Western and Central Pacific Fisheries Commission

SCIENTIFIC COMMITTEE
TWENTIETH REGULAR SESSION

Manila, Philippines
14–21 August 2024

Stock Assessment of South Pacific Albacore: 2024

WCPFC-SC20-2024/SA-WP-02-Rev 3
14 August 2024

T. Teears[1], C. Castillo-Jordán[1], N. Davies[2], J. Day[1], J. Hampton[1], A. Magnusson[1], T. Peatman[2], G. Pilling[1], H. Xu[3], T. Vidal[1], P. Williams[1] and P. Hamer[1]

[1]Oceanic Fisheries Programme of the Pacific Community
[2]Consultant for The Pacific Community, Oceanic Fisheries Programme
[3]Inter-American Tropical Tuna Commission, La Jolla, United States

Rev 1: error in calculation of MSY corrected in results table.

Rev 2: This version contains the full assessment report. There are some minor adjustments to some values for the reference points that include the estimation uncertainty at the bottom of Table 9.

Rev 3: This revision has an update to Figure 24 (fishery-specific LF likelihood profile) with formatting changes for clarity, Figure 19 (CPUE fits) and Figure 22 (CAAL fits) have been updated with standardised residuals, Figure 20 (aggregated LF composition fits) has additional text for each fishery with true effective sample size (as derived from robust normal likelihood), and Table 9 has an additional line with estimates of recent depletion for the WCPFC-CA only.

Contents

1 Executive Summary

This paper describes the 2024 stock assessment of albacore tuna (*Thunnus alalunga*) across the South Pacific Ocean (south of the equator), incorporating the Convention areas of the Western and Central Pacific Fisheries Commission (WCPFC–CA) and the Inter American Tropical Tuna Commission (IATTC). A further three years of data are available since the last stock assessment was conducted in 2021; the model time period now extends from 1954–2022. The 2021 assessment was the first to include both convention areas modelled jointly in a spatially structured South Pacific wide assessment. The 2024 assessment also includes both convention areas however, an areas-as-fleets approach was implemented in each of the convention areas in lieu of the explicit regional spatial structure used in 2021.

Based on recommendations from SC17 and the 2024 Pre-assessment workshop, there was a strong focus on simplifying the 2024 assessment compared to previous versions. Key changes made from the 2021 to the 2024 diagnostic case model include:

- Updating all data to the end of 2022 and applying a new version of MULTIFAN-CL (2.2.7.0)
- Conversion from a catch-errors to a catch-conditioned modelling framework, and the inclusion of a likelihood component for the CPUE from the index fisheries.
- Application of time-varying coefficient of variation (CV) for index fisheries.
- Collapsed the WCPFC-CA subregions and regions to a single region.
- Change from quarterly to annual recruitment, occurring in October of each year.
- Implementation of an areas-as-fleets approach for both the WCPFC-CA and the remaining area of the eastern Pacific Ocean (EPO) with fisheries structure informed by a regression tree spatial structure analysis of longline size composition data.
- Development of annual indices from operational longline data with the final indices for the WCPFC-CA restricted to the tropical area, a WCPFC-CA juvenile index from New Zealand troll fishery data, and a full EPO index from operational longline data.
- Growth estimation informed by conditional-age-at-length (CAAL) data based on validated otolith readings and fisheries size compositions. More detail on growth estimation changes below.
- Lorenzen natural mortality at age, with the A_{max} method used to provide average M values, with a max age of 15 years (Lorenzen, 2022).
- Effective sample sizes for size composition data calculated using the Francis weighting approach (Francis, 2017).
- Movement and recruitment distribution fixed to values derived from SEAPODYM (Senina et al., 2020).

- Minor updates to the weight-length a and b parameters.
- Uncertainty estimated using a Monte Carlo ensemble model approach in which 100 models incorporated uncertainty in average natural mortality, stock-recruitment steepness and estimation error for individual models.
- A comprehensive sensitivity analysis to support the structure of the ensemble and other modelling decisions.

These changes have been informed by a number of sources, including:

- The recommendations of the 2022 peer review of the WCPFC yellowfin tuna assessment (Punt et al., 2023) where they apply to South Pacific albacore.
- Comments received on the previous assessment of this stock (Castillo Jordan et al., 2021) delivered at SC17 and subsequently;
- The recommendations of the CAPAM "Tuna Stock Assessment Good Practices Workshop" held in Wellington, New Zealand in March 2023.
- The recommendations received from the 2024 SPC Pre-Assessment Workshop and follow-up inputs from PAW participants (Hamer, 2024).

The assessment is supported by the analysis of catch and effort data to develop CPUE abundance indices, and size composition data (Teears et al., 2024; Potts et al., 2024). Other papers of relevance to the assessment are the analysis of NZ troll fishery size data and CPUE (Neubauer and Hill-Moana, 2024).

In addition to the diagnostic model, a number of other models were investigated as sensitivities to assess the relative impacts of alternative data and model assumptions on the estimated assessment results and conclusions. These sensitivities included alternative CPUE indices, down weighting of NZ troll CPUE, removing troll size composition data, alternative weights for the CAAL data, alternative movement rates between the WCPFC-CA region and the EPO, effort creep in longline CPUE indices, alternative recruitment distributions between the WCPFC-CA region and the EPO, different numbers of age classes used in the model, selectivity and catchability time blocks, growth models, natural mortality (M) and steepness (h). Along with the commonly used diagnostic tools in previous assessments, additional diagnostic tools are applied to the diagnostic model including age structured production models and catch curve analysis.

From the extensive sensitivity analysis, it was determined that the key assumptions influencing uncertainty in management quantities were the M and h values. The uncertainty ensemble was therefore based on capturing the uncertainty in the average M for the Lorenzen M-at-age and steepness. In contrast to the 2021 assessment, size composition data weighting is not included in the uncertainty characterisation due to the use of Francis weighting in 2024. The new growth estimation used the CAAL data, fixed L1, and the estimation of offsets to the VB curve for ages

2, 3 and 4 years to better represent the almost linear growth across the juvenile age classes. This approach estimated growth that was consistent with the three annual size modes in the NZ troll data and the CAAL data, which could not be achieved satisfactorily with the standard VB model.

A key improvement to the uncertainty characterisation was the implementation of a Monte Carlo ensemble approach to incorporate the uncertainty in M and h. The method involved constructing prior distributions for average M and h and drawing independent replicates from these priors. One hundred pairs of average M and h were chosen to construct the uncertainty ensemble of 100 models. These models were then assessed for retention in the ensemble based on convergence and data fit criteria. All models were ultimately retained. Estimation uncertainty for the key management quantities for each of the 100 models in the ensemble was incorporated into the probability distributions and statistics for those quantities.

