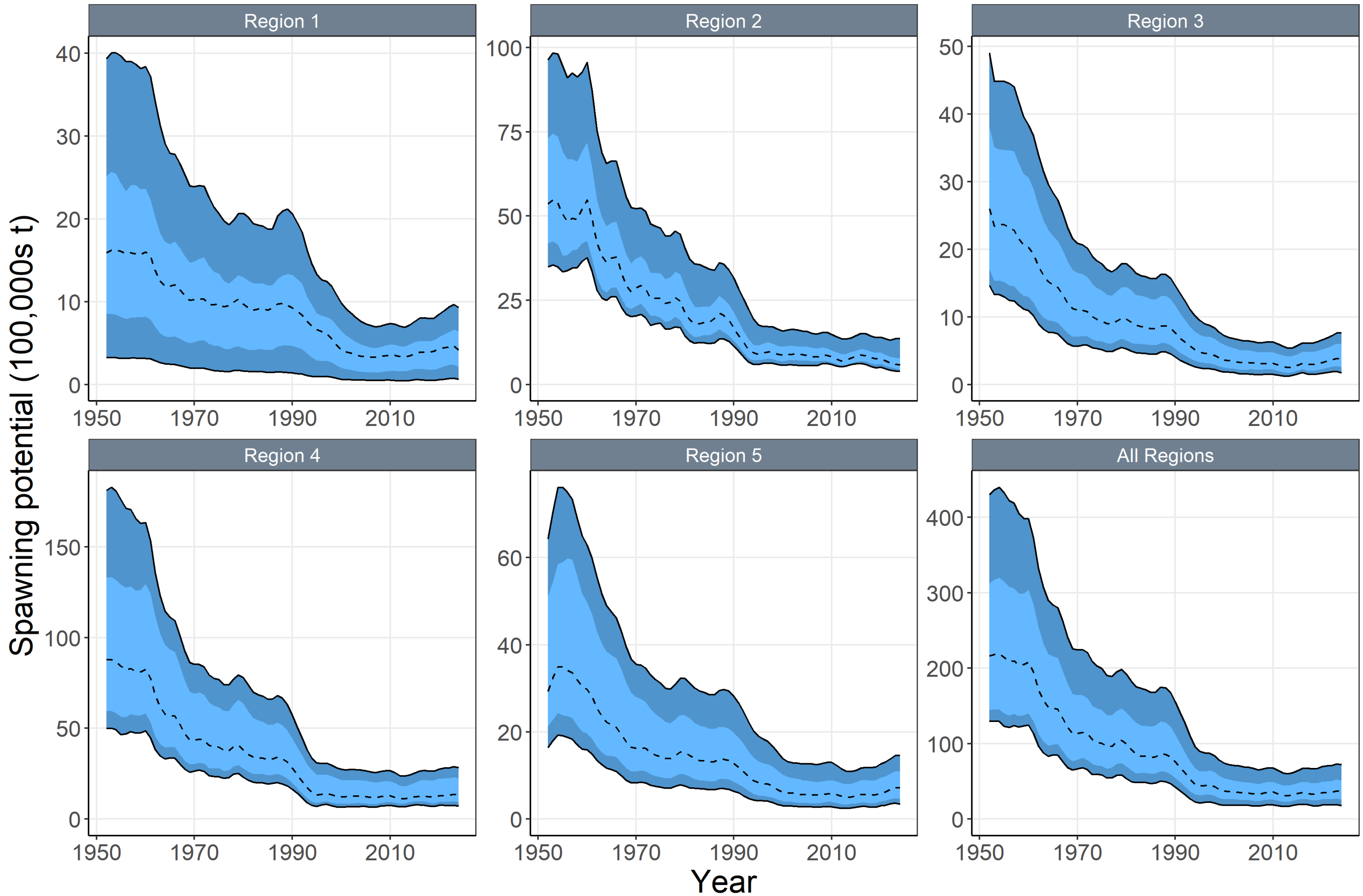

Figure 49: Ribbon plots of estimated dynamic spawning depletion, *SB*, for the 48 models retained for the uncertainty ensemble including the estimation uncertainty. 90% (dark blue) and 75% (light blue) quantiles, black dashed line is median.

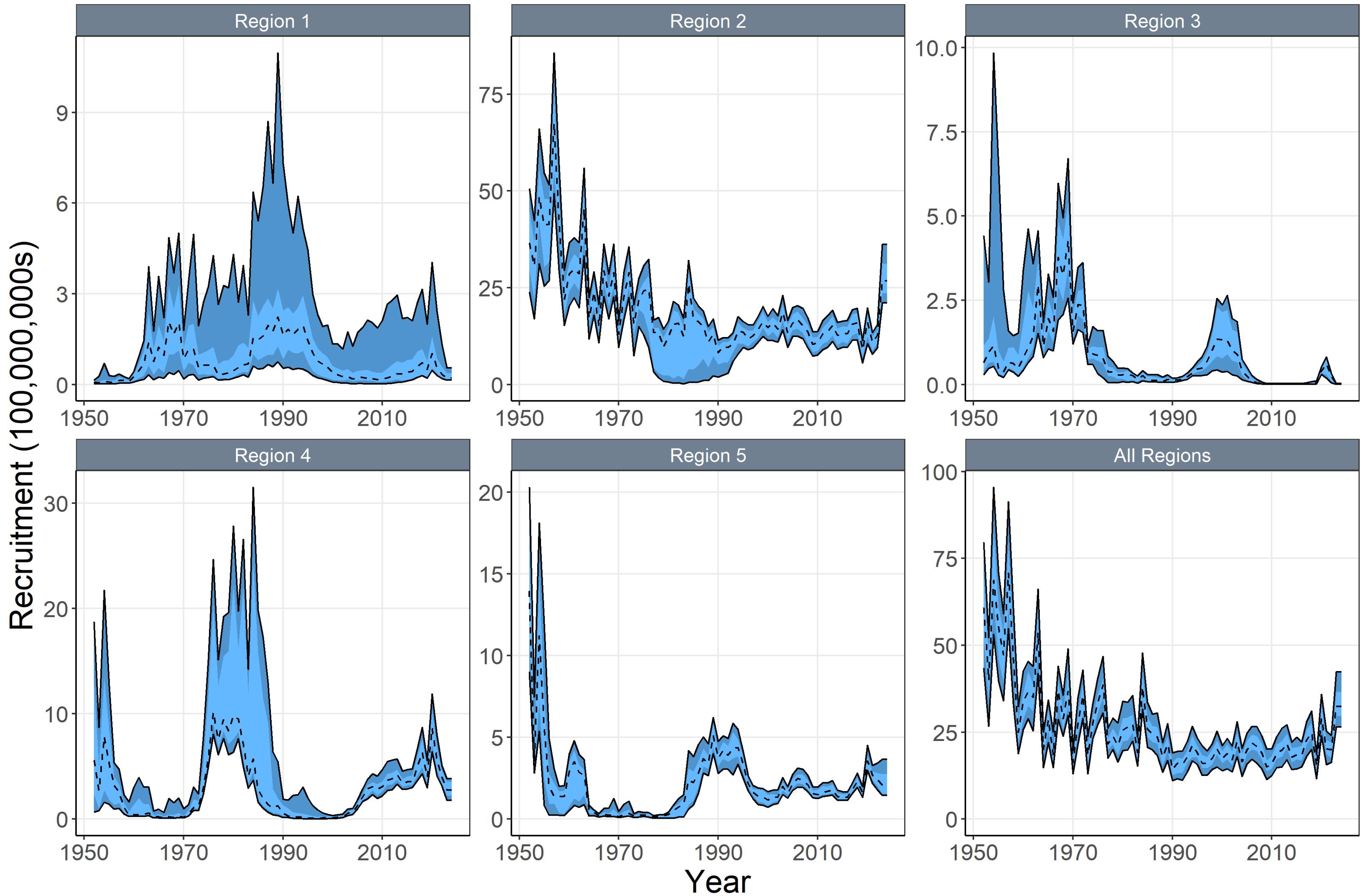

Figure 50: Ribbon plots of estimated dynamic recruitment for the 48 models retained for the uncertainty ensemble including the estimation uncertainty. 90% (dark blue) and 75% (light blue) quantiles, black dashed line is median.

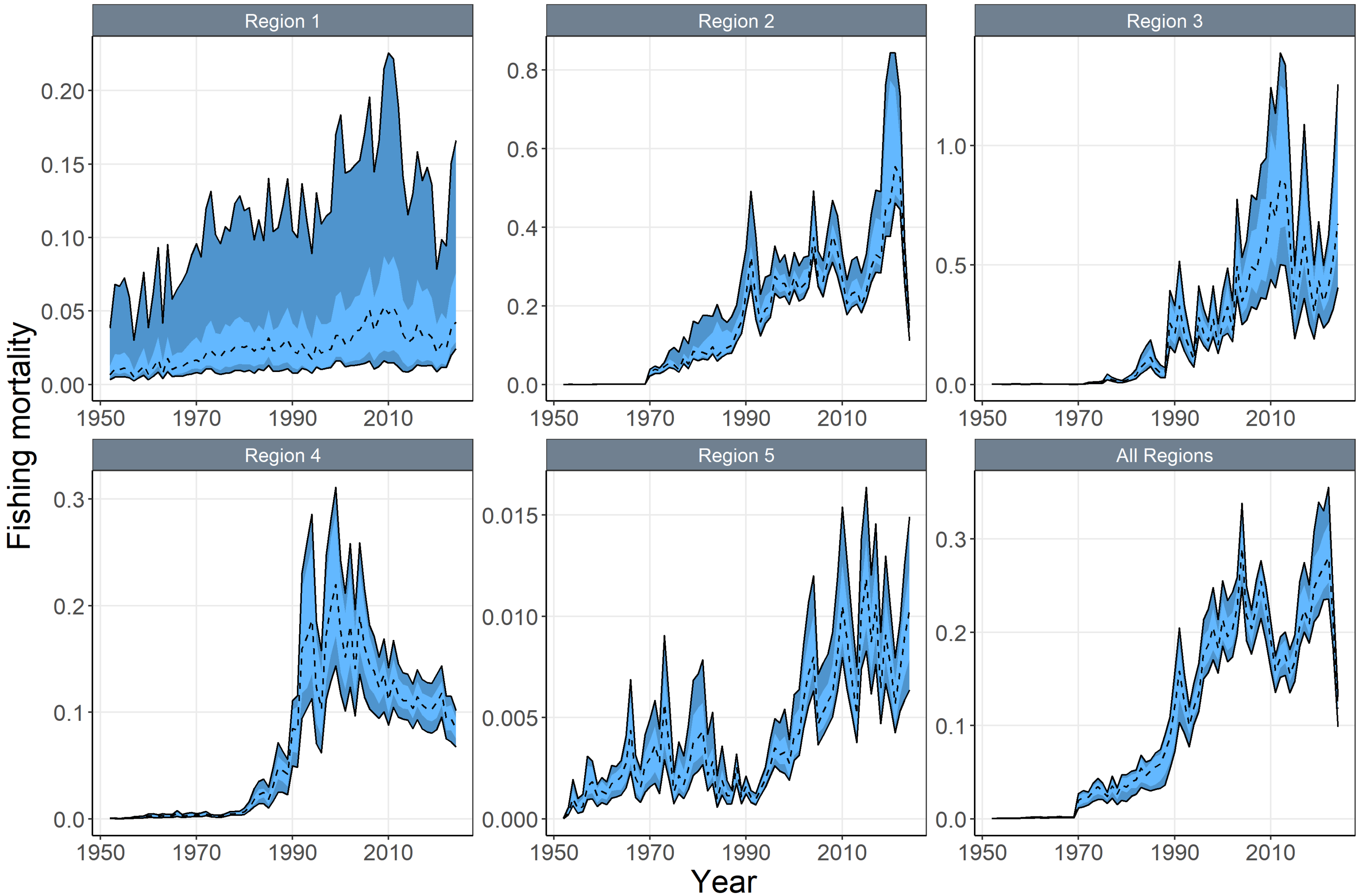

Figure 51: Ribbon plots of estimated dynamic fishing mortality, F, for the 48 models retained for the uncertainty ensemble including the estimation uncertainty. 90% (dark blue) and 75% (light blue) quantiles, black dashed line is median.

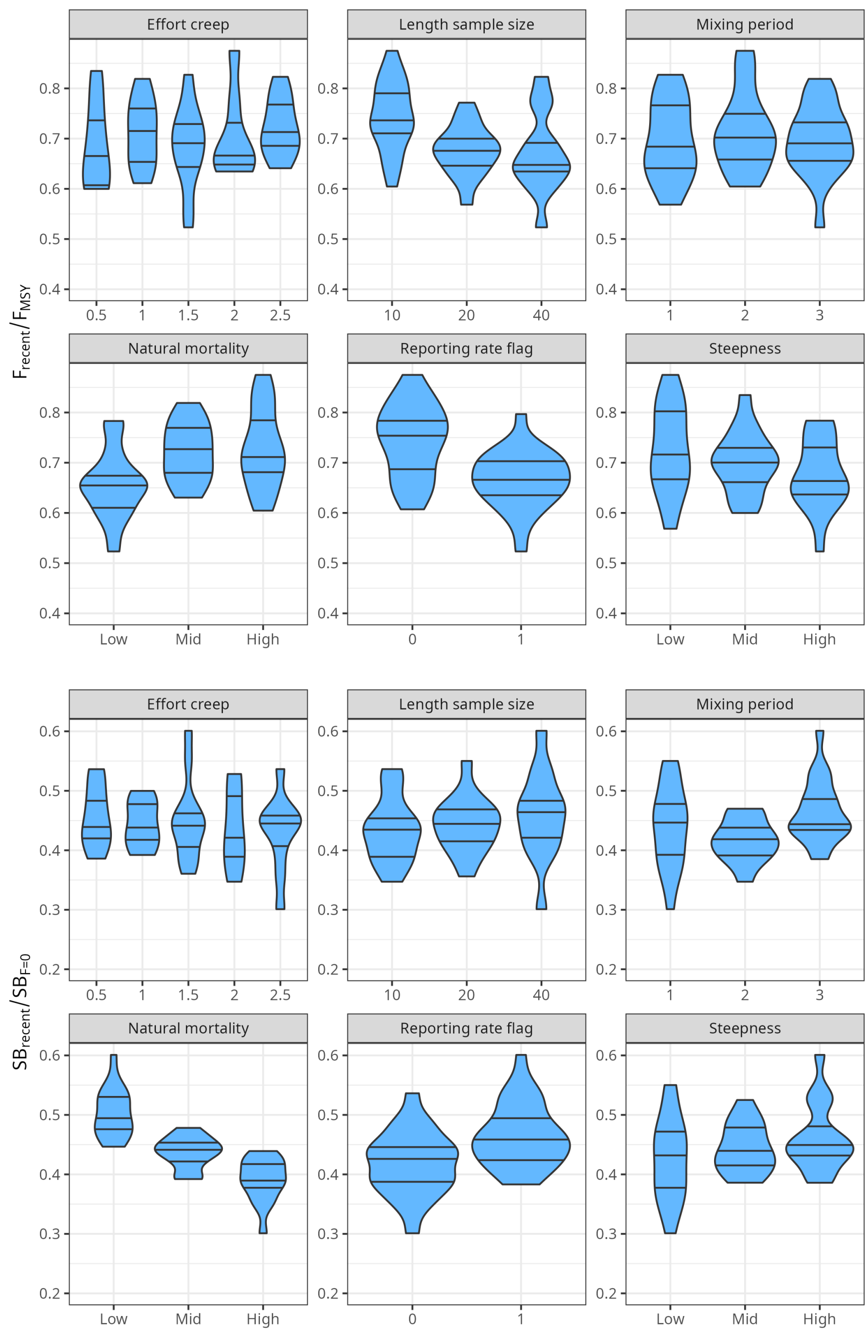

Figure 52: Box and violin plots summarizing the estimated $F_{\text{recent}}/F_{\text{MSY}}$ (top) and $SB_{\text{recent}}/SB_{F=0}$ (bottom) for each of the 48 models retained in the uncertainty ensemble by uncertainty axes: effort creep, length sample size (i.e. size data weighting divisors), tag mixing period, reporting rate flag (i.e. reporting rate adjustments in mixing period, deactivated = 1, activated = 0), steepness (h) and natural mortality (fixed M at age 40 quaters). For natural mortality and steepness (that use a beta distribution) we used the 1/3 and 2/3 quantiles, that produces 16 models in high, mid, and low categories, these quantiles being: M low is < 0.098 and high is > 0.117, mid is in between, and for steepness low is <0.807 and high is > 0.859, mid is in between.