Based on the ensemble of models, the general conclusions of this assessment are as follows:

- Consistent with the findings of the previous south Pacific albacore assessment (Castillo Jordan et al., 2021), the spawning biomass shows a sharp decline from the start of the model period until the mid-1970s after which it stabilises. The stock status, as indicated by the spawning biomass depletion, shows a more gradual long-term decline from the start of the model period.
- Although estimates for recent years of stock assessments have higher uncertainty and should be interpreted with caution, the dip in spawning biomass depletion that was a focus in the last assessment is moderated in the new assessment, and there are recent signs that the overall stock status has improved.
- Fishing mortality on adults continues to increase, while fishing mortality on juveniles remains low. Fishing mortality has increased sharply since 2010 in the EPO as the longline catches have increased, but has stabilised in the WCPFC-CA over a similar time period.
- Recruitment shows similar interannual variability across years, with an increasing trend from the late 1990s becoming more evident in the estimates.
- Overall, the median depletion from the model ensemble with estimation uncertainty for the recent period (2019-2022; $SB_{\text{recent}}/SB_{F=0}$) is 0.48 (80 percentile range 0.36–0.62, full range 0.23–0.77).
- The median recent fishing mortality from the model ensemble with estimation uncertainty is below the level for achieving MSY (median $F_{\text{recent}}/F_{\text{MSY}} = 0.18$, 80 percentile range 0.06–0.44, full range 0.03–1.00).
- The median recent spawning biomass from the model ensemble with estimation uncertainty is well above the spawning biomass to achieve MSY (median $SB_{\text{recent}}/SB_{\text{MSY}} = 3.02$, 80 percentile range 2.04–5.21, full range 1.20–8.96).
- All models in the uncertainty ensemble had $SB_{\text{recent}}/SB_{F=0} > 0.2$, the Limit Reference Point

for WCPFC key tuna stocks.

- For each model in the ensemble, the ratio of the $SB_{\text{recent}}/SB_{F=0}$ to the interim Target Reference Point (iTRP) estimated for that model was calculated. Across the 100 models the median ratio of $SB_{\text{recent}}/SB_{F=0}$:iTRP was 0.952, ranging from 0.899 to 1.016. Therefore, the recent stock status is close to the iTRP (estimated to be a median depletion of 0.50 across the model ensemble).

These results are broadly consistent with the previous 2021 stock assessment and suggest that the albacore stock across the South Pacific is not overfished nor undergoing overfishing. The recent stock status is around the iTRP.

2 Introduction

This paper presents the 2024 stock assessment of South Pacific albacore tuna (*Thunnus alalunga*). As requested by the Scientific Committee (SC) of the Western and Central Pacific Fisheries Commission (WCPFC), the 2021 assessment covered the entire South Pacific Figure 1. This assessment continues that coverage to include the albacore fisheries from the equator to 50°S in the Pacific Ocean, incorporating the convention areas of the WCPFC and the Inter American Tropical Tuna Commission (IATTC) (Figure 2). Since 1999, South Pacific albacore has been assessed regularly for the WCPFC convention area (Castillo Jordan et al., 2021; Harley et al., 2015; Hoyle and Davies, 2009; Hoyle et al., 2012; Tremblay-Boyer et al., 2018), with two assessments that covered the entire South Pacific (Castillo Jordan et al., 2021; Hoyle et al., 2012). The 2021 assessment has been the only previous assessment of albacore in the South Pacific with a separate model region for the south eastern Pacific Ocean (EPO) under the management of the IATTC. The 2012 assessment conducted by Hoyle et al. (2012) included the EPO as a fishery stratum in a single region areas-as-fleets model. In the 2024 assessment, a simplified two region structure is employed with one region defined as the WCPFC Convention Area (WCPFC-CA) and the second region defined as the EPO. Within each region, an areas-as-fleets approach was applied. As was done in 2021, the 2024 assessment was accomplished via collaboration between scientists from the Pacific Community (SPC, the Scientific Services Provider for the WCPFC) and the IATTC.

The 2024 assessment continues the development of stock assessment models for South Pacific albacore, facilitated by the ongoing development of the statistical stock assessment software, known as MULTIFAN-CL [4] (Fournier et al., 1998; Hampton and Fournier, 2001; Kleiber et al., 2019), that is routinely used by SPC for assessments of tuna and tuna-like species. Each new assessment can involve updates to fishery input data, implementation of new features in the MULTIFAN-CL modelling software, and consideration of new information on biology, population structure and other population vital rates. These changes are an important part of efforts to improve the modelling

[4] www.multifan-cl.org

procedures to more accurately estimate stock status, fishing impacts, biological and population processes, and the characterization of uncertainty. However, they can result in changes to the estimated status of the stock and fishing impacts from previous assessments. It is important to recognise that each new assessment represents a new estimation of the historical population dynamics, impacts of fishing and stock status. Advice from the SC on previous assessments, the annual pre-assessment workshops (PAW; Hamer (2024)), emerging scientific consensus on "good practice" (e.g., the CAPAM "Tuna Stock Assessment Good Practice Workshop" held in Wellington, New Zealand in March 2023 CAPAM Tuna Good Practices) and the recommendations of the 2022 peer review of the yellowfin tuna assessment (Punt et al., 2023) that are also appropriate to south Pacific albacore guided this ongoing process.

The objectives of this assessment were to estimate population parameters, such as time series of recruitment, biomass, biomass depletion and fishing mortality, which indicated the stock status and impacts of fishing. We summarized the stock status in terms of reference points (both limit and interim target reference points) adopted by the WCPFC for routine reporting in tuna assessments. The methodology used for the assessment was based on the general approach of integrated modelling (Fournier and Archibald, 1982), which was carried out using MULTIFAN-CL, and implemented a size-based, age- and spatially-structured population model. Model parameters were estimated by maximizing an objective function, consisting of both likelihood (data) and prior information components (penalties). The assessment used an ensemble approach to estimating uncertainty for the basis for management advice. The ensemble was structured to incorporate key sources of uncertainty identified by a series of sensitivity tests along with estimation uncertainty for the key management quantities.

This assessment report should be read in conjunction with several supporting papers, listed below:

- Background analysis and data inputs for the 2024 South Pacific albacore tuna stock assessment (Teears et al., 2024)
- Update to length-weight parameters (Macdonald et al., 2024b)
- Report from the SPC Pre-assessment Workshop - March 2024 (Hamer, 2024)
- Analysis of longline size frequency data for the 2024 South Pacific albacore and WCPO striped marlin assessments (Potts et al., 2024)
- Spatial structure, movement, and regional connectivity of South Pacific albacore tuna stocks in the WCPFC-CA and EPO (Macdonald et al., 2024a; Senina et al., 2020)

3 Background

3.1 Stock structure

Albacore are distributed globally with separate stocks in the Indian, Atlantic and Pacific Oceans (Nikolic et al., 2017, 2020). In the Pacific Ocean, albacore are thought to comprise discrete stocks north and south of the equator (Nikolic et al., 2017). In the South Pacific, the stock structure is not fully resolved. Tag recapture data for releases in the southern region of the WCPFC-CA show a high level of latitudinal mixing, and provide some evidence of individual movements from the WCPFC-CA to the southern EPO (Figure 4). The longest period at liberty for a recaptured tagged albacore in the South Pacific is 11 years, and for the North Pacific 15 years (ISC Albacore Working Group, 2011). Tagging mortality using either trolling or longline as the capture method is thought to be high in albacore and, as a result, there have been limited tagging programs for this species. Therefore, albacore is not an ongoing focus of tuna tagging in the South Pacific. The earlier tagging programmes that targeted albacore in the South Pacific occurred in the early 1990s and again in 2009-2011. For both these programmes the tag releases were all in the southern temperate latitudes (south of 30°S), with most releases being immature fish (<80cm FL), and most recapture displacements occurring in a northerly direction (Figure 4). There have been no tagging studies in the southern EPO. Tagging with PSATs (pop-up satellite tags) in the south western Pacific showed the capacity for large movements, with one individual moving more than 1000 km in 50 days (Williams et al., 2015).

The south-north mixing, and the prevalence of smaller albacore in the southern region of the WCPFC-CA suggests a southern juvenile feeding/nursery area, south of 25°S, with increased northerly movements and residency with age (Farley et al., 2014; Nikolic et al., 2017). In the WCPFC-CA, spawning and larval stages mostly occur where surface water temperatures (SST) are >24°C, and typically north of 25°S (Farley et al., 2014; Nikolic et al., 2017). In the EPO, size composition data also show a pattern of smaller juveniles occurring where SST is <24°C, typically closer to South America (see appendix E in Vidal et al. (2021); Potts et al. (2024)). Although spawning areas in the EPO are not documented, similar to the WCPFC-CA, larger fish appear to be more abundant in catches where SST is >24°C (Vidal et al., 2021). The level of mixing between the WCPFC and IATTC convention areas is poorly known, but based on the results of the high-resolution spatial model SEAPODYM (Senina et al., 2020) appears to be limited. .

Significant spatial variation in growth and observed length-at-age with longitude has been observed, which suggests a more complex spatial structure with longitude (Williams et al., 2012). Otolith chemistry studies have suggested that fish caught near French Polynesia in the central South Pacific originated from a separate larval source than fish caught further west (Macdonald et al., 2013). Similarly, for five sample groups from the western, central and southern Pacific, Anderson et al. (2019); Macdonald et al. (2024a) provided evidence of population structuring at adaptive loci, particularly a differentiation of samples from French Polynesia. Recent work by Macdonald et al.

(2024a) found that the populations around French Polynesia are not mixing extensively with the populations around New Caledonia. This is consistent with the EPO region being treated as a separate model region with low mixing with the WCPO. Further studies are needed to resolve the level of mixing between the WCPFC-CA and EPO regions to better inform assessments at the scale of the entire South Pacific.

Longline catch data indicate that adult albacore appear to migrate seasonally between tropical and sub-tropical waters (Langley, 2004; Nikolic et al., 2017). These data suggest that albacore in the southern hemisphere are most abundant in sub-equatorial waters during December-January and May-July, indicating that albacore migrate south during early summer, and north during winter. This movement tends to correspond with the seasonal shift in the 23–28°C sea surface temperature isotherm.

Understanding the potential extent of mixing across the South Pacific is also informed by recent developments in quantitative modelling of the spatial dynamics of South Pacific albacore across life history stages using the Spatial Ecosystem and Population Dynamics Model (SEAPODYM) (Lehodey et al., 2015; Senina et al., 2020). The SEAPODYM modelling framework is highly spatially resolved and provides predictions of spatio-temporal exchange of biomass by age class (by months), forced by environmental/habitat variables. SEAPODYM can potentially be used to predict the exchange rates among model regions to inform transfer rates in the stock assessment model. The application of SEAPODYM outputs to inform recruitment distribution and movement between model regions was discussed at the Pre-assessment Workshop (Hamer, 2024) and was utilized to inform recruitment distribution and movement between the WCPO and EPO model regions in the 2024 assessment.

3.2 Vertical movement behaviour

An important aspect of albacore behaviour for understanding their ecology and how they interact with fishing gear is vertical movement. For example, vessels that target albacore tuna at warmer latitudes typically set longlines at depths between 100 and 400m (Bigelow et al., 2006), while those at cooler latitudes typically set at shallower depths of around 100 m, and the troll fishery targeting smaller fish around New Zealand (NZ) typically fishes at <10m depth (Williams et al., 2015). Williams et al. (2015) used PSAT tags to study vertical movement behaviour of individual albacore tagged near New Caledonia, Tonga and NZ (tagged fish were 89–107cm fork length, FL). They found that for the tropical latitudes (i.e., New Caledonia and Tonga), albacore tuna showed a distinct diurnal pattern in vertical habitat use, occupying shallower, warmer waters above the mixed layer depth (MLD) at night (e.g., <150m deep), and deeper, cooler waters below the MLD during the day (e.g., 200–300m deep). However, there was little evidence of a diurnal pattern of vertical migration behaviour in albacore at temperate latitudes (NZ), with fish staying in shallow waters above the MLD (e.g., <150m deep) almost all of the time. This latitudinal variation in vertical migration was reflected in diets with the surface-dwelling fish near NZ consuming a low diversity

of prey consisting primarily of crustaceans, but albacore in the more tropical regions consuming primarily fish, with significantly more deep-water species and a greater diversity of prey species (Matsubara et al., 2024).

3.3 Biology

3.3.1 Reproductive biology

Migratory behaviours in the South Pacific have not yet been well defined however, adults are frequently caught above 30°S where albacore have a higher mean length (Potts et al., 2024). Albacore spawn in these tropical and sub-tropical waters between 5-25°S during the austral summer with peak spawning occurring from October through December (Nikolic et al., 2017; Farley et al., 2014). In the South Pacific, Farley et al. (2013b) and Farley et al. (2014) estimated ages at 50% and 100% maturity of 4.5 and 7 years respectively. The minimum reported size of mature females was 74cm FL with the length at 50% maturity of approximately 85cm FL, and the length at 100% maturity was 94cm FL. Juveniles are caught in surface fisheries in NZ coastal waters, and in the vicinity of the sub-tropical convergence zone (STCZ, at about 40°S and 130°W) in the South Pacific, at about one year old and at a size of 45–50cm FL (Hoyle and Davies, 2009). The troll fisheries in this southern region typically catch albacore aged 1–3 yr.

3.3.2 Growth

Daily otolith growth increments indicate that initial growth is rapid (Renck et al., 2014; Farley et al., 2021), with albacore reaching 45–50cm FL in their first year (Williams et al., 2012; Farley et al., 2021). Subsequent growth is slower, at approximately 12cm per year from 2 to 4 years age, and declining thereafter (Williams et al., 2012; Farley et al., 2021). Maximum recorded length is about 120cm FL and sex-combined von Bertalanffy (VB) growth models for both the South and North Pacific albacore predict an L_∞ of around 105cm (Williams et al., 2012; Farley et al., 2013a, 2014). Maximum age is around 15 years for males and females (Williams et al., 2012; Farley et al., 2021).