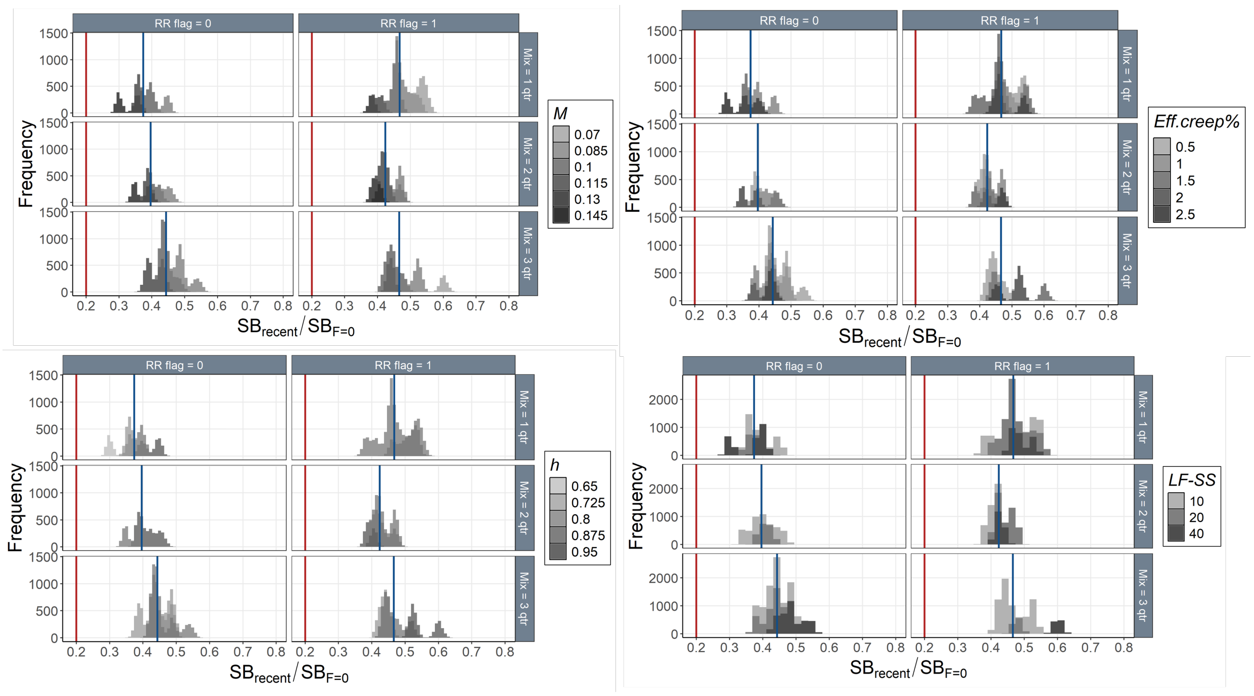

Figure 53: Histograms of Monte-Carlo estimated model uncertainty for $SB_{\mathrm{recent}}/SB_{F=0}$ by mixing period (1, 2, 3 quarter) and tag reporting rate activation (0) or deactivation (1) with separate plots for natural mortality (top left), steepness (bottom left), effort creep (top right), length data weighting divisor (10, 20, 40) (bottom right).

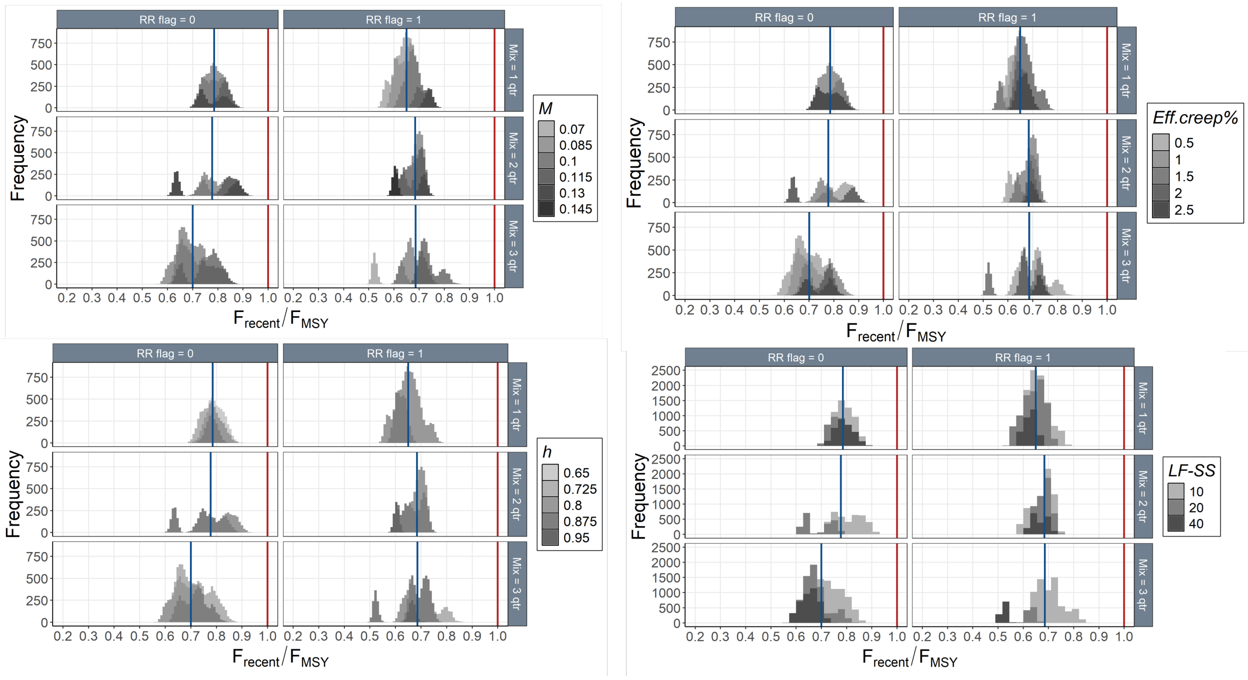

Figure 54: Histograms of Monte-Carlo estimated model uncertainty for $F_{\mathrm{recent}}/F_{\mathrm{MSY}}$ by mixing period (1, 2, 3 quarter) and tag reporting rate activation (0) or deactivation (1) with separate plots for natural mortality (top left), steepness (bottom left), effort creep (top right), length data weighting divisor (10, 20, 40) (bottom right).

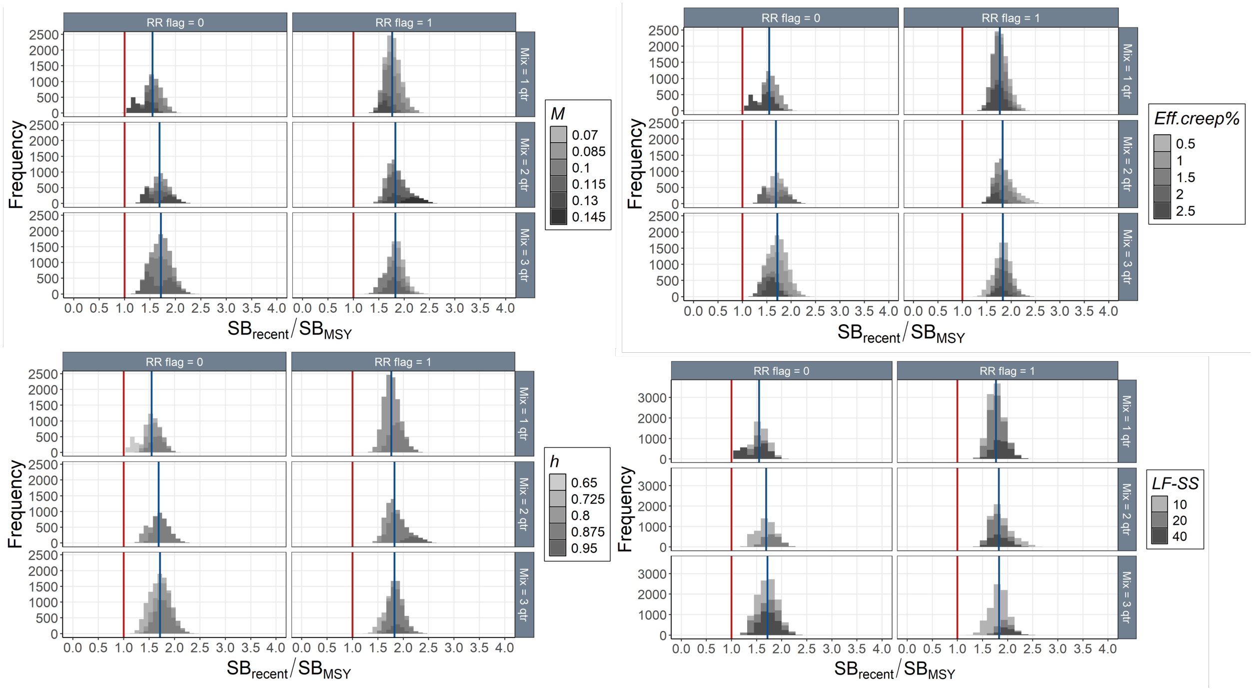

Figure 55: Histograms of Monte-Carlo estimated model uncertainty for $SB_{\text{recent}}/SB_{\text{MSY}}$ by mixing period (1, 2, 3 quarter) and tag reporting rate activation (0) or deactivation (1) with separate plots for natural mortality (top left), steepness (bottom left), effort creep (top right), length data weighting divisor (10, 20, 40) (bottom right).

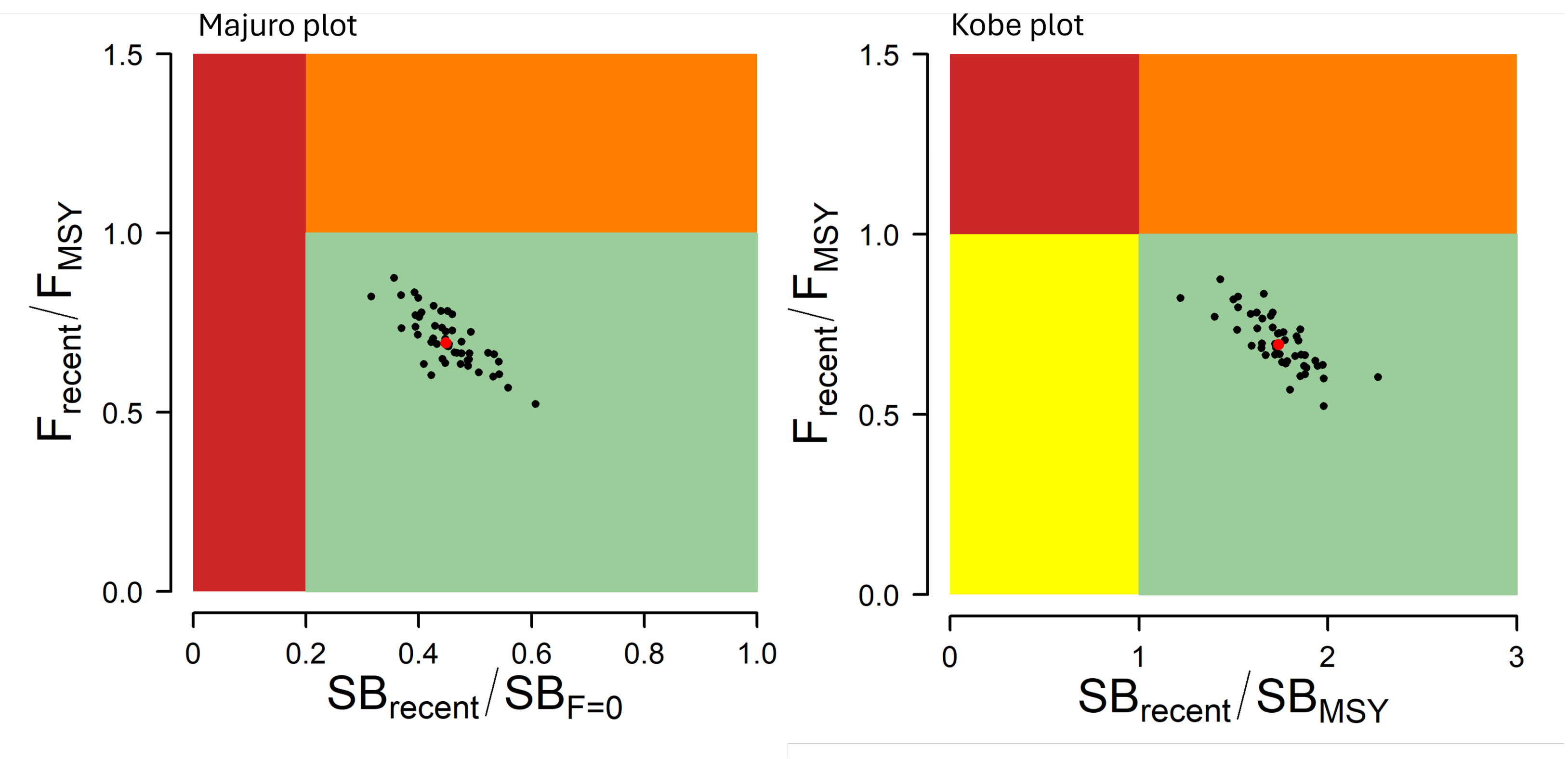

Figure 56: Stock status based on $SB_{\text{recent}}/SB_{F=0}$ (SBrecent 2021–2024) and $F_{\text{recent}}/F_{\text{MSY}}$ (Frecent 2020–2023). Panel (a) Majuro: $SB_{\text{recent}}/SB_{F=0}$ and $F_{\text{recent}}/F_{\text{MSY}}$ with the LRP of $SB_{\text{recent}}/SB_{F=0} = 0.20$. Panel (a) Kobe: is relative to $SB_{\text{recent}}/SB_{\text{MSY}}$ and $F_{\text{recent}}/F_{\text{MSY}}$ panel Points are the the 48 retained ensemble model estimate, red points are the medians

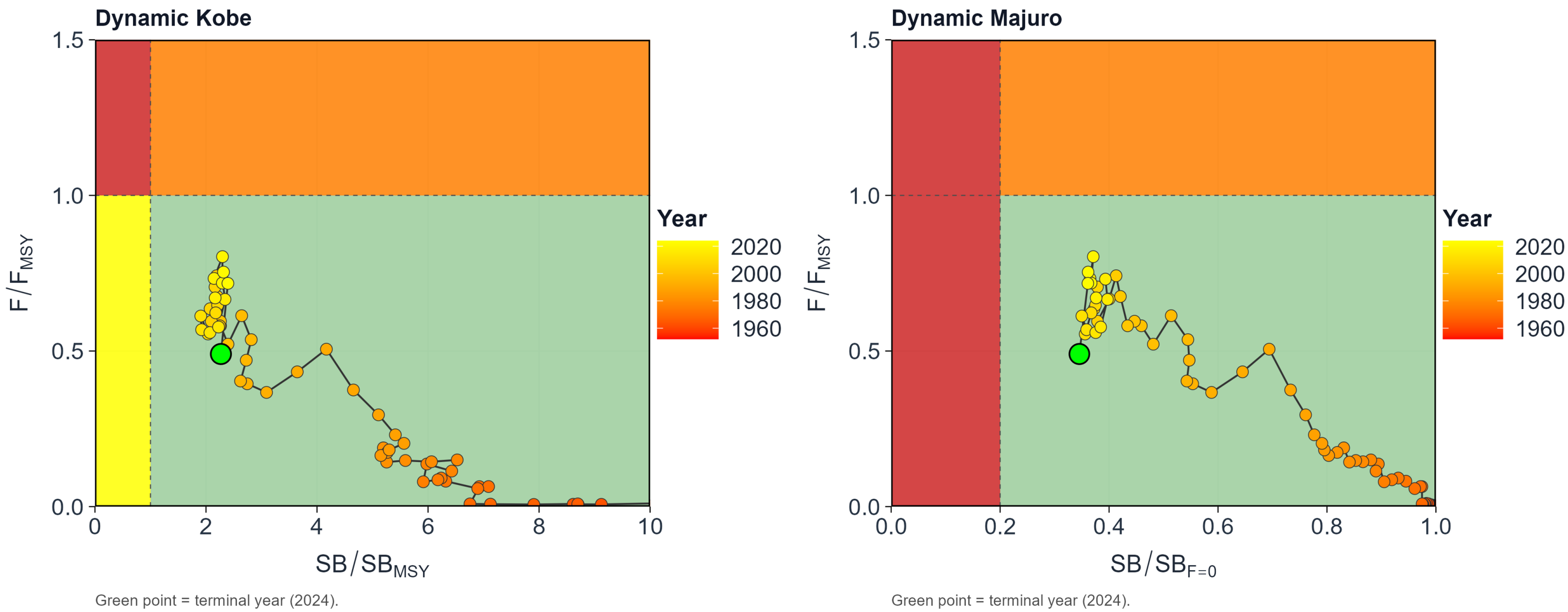

Figure 57: Time-dynamic Majuro and Kobe trajectories for the 2026 diagnostic model.