Recent analyses of age-at-length from otolith data have identified important patterns in South Pacific albacore growth (Williams et al., 2012; Farley et al., 2021). Males grow to larger sizes than females, and their lengths-at-age begin to diverge above 85cm FL, when they reach maturity. Lengths-at-age of both sexes also appear to vary with longitude, with growth rates and maximum sizes increasing toward the east. In the NZ troll fishery, there are clear modes separated by about 10cm in the length frequency data for juveniles between 50 and 80cm. These modes are likely to represent annual age classes, based on the seasonal spawning cycle peaking in January (Farley et al., 2014). Farley et al. (2021) re-analysed otoliths from Farley et al. (2013b) and applied a new decimal age algorithm (developed for WCPFC-CA Pacific bigeye and yellowfin) to obtain updated decimal age estimates for albacore. The updated growth estimates based on these data were not used in the 2024 assessment however, the otolith integer ages were incorporated in the assessment

to inform internal growth estimation.

3.3.3 Natural mortality

The instantaneous natural mortality (M) rate is thought to be between 0.2 and 0.5 per year, with significant numbers of fish living 10 years or more. The default M of 0.4 used in previous assessments was updated in 2015 to 0.3 to match that used in other stocks, including the North Pacific. In 2016, a meta-analysis of mortality for the North Pacific stock indicated M should be closer to 0.39, for albacore age 6 and above (Kinney and Teo, 2016). The value of M is an important biological uncertainty that is treated in detail in the current assessment.

3.4 Fisheries

Distant-water longline fleets from Japan, Korea, Chinese Taipei, and China, and the domestic longline fleets from a number of Pacific Island Countries and Territories (PICTs), catch albacore over a large area of the South Pacific (Figure 5, Figure 6). Most of the catch is taken by longline. Significant catches began in 1954, when Japan began to expand the range of their tuna fleets post World War 2. Smaller albacore have been targeted by a troll fishery around NZ since the 1960s, the United States troll fishery (with Canadian-flagged vessels also participating in some years) operating further east, and a small fishery occurring along the east coast of Australia. Catches from the troll fishery are relatively small, generally less than 10,000 mt per year. The Chinese Taipei fleet, in particular, has targeted albacore consistently since the 1960s (Figure 6).

Since the mid-1990s, longline catch and effort (Figure 7) has increased considerably with the development and expansion of longline fisheries targeting albacore in several Pacific Island EEZs, notably those of American Samoa, Cook Islands, Fiji, French Polynesia, New Caledonia, Samoa, Solomon Islands, Tonga, and Vanuatu. Driftnet vessels from Japan, Korea, and Chinese Taipei also targeted albacore in the central Tasman Sea and in the central Pacific near the sub-tropical convergence zone for a short period during the 1980s and early 1990s (Figure 5). Driftnet catch reached approximately 22,000 mt in 1989, but rapidly declined to zero following a United Nations moratorium on industrial-scale drift-netting.

The areas of highest albacore catch rates have typically been in waters between 10–35°S in the WCPFC-CA. The spatial pattern of CPUE has been relatively consistent over time, although overall CPUE has reduced (Figure 7, Figure 8).

Longline fisheries operate throughout the year, although there is a strong seasonal trend in the catch distribution by latitude, with the fishery operating in southern latitudes (south of 35°S) during late summer and autumn, moving northwards during winter. Surface troll fisheries around NZ are highly seasonal, occurring mainly from December-April.

4 Data compilation

Data used in this South Pacific albacore assessment consist of fishery-specific catch, effort, length-frequency (LF), and age-length data. Details of these data and their stratification are described below. In the 2021 assessment, tag release-recapture data were included however, these data were not utilized in the current assessment due to large uncertainties around tagging mortality and tag reporting rates. A summary of the available data is in Table 1.

4.1 Spatial stratification

The 2021 assessment was the first attempt at a spatially structured assessment of albacore across the entire South Pacific encompassing the southern hemisphere area of the WCPFC-CA and the EPO under the management jurisdictions of the WCPFC and the IATTC (Castillo Jordan et al., 2021). The 2012 assessment (Hoyle et al., 2012) considered the entire South Pacific as a single region model with the fisheries stratified into six areas (i.e., an areas-as-fleets approach). The 2015 assessment (Harley et al., 2015) introduced a fully spatially structured model for the WCPFC-CA south of the equator with eight regions, however, there were difficulties with the complexity of this structure and the spatial structure was therefore simplified to five regions for the 2018 assessment focussing on the WCPFC-CA (Tremblay-Boyer et al., 2018).

The 2021 spatial structure was defined by four regions; three regions in the WCPFC-CA and one region in the EPO with sub-regions to delineate areas of overlap between the WCPFC-CA and EPO (Figure 1). An areas-as-fleets approach was applied in the current assessment. This necessitated changes to the spatial structure as well as the fishery definitions for both the extraction and index fisheries. To develop the spatial structure and fishery definitions for the current assessment a range of information was considered including: the previous assessment structure, tagging data, size composition data, genetics research (Anderson et al., 2019; Macdonald et al., 2024a), review of biology, fisheries, and management (Nikolic et al., 2017), modelling of spatial dynamics (SEAPODYM; Senina et al. (2020)), fishery structural regression tree analyses (Potts et al., 2024; Teears et al., 2024), fishery coverage by fleets/gears; and management jurisdictional boundaries.

The spatial boundaries for the current assessment model span from 0° to 50°S and from 140°E to the western coast of South America (approximately 70°W). The regional structure for the WCPFC-CA includes a single region and the overlap between the WCPO and the EPO was included as part of the WCPFC-CA (Figure 2). The latitudinal boundaries for fisheries defintions in the WCPFC-CA and EPO were considered on the basis of biological hypotheses of seasonal movement, spatial structuring of the population by age, and patterns of fishing activity.

4.2 Temporal stratification

The time period covered by this assessment was the first quarter of 1954 to the 3rd quarter of 2022. The 4^{th} quarter of 2022 was not included because the definition of "year" in the population

dynamics is Oct–Sep, in line with the assumption of recruitment occurring in Oct. Within this overall period, data were compiled by quarter for longline and driftnet fisheries and by month for troll fisheries. Expansion of the time-series were a deviation from the 2021 assessment and better met the assumption of an unexploited population at the beginning of the modelling period.

4.3 Fisheries definitions

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

The simplification of the spatial regions and the application of an areas-as-fleets approach, resulted in significant changes to the fishery definitions applied in the 2021 assessment. Within the two model regions, fisheries areas were defined as a six area fleet structure to define the WCPFC-CA fisheries and a three area fleet structure to define the EPO fisheries (Figure 3).

The boundary at 10°S in the WCPFC-CA was established to facilitate the exploration of management options (i.e., stock projections that apply different fishery management options) and also to ensure compatibility with future mixed fishery strategies in MSE modelling. Similar to the approach taken in 2021, the fleets associated with each gear sector and each region were further disaggregated based on vessel flag to improve the model fit to catch and size composition data due to differing selectivities. Flags[5] were grouped into three main groups: distant water fishing nations (DWFNs), Pacific island countries and territories (PICTs) and Australia and NZ (AU/NZ).