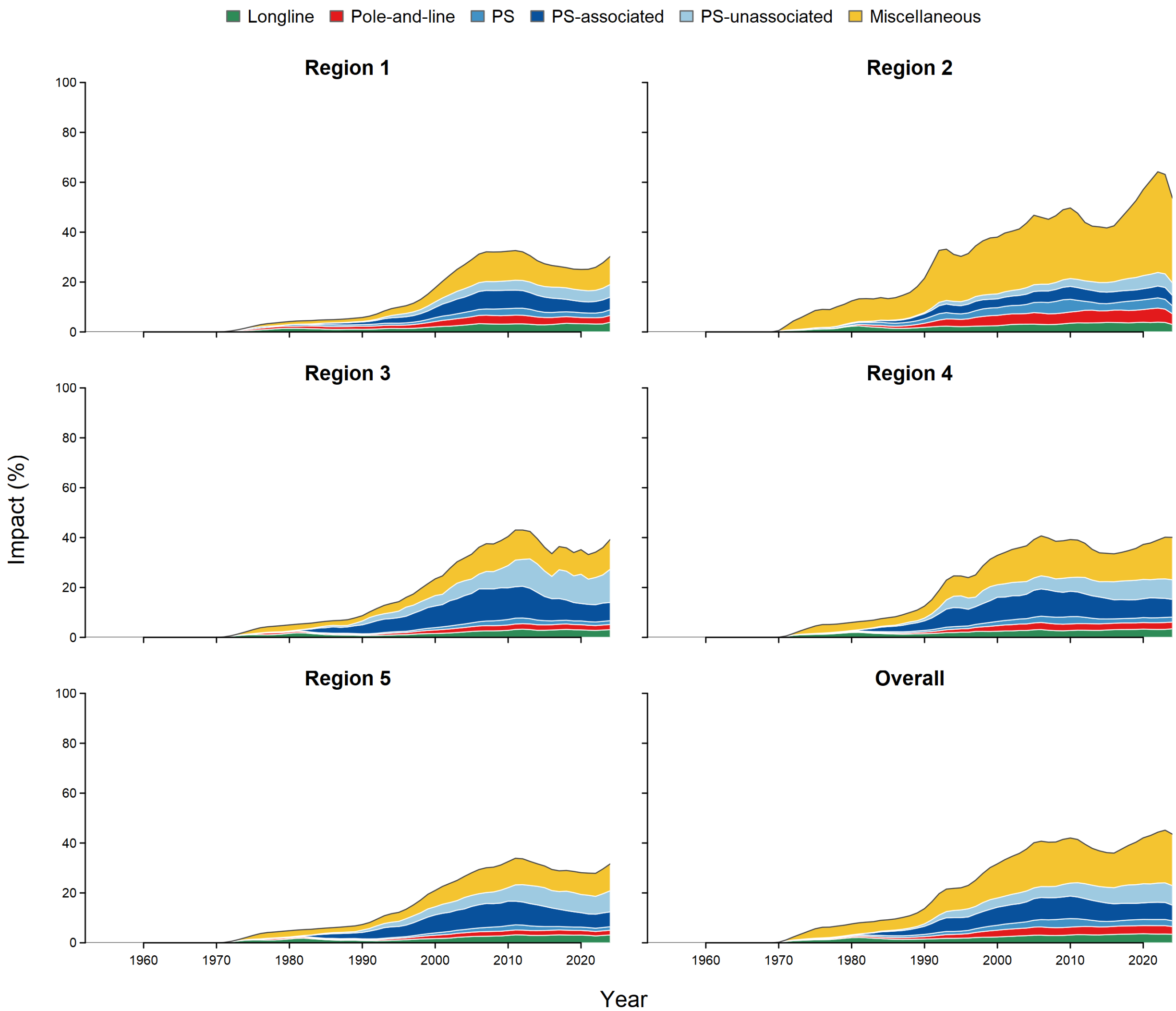

Figure 58: Estimates of fishery impact, or reduction in spawning potential due to fishing (Fishery Impact $= 1 - SB_t/SB_{t,F=0}$) by region, and over all regions (bottom right), attributed to various fishery groups for the diagnostic model.

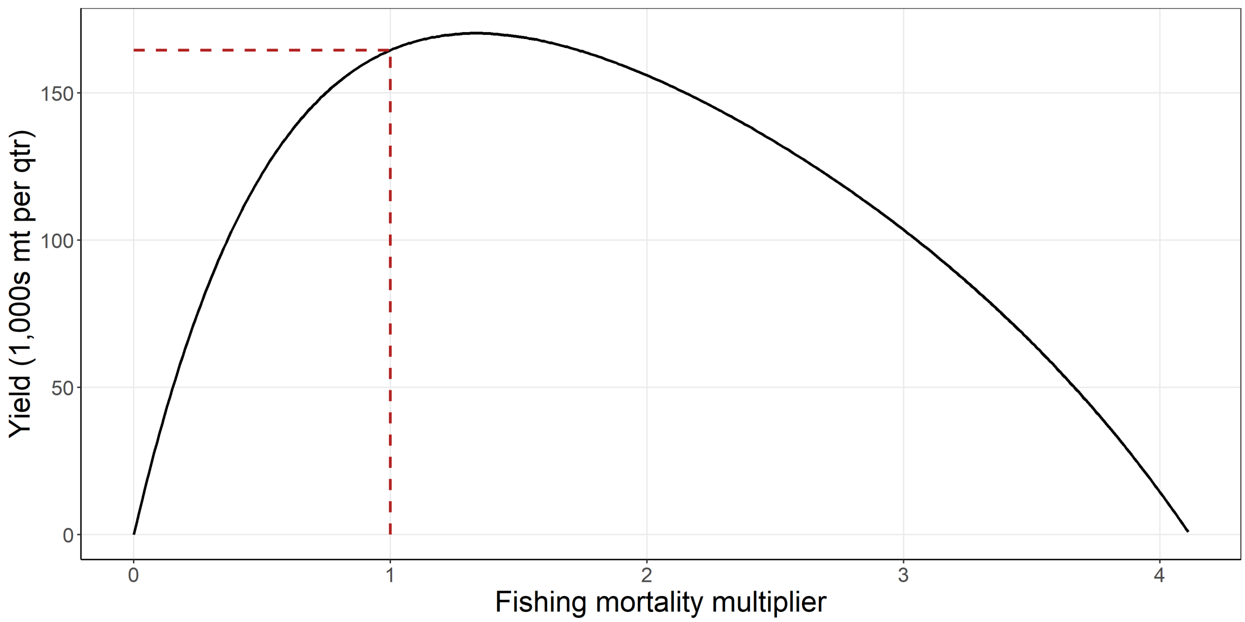

Figure 59: Estimated yield as a function of fishing mortality multiplier for the diagnostic model and the alternative mixing scenario model. The red dashed line indicates the equilibrium yield at current fishing mortality.