4.3.1 Extraction fisheries

The fishery definitions for the 2024 assessment are detailed in Table 2. In summary there are 17 extraction fisheries, that include groupings of longline fisheries for the DWFNs, PICTs, and AU/NZ, along with two troll fisheries and one driftnet fishery in the WCPFC-CA and one troll fishery in the EPO. The extraction fisheries are dominated by the longline sector (13 out of 17), as it is the dominant gear sector for South Pacific albacore. The extraction fisheries are stratified according to the model regions and spatial sub-region in Figure 3, resulting in 13 fisheries for the WCPFC-CA and 4 fisheries for the EPO (IATTC) region.

4.4 Catch-conditioned approach

In previous MULTIFAN-CL assessments of albacore, catch was predicted by the model (termed a "catch-errors" method) with observation error allowed, and the standard deviation of the log-catch deviates assumed to be very small (equivalent to a CV of 0.002). This produced accurate predictions of observed catches and therefore only a small contribution of the catch to the overall objective

[5]Flags defined as where flag of chartered vessels is considered to be that of the chartering nation, which is consistent with the attribution of catch histories under WCPFC management measures

function. However, the cost of treating the catch in this way was that effort deviation coefficients had to be estimated as model parameters for each catch observation. While these parameters were constrained by penalties and estimation was feasible, it resulted in very large numbers of parameters needing to be estimated by the function minimiser with many of these being effort deviation coefficients and parameters relating to catchability.

In order to reduce the complexity and number of parameters in the estimation, the "catch-conditioned" approach was developed (Davies et al., 2022) and has been previously applied in the most recent skipjack, yellowfin, and bigeye assessments (Castillo Jordan et al., 2022; Day et al., 2023; Magnusson et al., 2023). The catch-conditioned approach allows for the exact solution of the catch equation for fishing mortality by using a Newton-Raphson sub-iterative procedure. The most significant outcome of this method is the removal of the effort deviation coefficients and catchability parameters from the estimation procedure. This improves convergence diagnostics, in particular the achievement of positive definite Hessian matrices (an important convergence diagnostic) due to the simplification of the variance-covariance matrix to one with significantly fewer parameters. This approach also reduces the run times for estimation and model convergence, allowing for more exploration of various model configurations. Additionally, effort data no longer need to be included but can be input for the purpose of estimating the fishing mortality and effort relationship (via regression) for making stock projections.

4.5 Catch data

Catch data were compiled according to the fisheries defined in Table 2. See *Appendix 1: Catch and length frequency data summaries by fishery* in Teears et al. (2024) for detailed plots. All catches were expressed in numbers of fish, with the exception of the troll and driftnet fisheries, where catches were expressed in weight (metric tonnes). As mentioned in Catch-conditioned approach, effort data are not a necessary input using the catch-conditioned approach and were not included in the MULTIFAN-CL input summary.

Annual catches by flag for the South Pacific, and the WCPFC-CA and EPO regions are provided in Figure 6 and historically by gear across model regions in Figure 5. Catch for the entire South Pacific has generally been between 80,000–100,000 mt since 2009, after an increasing trend from 1990. The majority of the catch has always been from the WCPFC-CA, although the catch from the EPO has been increasing since 2000, most notably since 2010. Catches in the EPO peaked at slightly above 20,000mt in 2014, largely due to increased catches by the Chinese fleet since 2010. The Pacific-wide catch is almost all taken by longline, except for the troll catches primarily around NZ (Figure 5). Driftnet catches only occurred over a few years in the late 1980s and early 1990s.

Most of the recent catch and effort has occurred in the area almost exclusively comprised of EEZs between 10-25°S in the WCPFC-CA (Figure 7). Over the last two decades, catch and effort has increased in the south-central Pacific extending into the overlap and EPO region

4.5.1 CPUE

Given the dynamic and patchy effort of longline fleets in the South Pacific over time, there are distinct advantages to developing index fisheries from a combined-fleets dataset. This makes maximum use of the fully integrated, multi-fleet standardised CPUE analyses by providing the best possible spatial and temporal coverage for the indices of relative abundance in the assessment, and avoids assigning the multi-fleet standardised CPUE time series to only one fleet component within the assessment. The multi-fleet regional abundance indices were calculated using the sdmTMB package for implementing spatial delta-generalized linear mixed models (Anderson et al., 2022) and is discussed further below and in the supporting paper by Teears et al. (2024).

We defined two index fisheries for the WCPFC-CA and one index fishery for the EPO for which catchability was assumed constant across years. The index fisheries in the WCPFC-CA are characterized by the longline fishery from 0°–25°S and the NZ troll fishery. The index fishery in the EPO was characterized by the longline fishery over the full region 2. The effort for each time step was adjusted such that the original standardised CPUE from the supporting analysis (Teears et al., 2024) was preserved. However, a change from the previous assessment was the application of time varying CVs for the index fisheries. This addresses the concern that changes in longline fleet targeting, gear configuration and material, and fleet dynamics could result in temporal variation in uncertainty that would be important to capture within the estimation framework. There was considerable overlap in the fleet composition for the longline index fisheries, but there were also some differences. Data for each of the longline index fisheries spanned the full time series, 1954-2022 however; the NZ troll index spanned the years from 1992-2022 (Neubauer and Hill-Moana, 2024).

In the previous two assessments, indices of relative abundance based on longline data were developed using a spatio-temporal modelling approach implemented in the VAST R package (Thorson et al., 2015). The current assessment uses a similar spatio-temporal modelling approach implemented in the sdmTMB R package (Anderson et al., 2022) as was done in 2023 for the estimation of bigeye and yellowfin tuna (Teears et al., 2023). The sdmTMB geostatistical software was selected as it has been developed to be computationally efficient, flexible, and user-friendly with online community support (Anderson et al., 2022) and thus, represents a reasonable alternative for improving reproducibility and efficiency. The modelling process is described in detail in Teears et al. (2024).

The sdmTMB framework was used to implement a spatio-temporal delta generalized linear mixed model (GLMM), from which area-weighted abundance indices were generated after "standardizing out" the influence of the catchability covariates. The modelling approach explicitly addresses the spatial structure in the response variable, that is, the fact that observations closer in space are more likely to be similar. This allows the spatial autocorrelation to be accounted for, which increases the precision in estimates and in some instances makes it easier to identify a relationship between the response and candidate explanatory variables.

In the 2021 assessment, the CPUE standardization model used targeting cluster as a catchabil-

ity covariate, which was derived using longline catch species composition (i.e., yellowfin, bigeye, albacore, and other species) and hooks between floats (HBF). This approach follows the assumption that changes in species‘ composition of longline sets represent changes in targeting behaviour. However, species composition may be confounded with changes in abundance of each species. Additionally, vessel flag was predicted to be an important covariate as there can be fleet effects and differences in fishing strategies employed. Therefore, the final model used to generate the CPUE indices included HBF in place of targeting cluster and vessel flag as catchability covariates. Testing of standardization models with HBF gave similar results to those with targeting cluster.

In the WCPFC-CA, separate indices were developed (including all flags) in the northern (north of 25°S) and southern (south of 25°S) areas of the region to index adults and juveniles, respectively. In the southern area, NZ troll fishery CPUE data were standardised (along with corresponding size compositions) to index relative abundance (Neubauer and Hill-Moana, 2024). In the northern area, an annual index was developed that included HBF, flag, and season (i.e., month) as a cyclic spline; all of which were treated as catchability covariates. Season was considered a catchability covariate based on the seasonal dynamics of the fishery driven by fluctuating local availability and market demand (R. Dunham, Tri Marine Group, personal communication, March 13, 2024) for the main longline species (i.e., albacore, bigeye, swordfish, and yellowfin).

In the EPO, a single index fishery was developed for the whole region. In the 2021 assessment, only the Japanese data were included in the standardization of the EPO index. However, in addition to changes in apparent targeting among the Japanese fleet over time, there were multiple flags over more recent years ($\geq$2000; e.g. China and Chinese-Taipei) that indicated differing signals in relative abundance (see Teears et al. (2024) for further details) which led to the decision to include other flags within the analysis as well.

The MULTIFAN-CL model was configured to allow for the time-varying nature of the CVs such that the fit to the CPUE data was given greater influence in the likelihood in time-steps with more precise estimates of relative abundance. However, in contrast to the previous assessment, regional weights were not informed by the CPUE standardization model, which were input as mean-centred in the data input file as it was considered an improvement to allow the recruitment distribution to provide information on regional scaling since these were fixed at estimates from the SEAPODYM model.

4.6 Size compositions and length data treatment

Available LF data for each of the defined fisheries were compiled into 50, 2-cm size classes (30–128cm). This was a change from the 2021 assessment that used 100, 1-cm size classes (30–129cm). An investigation into the length composition data indicated frequent misreporting of bin size resolution in data submissions. Briefly, there were an inordinate number of frequencies in even numbered bins, odd numbered bins, or, to a much lesser extent, length bins that were multiples of five. Subsequently, efforts to identify and remove “contaminated” length compositions at the trip level were performed

to improve the quality of the data while preserving as much of the "uncontaminated" data as possible (see Teears et al. 2024 for further details). However, the change from 1-cm to 2-cm length bins ultimately allowed more data to be included, particularly for the longline fisheries. Data were collected from a number of sources and can be summarized as follows.

4.6.1 Troll and other fisheries

NZ domestic troll size composition data (sub-region 1-e) were collected from port sampling programmes conducted by the NZ Ministry of Fisheries and, more recently, the NZ National Institute of Water and Atmospheric Research (NIWA).

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

Driftnet data were provided by the National Research Institute of Far Seas Fisheries (NRIFSF) for Japanese driftnet vessels. Data from Japanese vessels were also collected by observers and by port sampling in Nouméa, New Caledonia. It was assumed that these data are representative of all driftnet activity.

4.7 Re-weighting size compositions

Statistical correction of size composition data is required as length and weight samples are often collected unevenly in space and time. The methods for re-weighting of the size composition data are detailed in Teears et al. (2024) and are based on those developed by McKechnie (2014) for longline extraction fisheries, and Tremblay-Boyer et al. (2018) for longline index fisheries. For the extraction fisheries, re-weighting of composition data is required to ensure that sampling biases in space, time, and the fleets providing data, are minimised so that size composition data better reflect the composition of the overall removals. Strata-specific size data samples were therefore re-weighted by catch for the extraction fisheries. For the index fisheries, re-weighting of composition data is required to ensure that the size composition of the abundance indices reflect the size component of the population that is being sampled by the index fisheries through space and time. Strata-specific samples are therefore re-weighted by relative abundance using the CPUE. Given that the same composition data were used for both the extraction and index fisheries, the observed number of size-frequency samples input into the assessment was divided by two for both the extraction and index fisheries where these are the same fisheries.

4.8 Age-at-length

As in the previous assessment, age-length data from otolith readings were utilized from the ageing study of South Pacific albacore by Farley et al. (2021). The work used the previously collected and

prepared otoliths (collected in 2009-2010; Farley et al. 2013a) with the addition of new data on daily ages from 60 small fish (43–47cm FL) sampled from NZ waters and the re-reading and measuring of increment zones on 600 previously collected 'high confidence' otoliths from the WCPFC-CA. The reanalysis applied an age algorithm (developed for WCPFC-CA bigeye and yellowfin; Farley et al. 2020) to obtain updated decimal age estimates for albacore that were applied in the assessment.

Age-at-length were input as conditional-ages-at-length as this method relaxes the assumption that the data were representative of the entire age range and as recommended by the 2022 peer review of the yellowfin assessment (Punt et al., 2023). In all, 654 age-at-length samples were aggregated into 33 fishery and sampling year/quarter strata.

5 Model description

5.1 General characteristics

The model comprises of several components, (i) the dynamics of the fish population; (ii) the fishery dynamics; (iii) the observation models for the data; (iv) the parameter estimation procedure; (v) the uncertainty estimation procedure (both parameter and model uncertainty); and (vi) stock assessment interpretations. Detailed technical descriptions of components (i)–(iv) are given in Hampton and Fournier (2001) and Kleiber et al. (2019). In addition, we describe the procedures followed for estimating the parameters of the model, the uncertainty, and the way in which stock assessment conclusions are drawn using a series of reference points. In this section, model settings primarily refer to those used within the "diagnostic case" model. Some of these settings are later varied in sensitivity analyses.

5.2 Population dynamics

The model partitions the population into two spatial regions (under the 2024 regional structure) and 12 annual 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 annual time steps, extending through a time window of 1954–2022. The main population dynamics processes are as follows.

5.2.1 Temporal structure

The estimation model was configured to estimate population dynamics on an annual basis with corresponding annual recruitment. However, catches were extracted from the population at specified quarters or months, and within-year natural mortality applied to correctly implement the catch equations. The beginning of the year was initialized at week 37, which corresponds to the mid-point of the third quarter. This was based on a mean length of the first age class of ∼45cm (i.e., age 3 quarters) and a spawning season mid-point in January (Farley et al., 2013b). This temporal

structure was a simplification of the 2021 assessment, which was at a quarterly time-step beginning in the first week of the calendar year.

5.2.2 Recruitment

Recruitment was defined as the appearance of age-class 1 (year) fish in the population. In contrast to the tropical tunas, spawning of South Pacific albacore occurs during the Austral summer. It was assumed that recruitment occurs instantaneously at the beginning of each model year (i.e., week 37 of the calendar year). Spatially aggregated (over all model regions) recruitment was assumed to have a weak relationship with spawning potential via a Beverton and Holt stock-recruitment relationship (SRR) with a fixed value of steepness (h). The h parameter 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 for the h parameter in the SRR (ISSF, 2011). In the previous assessment, values of 0.65, 0.80 and 0.95 were included in the uncertainty grid. In the current assessment, a Monte Carlo model ensemble approach was adopted where h values (along with average M-at-age) were sampled from an assumed prior distribution (see Monte–Carlo model ensemble uncertainty estimation methods for further details).