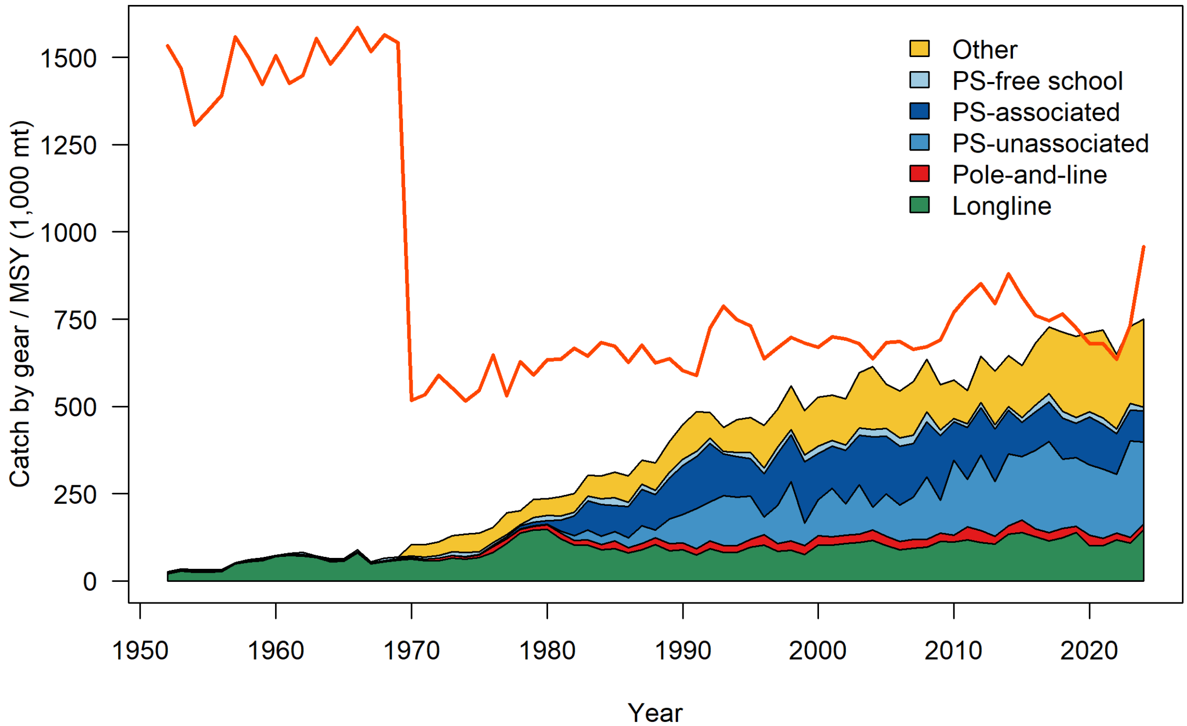

Figure 60: History of the annual estimates of MSY (red line) for the diagnostic model compared with annual catch by the main gear types. Note that this is a 'dynamic' MSY.

12 References

13 References

Aguila, R. D., Perez, S. K. L., Catacutan, B. J. N., Lopez, G. V., Barut, N. C., and Santos, M. D. (2015). Distinct yellowfin tuna (thunnus albacares) stocks detected in the western and central pacific ocean (WCPO) using DNA microsatellites. *PLOS ONE*, 10(9):e0138292.

Anderson, G., Lal, M., Hampton, J., Smith, N., and Rico, C. (2019-07). Close kin proximity in yellowfin tuna (thunnus albacares) as a driver of population genetic structure in the tropical western and central pacific ocean. *Front. Mar. Sci.*, 6:341.

Anderson, S. C., Ward, E. J., English, P. A., and Barnett, L. A. K. (2022-03). sdmTMB: an r package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. preprint, Ecology.

Andrews, A. H., Eveson, J. P., Welte, C., Okamoto, K., Satoh, K., Krusic-Golub, K., Lougheed, B. C., Macdonald, J. I., Roupsard, F., and Farley, J. H. (2024-08-19). Age validation of yellowfin and bigeye tuna using post-peak bomb radiocarbon dating confirms long lifespans in the western and central pacific ocean. *ICES Journal of Marine Science*, 81(6):1137–1149.

Berger, A. M., McKechnie, S., Abascal, F., Kumasi, B., Usu, T., and Nichol, S. J. (2014-08). Analysis of tagging data for the 2014 tropical tuna assessments: data quality rules, tagger effects, and reporting rates. Technical Report WCPFC-SC10-2014/SA-IP-06, WCPFC.

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.

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K. R., Maunder, M. N., Taylor, I., Wetzel, C. R., Doering, K., Johnson, K. F., and Methot, R. D. (2021-08). A cookbook for using model diagnostics in integrated stock assessments. *Fisheries Research*, 240:105959.

Castillo Jordan, C., Teears, T., Hampton, J., Davies, N., Scutt Phillips, J., McKechnie, S., Peatman, T., Macdonald, J., Day, J., Magnusson, A., Scott, R., Scott, F., Pilling, G., and Hamer, P. (2022). Stock assessment of skipjack tuna in the western and central pacific ocean: 2022. Technical Report WCPFC-SC18-SA-WP-01, WCPFC.

Davies, N., Bouye, F., Kim, K., Magnusson, A., and Hampton, J. (2026a). Developments in the MULTIFAN-CL software 2025-26. Technical Report WCPFC-SC22-2026/SA-IP-02, WCPFC.

Davies, N., Bouye, F., Kim, K., Magnusson, A., and Hampton, J. (2026b). Developments in the MULTIFAN-CL software 2025-26. Technical Report WCPFC-SC22-2026/SA-IP-02, WCPFC.

Davies, N., Fournier, D., and Hampton, J. (2019). Developments in the MULTIFAN-CL software 2018-2019. Technical Report WCPFC-SC15-2019/SA-IP-02, WCPFC.

Davies, N., Fournier, D. A., Hampton, J., Kleiber, P., Bouye, F., Kim, K., Magnusson, A., and Hoyle, S. (2026c). MULTIFAN-CL user's guide version 2.2.7.8. Technical report, WCPFC.

Davies, N., Harley, S., Hampton, J., and McKechnie, S. (2014-08). Stock assessment of yellowfin tuna in the western and central pacific ocean. Technical Report WCPFC-SC10-2014/SA-WP-04, WCPFC.

Day, J., Magnusson, A., Teears, T., Hampton, J., Davies, N., Castillo Jordán, C., Peatman, T., Scott, R., Scutt Phillips, J., McKechnie, S., Scott, F., Yao, N., Natadra, R., Pilling, G., Williams, P., and Hamer, P. (2023). Stock assessment of bigeye tuna in the western and central pacific ocean: 2023. Technical Report WCPFC-SC19-2023/SA-WP-05 (Rev. 2), WCPFC.

Ducharme-Barth, N., Vincent, M., Hampton, J., Hamer, P., Williams, P., and Pilling, G. (2020a). Stock assessment of bigeye tuna in the western and central pacific ocean. Technical Report WCPFC-SC16-2020/SA-WP-03 [REV3], WCPFC.

Ducharme-Barth, N., Vincent, M., Vidal, T., and Hamer, P. (2020b). Analysis of pacific-wide operational longline dataset for bigeye and yellowfin tuna catch-per-unit-effort (CPUE). Technical Report SC16-SA-IP-07, WCPFC.

Evans, K., Grewe, P. M., Foster, S. D., Gosselin, T., Gunasekera, R. M., and Lansdell, M. (2019). Connectivity of tuna and billfish species targeted by the australian eastern tuna and billfish fishery with the broader western pacific ocean. Technical Report WCPFC-SC15-2019/SA-IP13, WCPFC.

Evans, K., Langley, A., Clear, N., Williams, P., Patterson, T., Sibert, J., Hampton, J., and Gunn, J. (2008). Behaviour and habitat preferences of bigeye tuna (thunnus obesus) and their influence on longline fishery catches in the western coral sea. *Canadian Journal of Fisheries and Aquatic Sciences*, 65:2427–2443.

Eveson, P., Vincent, M., Farley, J., Krusic-Golub, K., and Hampton, J. (2020). Integrated growth models from otolith and tagging data for yellowfin and bigeye tuna in the western and central pacific ocean. Technical Report SC16-SA-IP-03, WCPFC.

Farley, J., Eveson, P., Krusic-Golub, K., Sanchez, C., Rouspard, F., McKechnie, S., Nichol, S., Leroy, B., Smith, N., and Chang, S.-K. (2017). Age, growth and maturity of bigeye tuna in the pacific. Technical Report WCPFC-SC13- 2017/SA-WP-01, WCPFC.

Farley, J., Krusic-Golub, K., Eveson, P., Clear, N., Rouspard, F., Sanchez, C., Nicol, S., and Hampton, J. (2020). Age and growth of yellowfin and bigeye tuna in the western and central pacific ocean from otoliths. Technical Report SC16-SA-WP-02, WCPFC.

Fonteneau, A. (2002). Estimated sex ratio of large yellowfin taken by purse seiners in the indian ocean: comparison with other oceans. Technical Report IOTC Proc., 5: 279-281., IOTC.

Fournier, D. and Archibald, C. P. (1982). A general-theory for analyzing catch at age data. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(8):1195–1207.

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

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

Francis, R. I. C. C. (1992). Use of risk analysis to assess fishery management strategies: A case study using orange roughy (ıt hoplostethus atlanticus) on the chatham rise, new zealand. *Canadian Journal of Fisheries and Aquatic Science*, 49:922–930.

Grewe, P. M., Feutry, P., Hill, P. L., Gunasekera, R. M., Schaefer, K. M., Itano, D. G., Fuller, D. W., Foster, S. D., and Davies, C. R. (2015). Evidence of discrete yellowfin tuna (thunnus albacares) populations demands rethink of management for this glglobal important resource. *Nature Science Reports*, 5(16915).

Gunn, J. S., Clear, N. P., Ward, R., Appleyard, S., Grewe, P. M., and Innes, B. (2002). The origin of recruits to the east coast yellowfin tuna fishery and the delineation of the structure of yellowfin stocks in the western pacific. Final report to the fisheries research and development corporation, CSIRO Marine Research, Hobart, TAS, Australia.

Hamer, P. (2026). Summary report from the SPC pre-assessment workshop - march 2026. Technical Report WCPFC-SC22-2026/SA-IP-01, WCPFC.

Hamer, P., Castillo Jordán, C., Panizza, A., and Peatman, T. (2026). WCPFC project127/127a - review and reconciliation of size data collected in the WCPFC-CA for stock assessment purposes: phase 2 progress report. Technical Report WCPFC-SC22-2026/ST-WP-02, WCPFC.

Hamer, P., Potts, J., Macdonald, J., and Senina, I. (2023). Review and analyses to inform conceptual models of population structure and spatial stratification of bigeye and yellowfin tuna assessments in the western and central pacific ocean. Technical Report SC19-SA-WP-02, WCPFC.

Hamer, P., Schneiter, E., Vidal, T., and Williams, P. (2025). Progress report: Review and reconciliation of size data collected in the WCPFC-CA for stock assessment purposes (WCPFC project: 127). Technical Report WCPFC-SC21-ST-WP-02, WCPFC.

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

Hampton, J., Davies, N., and Hamer, P. (2026a). Investigation of the impacts of regional growth variation on the 2023 bigeye tuna stock assessment. Technical Report WCPFC-SC22-2026/SA-IP11, WCPFC.

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 Kleiber, P. (2003). Stock assessment of yellowfin tuna in the western and central pacific ocean. Technical Report WP YFT-1, SCTB 16, Mooloolaba, Australia, 9-16 July 2003., SCTB.

Hampton, J., Kleiber, P., Langley, A., Takeuchi, Y., and Ichinokawa, M. (2005). Stock assessment of yellowfin tuna in the western and central pacific ocean. Technical Report WCPFC SC1 SA WP-1, Noumea, New Caledonia, 8-19 August 2005., WCPFC.

Hampton, J., Teears, T., Kim, K., Magnusson, A., Davies, N., and Hamer, P. (2026b). Investigation of single region fleets as areas models for bigeye and yellowfin tuna in the WCPO. Technical Report WCPFC-SC22-2026/SA-IP-09, WCPFC.

Hampton, J. and Williams, P. (2016). Annual estimates of purse seine catches by species based on alternative data sources. Technical Report WCPFC-SC12-2016/ST-IP-03, WCPFC.

Hare, S., Vidal, T., Castillo Jordán, C., Day, J., Hamer, P., Hill, N., Machful, P., Magnusson, A., Nicol, S., Odongo, V., Panizza, A., Peatman, T., Scott, R., Senina, I., Sloyan, B., Teears, T., and Pilling, G. (2025). The western and central pacific una fishery: 2024 overview and status of stocks. Tuna fisheries assessment report no. 25, SPC.

Hare, S., Vidal, T., Scott, R., and Pilling, G. (2026). A compendium of fisheries indicators for target tuna stocks in the WCPFC convention area. Technical Report WCPFC-SC22-2026/SA-WP-07, WCPFC.

Harley, S. J. (2011). A preliminary investigation of steepness in tunas based on stock assessment results. Technical Report WCPFC-SC7-2011/SA-IP-08, WCPFC.

Harley, S. J., Davies, N., Hampton, J., and McKechnie, S. (2014-08). Stock assessment of bigeye tuna in the western and central pacific ocean. Technical Report WCPFC-SC10-2014/SA-WP-01, WCPFC.

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

Hasegawa, T., Okamoto, K., Ijima, H., Kim, K., Hamer, P., Wakefield, C., Pons, A., Widdrington, J., and Davies, N. (2026). New age-at-length data for WCPO bigeye and yellowfin tuna from the temperate north pacific and approaches to address spatial sampling bias in growth estimation within MFCL. Technical Report WCPFC-SC22-2026/IP-08, WCPFC.

Hoyle, S. and Nicol, S. (2008). Sensitivity of bigeye stock assessment to alternative biological and reporductive assumptions. Technical Report WCPFC-SC4-2008/ME-WP-01, WCPFC.

Hoyle, S. D. (2008). Adjusted biological parameters and spawning biomass calculations for south pacific albacore tuna, and their implications for stock assessments. Technical Report WCPFC-SC4-2008/ME-WP-02, WCPFC.

Hoyle, S. D., Williams, A. J., Minte-Vera, C. V., and Maunder, M. N. (2023-01). Approaches for estimating natural mortality in tuna stock assessments: Application to global yellowfin tuna stocks. *Fisheries Research*, 257:106498.

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, WCPFC.

Ijima, H. and Jusup, M. (2023-04). Tuna and billfish larval distributions in a warming ocean.

ISSF (2011). Report of the 2011 ISSF stock assessment workshop. Technical Report ISSF Technical Report 2011-02, ISSF.

Itano, D. (2000). The reproductive biology of yellowfin tuna (ıt thunnus albacares) in hawaiian waters and the western tropical pacific ocean: Project summary. JIMAR Contribution 00-328 SOEST 00-01, University of Hawai'i.

Kolody, D. and Hoyle, S. (2014). Evaluation of tag mixing assumptions in western pacific ocean skipjack tuna stock assessment models. *Fisheries Research*, http://dx.doi.org/101016/j.fishres.2014.05.008.

Langley, A., Hoyle, S. D., and Hampton, J. (2011). Stock assessment of yellowfin tuna in the western and central pacific ocean. Technical Report WCPFC-SC7-2011/SA-WP-03, WCPFC.

Magnusson, A., Day, J., Teears, T., Hampton, J., Davies, N., Castillo Jordán, C., Peatman, T., Scott, R., Scutt Phillips, J., McKechnie, S., Scott, F., Yao, N., Natadra, R., Pilling, G., Williams, P., and Hamer, P. (2023). Stock assessment of yellowfin tuna in the western and central pacific ocean: 2023. Technical Report WCPFC-SC19-2023/SA-WP-04 (Rev. 2), WCPFC.

McKechnie, S., Harley, S. J., Davies, N., Rice, J., Hampton, J., and Berger, A. (2014). Basis for regional structures used in the 2014 tropical tuna assessments, including regional weights. Technical Report WCPFC-SC10-2014/SA-IP-02, WCPFC.

McKechnie, S., Ochi, D., Kiyofuji, H., Peatman, T., and Caillot, S. (2016). Construction of tagging data input files for the 2016 skipjack tuna stock assessment in the western and central pacific ocean. Technical Report WCPFC-SC12-2016/SA-IP-05, WCPFC.

McKechnie, S., Pilling, G., and Hampton, J. (2017). Stock assessment of bigeye tuna in the western and central pacific ocean. Technical Report WCPFC-SC13-2017/SA-WP-05, WCPFC.

McKechnie, S., Tremblay-Boyer, L., and Harley, S. J. (2015). Analysis of pacific-wide operational longline CPUE data for bigeye tuna. Technical Report WCPFC-SC11-2015/SA-WP-03, WCPFC.

McPherson, G. (1991). Reproductive biology of yellowfin tuna in the eastern australian fishing zone, with special reference to the north-western coral sea. *Mar. Freshwater Res.*, 42(5):465.

Moore, B. R., Bell, J. D., Evans, K., Farley, J., Grewe, P. M., Hampton, J., Marie, A. D., Minte-Vera, C., Nicol, S., Pilling, G. M., Scutt Phillips, J., Tremblay-Boyer, L., Williams, A. J., and Smith, N. (2020-10). Defining the stock structures of key commercial tunas in the pacific ocean i: Current knowledge and main uncertainties. *Fisheries Research*, 230:105525.

Mullins, R. B., McKeown, N. J., Sauer, W. H. H., and Shaw, P. W. (2018-12). Genomic analysis reveals multiple mismatches between biological and management units in yellowfin tuna (thunnus albacares). *ICES Journal of Marine Science*, 75(6):2145–2152.