The SRR was incorporated mainly so that yield analyses, equilibrium– and depletion–based reference points and population projections could be undertaken for stock assessment purposes. As was done in the previous assessment, a penalty (equivalent to a CV of 0.71) was applied to recruitment deviations from the SRR so that it would have a minor effect on the recruitment estimates and other model estimates (Hampton and Fournier, 2001), but still allow the estimation of asymptotic recruitment. This approach was recommended by the review of the 2011 bigeye stock assessment (Ianelli et al., 2012). Recruitment deviations were estimated on the log scale for the full model period, excluding the terminal annual recruitment (which was not freely estimated and was set equal to the arithmetic mean of the estimated recruitments).

The regional recruitment distribution was configured to use fixed proportions based on the average distribution obtained from the SEAPODYM model (Senina et al., 2020). The distribution was assumed to be time invariant at 0.819 and 0.181 in regions 1 and 2, respectively.

5.2.3 Initial population

The population age structure in the initial time period in each region was assumed to arise from an equilibrium unexploited population. This assumption was consistent with the catch history and avoids having to estimate the initial age structure, which is generally poorly determined, as independent parameters in the model.

5.2.4 Length-weight relationship

Since the 2005 assessment (Langley and Hampton, 2005), length-weight relationship estimates were derived from:

$$WW = a * FL^b$$

where WW was whole weight, FL was fork length, $a = 6.9587e^{-6}$, and $b = 3.2351$ as estimated from available length-weight data collected.

A significant quantity of new length-at-weight data have become available since then and has been analysed along with previously available data for updating the length-weight parameters as input in the current stock assessment (Macdonald et al., 2024b). The updated length-weight parameters are $a = 1.708e^{-5}$ and $b = 3.0483$.

5.2.5 Growth

The standard growth assumptions applied in MULTIFAN-CL concerning age and growth are: i) the lengths-at-age are normally distributed for each age-class; ii) the standard deviations of length for each age-class were a log-linear function of the mean lengths-at-age; and 3) the probability distributions of weights-at-age were a deterministic function of the lengths-at-age and the specified weight-length relationship. These processes are assumed to be spatially and temporally invariant.

This assessment included two predominant data sources that are potentially informative regarding growth – the length frequency data and the conditional age-at-length data. In respect of the LF data, the NZ troll data in particular demonstrate periods of strong modal structure that almost certainly represent annual age classes (Figure 9). We initially employed a fixed growth parameterisation informed by an external analysis of the otolith-based length-age data (Farley et al., 2021). This analysis indicated that a logistic model provided a superior fit to the data; however this growth formulation is not currently available in MULTIFAN-CL. For initial model exploration, we therefore used a VB model fitted to the same data set, providing parameter estimates of $L_\infty = 110.2$, $k = 0.268yr^1$ and $a_0 = -1.24yr$ (J. Farley, pers.comm., 31/03/2021). However, we found that this model did not provide adequate fits to the LF data, and in particular was not able to represent the modal structure shown in Figure 9 well.

On the advice of Punt et al. (2023), we then incorporated the otolith ageing data into the assessment model in a conditional age-at-length (CAAL) structure, as described in Conditional-age-at-length, and estimated growth parameters internally. However, it was clear that the estimated VB model was still not capable of adequately fitting either the LF data or the CAAL data.

We re-examined the original growth models estimated by Farley et al. (2021) and noted that the preferred logistic model demonstrated approximately linear growth to around age 4, which is a key departure from VB growth. Anticipating that this was responsible for the lack of fit to both the LF and CAAL data that was problematic with the VB model, we then extended the VB model

to include offsets from the VB growth pattern for age classes 2, 3 and 4 years. This involved the estimation of three additional parameters, but we found that the VB-with-offsets model was able to adequately replicate the linear growth for the young age classes and provide better fits to both the LF data (and in particular to the NZ troll data segments showing strong modal structure) and the CAAL data (see later results section). This growth model was therefore adopted for the diagnostic case (and model ensemble). This resulted in a growth model with eight parameters. After experimenting with various combinations for estimated and fixed parameters, we adopted the following approach:

- The mean length of the first age class, $L1$, was fixed at 45.5cm. This value implied an age of approximately 9 months, consistent with the assumption of recruitment occurring in October and January being the approximate mid-point of the spawning season.
- The growth coefficient, k, and mean length of the oldest age class, $L2$ (age-class 12+) were estimated in the assessment model.
- Offsets from VB growth were estimated for age classes 2, 3 and 4.
- The standard deviation of length at age is determined by two parameters, a generic standard deviation (SD) parameter, which was estimated in the assessment model, and a parameter describing how the SD increases with age. This parameter was fixed at a value that kept the SD for the youngest age class to about 2.5cm, consistent with the length distribution seen in this age class in the NZ troll LF data.

5.2.6 Movement

Movement was assumed to occur instantaneously at the beginning of each quarter via age-specific movement probabilities that specify the proportion of fish in a given region that move to the adjacent region. The movement probabilities were estimated from the SEAPODYM model (Figure 10; Senina et al. 2020).

5.2.7 Natural mortality

CAPAM at the 2023 "Tuna Stock Assessment Good Practices Workshop" recommended applying an age-specific pattern in M by using the inverse mean length-at-age method developed by Lorenzen (1996). Furthermore, average M should be scaled using a maximum age (A_{max}) approach when estimating M-at-age internally is not possible, as was the case with the current assessment. For this assessment, we have adopted this approach for specifying M-at-age and have followed the Hamel and Cope (2022) method ($5.4/A_{max}$). We applied an assumed A_{max} of 15 yrs based mainly on the observations of the oldest fish in the age-at-length data set (15.16 yrs), with two other observations >14 yrs. Furthermore, bomb radiocarbon observations from other oceans suggest an A_{max} in the "mid-teens" to be appropriate for South Pacific albacore (A. Andrews, pers.comm., 07/07/2024). An A_{max} of 15 yr results in an expected value for average M of 0.36. Following the

CAPAM Workshop advice, we allowed M-at-age to be inversely proportional to mean length-at-age following Lorenzen (1996). We assumed that the value of 0.36 could be applied as a simple average M ($\overline{M}$) for the mature age classes (4–12+) in the model (Hamel and Cope, 2022; Hoyle et al., 2023) and configured the Lorenzen function to meet this assumption.

5.2.8 Reproductive potential-at-age (sexual maturity)

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 biomass (SB). Reproductive output at age, which is used to derive spawning potential, attempts to provide a measure of the relative contribution of fish at different ages to the next generation. We specified estimates of length-specific spawning potential, which is then converted internally within the model to an age-based vector dependent upon the growth model (Figure 11). Spawning potential-at-length was specified as the product of sex ratio and maturity-at-length, rescaled to a maximum of one. Sex ratio at length was obtained by fitting a spline to the observed sex-ratio from SPC-held longline observer data, stratified by flag and 10° cell and weighted by longline catch to account for uneven observer coverage amongst fleets in space. Longline observer data only covered lengths from 70–110cm so lengths beyond this range were extrapolated to cover the stock assessment range by: (1) setting sex ratio for lengths <70cm at the 70cm value and (2) assuming sex ratio for females declined linearly from the value observed at 110cm to 0 at 130cm. The maturity proportion at length was obtained from the weighted maturity ogives presented in Farley et al. (2014), smoothed via a logistic curve.