Nishikawa, Y., Honma, M., Ueyanagi, S., and Kikawa, S. (1985). Average distribution of larvae of oceanic species of scombroid fishes, 1956-1981. *Far Seas Fish.Res.Lab., 99 p.*

Peatman, T., Gislard, S., Scutt Phillips, J., Potts, J., and Nicol, S. (2026a). Analysis of tag seeding data for the 2026 yellowfin & bigeye assessments: Reporting rates for purse seine fleets. Technical Report WCPFC-SC22-2026/SA-IP-05, WCPFC.

Peatman, T., Kim, K., Magnusson, A., Castillo Jordan, C., and Hamer, P. (2026b). Analysis of size frequency data for the 2026 bigeye and yellowfin assessments. Technical Report WCPFC-SC22-SA-IP06, WCPFC.

Peatman, T., Scutt Phillips, J., Potts, J., and Nicol, S. (2026c). Analysis of tagging data for the 2026 yellowfin & bigeye assessments: Corrections to tag releases for tagging conditions. Technical Report WCPFC-SC22-2026/SA-IP-04, WCPFC.

Peatman, T., Williams, P., and Nicol, S. (2021). Project 60: Progress towards acheiving SC16 recomendations. Technical Report WCPFC-SC17-2021/ST-IP-04, WCPFC.

Pecoraro, C., Babbucci, M., Franch, R., Rico, C., Papetti, C., Chassot, E., Bodin, N., Cariani, A., Bargelloni, L., and Tinti, F. (2018-12). The population genomics of yellowfin tuna (thunnus albacares) at global geographic scale challenges current stock delineation. *Sci Rep*, 8(1):13890.

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. Technical Report WCPFC-SC12-2016/MI-IP-04, Bali, Indonesia, 3-11 August 2016.

Proctor, C., Lester, R., Clear, N. P., Grewe, P. M., Moore, B. R., Eveson, J. P., Lestari, P., Wudji, A., Taufik, M., Wudianto, Lansdell, M., Hill, P., Dietz, C., Thompson, J., Cutmore, S., Foster, S., Gosselin, T., and Davies, C. R. (2019). Population structure of yellowfin tuna (thunnus albacares) and bigeye tuna (t. obesus) in the indonesian region. Technical Report Final Report as output of ACIAR Project FIS/2009/059, ACIAR.

Punt, A. E., Maunder, M. N., and Ianelli, J. (2023). Independent review of recent WCPO yellowfin tuna assessment. Technical Report SC19-SA-WP-01, WCPFC.

Reglero, P., Tittensor, D., Álvarez Berastegui, D., Aparicio-González, A., and Worm, B. (2014-03). Worldwide distributions of tuna larvae: revisiting hypotheses on environmental requirements for spawning habitats. *Mar. Ecol. Prog. Ser.*, 501:207–224.

Rooker, J. R., David Wells, R. J., Itano, D. G., Thorrold, S. R., and Lee, J. M. (2016-05). Natal origin and population connectivity of bigeye and yellowfin tuna in the pacific ocean. *Fish. Oceanogr.*, 25(3):277–291.

Satoh, K., Katayama, S., Tanabe, T., and Okamoto, K. (2023-06-29). Connectivity in juvenile yellowfin tuna thunnus albacares between temperate and tropical regions of the western pacific ocean. *Marine Ecology Progress Series*, 713:151–172.

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

Schaefer, M.B., Broadhead, G.C., and Orange, C.J. (1963). Synopsis on the biology of yellowfin tuna thunnus (neothumus) albacares (bonnaterre) 1788 (pacific ocean). Technical report, FAO.

Scutt Phillips, J., Lehodey, J., Hampton, J., Senina, I., and Nicol, S. (2022). Quantifying rates of mixing in tagged, WCPO skipjack tuna. Technical Report WCPFC-SC18-2022/SA-WP-04, WCPFC.

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.

Scutt Phillips, J., Senina, I., and Bonin, L. (2026). Estimation of tag mixing periods for teh 2026 WCPO tuna stock assessments. Technical Report WCPFC-SC22-2026-SA-IP-10, WCPFC.

Servidad-Bacordo, R., Dickson, A., Nepomuceno, L., and Ramiscal, R. (2012). Composition, distribution and abundance of fish eggs and larvae in the philippine pacific seaboard and celebes sea with focus on tuna larvae (family: Scombridae). Information paper WCPFC-SC8-2012/SA-IP-03, WCPFC.

Sun, C.-L., Chu, S.-L., and Yeh, S.-Z. (2006). Reproductive biology of bigeye tuna in the western and central pacific ocean. Technical Report WCPFC-SC2-2006/BI-WP-01, WCPFC.

Teears, T. (2026). Background analyses and data inputs for the 2026 WCPO yellowfin and bigeye tuna stock assessment. Technical Report WCPFC-SC22-2026-SA-IP03, WCPFC.

Teears, T., Castillo Jordán, C., Davies, N., Day, J., Hampton, J., Magnusson, A., Peatman, T., Pilling, G., Xu, H., Vidal, T., Williams, P., and Hamer, P. (2024). Stock assessment of south pacific albacore: 2024. Technical Report WCPFC-SC20-2024/SA-WP-02-Rev 2, WCPFC.

Teears, T., Davies, N., Hamer, P., Magnusson, A., Peatman, T., Scutt Phillips, J., Vidal, T., and Hampton, J. (2025). Stock assessment of skipjack tuna in the western and central pacific ocean: 2025 (rev 3). Technical Report WCPFC-SC21-2025/SA-WP-02_rev_3, WCPFC.

Teears, T., Day, J., Hampton, J., Magnusson, A., McKechnie, S., Peatman, T., Scutt Phillips, J., Williams, P., and Hamer, P. (2023). CPUE analysis and data inputs for the 2023 bigeye and yellowfin tuna assessments in the WCPO. Technical Report WCPFC-SC19-2023/SA-WP-03, WCPFC.

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

Thorson, J. T., Shelton, A. O., Ward, E. J., and Skaug, H. J. (2015). Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for west coast groundfishes. *Ices Journal of Marine Science*, 72(5):1297–1310.

Tremblay-Boyer, L., McKechnie, S., Pilling, G., and Hampton, J. (2017). Stock assessment of yellowfin tuna in the western and central pacific ocean. Technical Report WCPFC-SC13-2017/SA-WP-06, WCPFC.

Vidal, T. and Hamer, P. (2020). Developing yellowfin tuna recruitment indices from drifting FAD purse seine catch and effort data. Technical Report WCPFC-SC16-SA-IP-08, WCPFC.

Vidal, T., Heather, F., Satria, F., Hai, V., Barcoma, S., Duong, N., Sadiyah, L., Anas, M., Januar, A., Rachmadi, D., and Suadela, P. (2026). Data improvements in the west pacific east asia (WPEA) region of WCPFC. Technical Report WCPFC-SC22-2026/ST-IP-05, WCPFC.

Vincent, M. and Ducharme-Barth, N. (2020). Background analyses for the 2020 stock assessments of bigeye and yellowfin tuna in the western and central pacific ocean. Technical Report SC16-SA-IP-06, WCPFC.

Vincent, M., Ducharme-Barth, N., Hamer, P., Hampton, J., Williams, P., and Pilling, G. (2020). Stock assessment of yellowfin tuna in the western and central pacific ocean. WCPFC-SC16-2020/SA-WP-04-rev2, WCPFC.

Vincent, M., Pilling, G., and Hampton, J. (2018). Incorporation of updated growth information within the 2017 WCPO bigeye stock assessment grid, and examination of the sensitivity of estimates to alternative model spatial structures. Technical Report WCPFC-SC14-2018/ SA-WP-03, Oceanic Fisheries Programme, The Pacific Community.

Wakefield, C., Widdrington, J., Krusic-Golub, K., Pons, A., Reis-Santos, P., Gillanders, B., Farley, J., Nicol, S., and Hamer, P. (2026). Project 129 update: Building an age-length data stream for tuna assessments. Technical Report WCPFC-SC22-2022/SA-WP-12, WCPFC.

Ward, R., Elliott, N., and Grewe, P. (1994). Allozyme and mitochondrial DNA variation in yellowfin tuna (thunnus albacares) from the pacific ocean. *Marine Biology*, 118:531–539.

Wells, R., Rooker, J., and Itano, D. (2012-08). Nursery origin of yellowfin tuna in the hawaiian islands. *Mar. Ecol. Prog. Ser.*, 461:187–196.

Zhu, G., Xu, L., Zhou, Y., and Song, L. (2008-08). Reproductive biology of yellowfin tuna t. albacares in the west-central indian ocean. *J. Ocean Univ. China*, 7(3):327–332.