5.3 Fishery dynamics

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

5.3.1 Selectivity

Selectivity is fishery-specific and assumed to be year-invariant. As was done in the 2021 and 2018 assessments, selectivity was based on the cubic spline interpolation technique. This is a form of smoothing, but the number of parameters for each fishery is the number of cubic spline "nodes" that are deemed to be sufficient to characterise the selectivity over the age range. Cubic splines were configured with three nodes, which was sufficient to allow for reasonably complex selectivity patterns with the exception of the longline fisheries below 25°S in the WCPFC-CA (fisheries 8, 9, and 10), which were configured with four nodes.

As noted in Fisheries definitions, the division between fisheries in the WCPFC-CA above and below 10°S was designed for the purpose of exploring management options. Subsequently, fleet-specific selectivity for longline fisheries above and below 10°S were grouped (i.e., fisheries 1, 2, and

5 were grouped with fisheries 3, 4, and 6, respectively). Additionally, selectivity for troll fisheries 12 and 17 were grouped together as they were expected to share similar selectivity curves. In the previous assessment, the index fisheries were grouped to have shared selectivity, which also carried the assumption that differences in length frequency observations were due to differences in the underlying size structure of the population. However, for this assessment, this assumption was relaxed due to likely differences in targeting that were difficult to standardise out in the CPUE model due to lack of adequate gear configuration information (see Teears et al. (2024) for more information).

In the previous assessment, spatial structure and movement were explicitly modelled, and seasonal changes in size composition of catches were modelled through movement. In the current assessment, an areas-as-fleets approach was adopted within each of the WCPFC-CA and EPO regions. In order to allow for seasonal changes in size composition of catches, we allowed selectivity to vary seasonally for those fisheries that operated substantially throughout the year in the tropical sub-regions (1-a, 1-b, 1-c, 1-d and 2-a).

For the longline index fisheries, selectivity was configured to be non-decreasing for successively older age-classes.

The configuration of the diagnostic case model with respect to selectivity and other model parameters is provided in Table 3.

5.4 Likelihood components

There are three main data components that contribute to the log-likelihood function for the current assessment: the LF data, the conditional-age-at-length data, and the CPUE data.

5.4.1 Length-frequency

The probability distributions for the LF proportions are assumed to be approximated by log-normal distributions, with the variance determined by the effective sample size (ESS) and the observed LF proportion. Lower ESS values account for the fact that (i) LF 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 underestimation of variances.

In previous assessments, a down-weighting factor was arbitrarily applied to the observed sample sizes to produce the ESS (and the down-weighting factor usually considered as an uncertainty component in an uncertainty grid). In the current assessment, LF samples were assigned an ESS calculated using the data-weighting method developed by Francis (2011) and referred to hereafter as the "Francis method". The Francis method involves calculating a multiplication factor for the LF data for each fishery from a previous model run (with initial arbitrary ESS values) with the aim of matching the input variance (from the observed data) to the output variance (from the

standardised residuals). This multiplication factor is transformed into the ESS divisor values and adjusted accordingly in the input files for each fishery. Note that the ESS weighting is performed after the observed sample size has been capped at 1000 samples. The model is run again with the new weightings, and the resulting output is used to recalculate the ESS values once more. This iterative weighting process is performed multiple times if necessary until the weightings stabilize resulting in approximately matching input and output variances for each fishery. The resulting effective sample sizes were maintained for sensitivity analyses.

5.4.2 Conditional-age-at-length

The CAAL data is modelled by using the growth function to predict the age distribution by length class for each fishery-specific CAAL sample. The key advantages of this approach relative to the traditional approach of using the age-length observations to estimate a growth curve external to the assessment model are: 1) The predictions account for the age-specific selectivity of the fishery from which the samples were taken; 2) The method is consistent with the length-stratified sampling design and does not require representative sampling of ages from the population; and 3) uncertainty in the estimates of growth parameters are a part of the overall estimation uncertainty and are therefore propagated through to the stock assessment results of key interest.

The observed age composition in each length class is assumed to follow a multinomial distribution with the negative log-likelihood summed for all length classes and across all samples.

The current MULTIFAN-CL model configures the ESS for CAAL by applying a scalar to the observed sample size (OSS) for each sample. For the diagnostic case model, the ESS for all samples was set as $OSS * 0.75$. Sensitivity to this scalar was explored.

5.4.3 Index fishery CPUE

The CPUE indices were modelled as annual log-normally distributed observations. The contribution of each observation to the log likelihood is made up of a time series component, or relative CV, and a generic component, or scaled CV, that is specified for each index. We used the CPUE standardisation analysis estimates to specify both components for the longline indices (the generic components of which were approximately 0.20 for both longline indices); however, the generic CV values obtained for the troll index were considered unreasonably small ($\sim$0.02). We therefore specified the generic CV for the troll index to be 0.20 but retained the time series component informed by the troll CPUE analysis.

5.5 Parameter and uncertainty estimation

The parameters of the model (Table 3) were estimated by maximizing the log-likelihood of all data components plus the log of the probability density functions of the penalties specified in the model. The maximization to a point of model convergence was performed by an efficient optimization using exact derivatives with respect to the model parameters (auto-differentiation, Fournier

et al. (2012)). Estimation was conducted in a series of phases, the first of which used relatively arbitrary starting values for most parameters. A bash shell script, "doitall" file (see Appendix 1 for diagnostic case model doitall file), implements the phased procedure for fitting the model. After obtaining a converged model, the Hessian matrix was computed and its positive definite status verified by the absence of negative eigenvalues. Then, the estimation errors for the important stock-assessment-related dependent variables (including time series of recruitment, SB, dynamic depletion and depletion and MSY-based reference points) were computed using the Delta method.

6 Diagnostics methods

As has been suggested in Carvalho et al. (2017) and Carvalho et al. (2021), diagnostic tools are vital for evaluating the quality of integrated stock assessment models for informing management advice and there is no single diagnostic tool that is capable of comprehensive evaluation for all models. As such, a suite of tools from the "diagnostic toolbox" was applied to the current model assessment.

6.1 Convergence diagnostics

All models in this assessment were considered to be converged if 1) a maximum parameter gradient of 1e-05 was achieved; 2) the Hessian matrix was positive definite; 3) the model fit could not be significantly improved by jittering the estimated parameters (Carvalho et al., 2021); (4) estimated parameters should not be on their bounds; or (5) be highly correlated with each other ($-0.9 > r > 0.9$).

6.2 Model fit

Plots of observed and predicted index, LF and CAAL data were examined, including residuals plots. Particular attention was paid to the fits to the NZ troll LF data during model development, which indicated distinct annual age class modes representing the younger ages.

6.3 Age-Structured Production Model (ASPM)

ASPM diagnostics (Carvalho et al., 2017, 2021; Maunder and Piner, 2015; Minte-Vera et al., 2017) for the diagnostic case model were estimated by 1) fixing growth and selectivity parameters at their estimated values; 2) removing the LF and CAAL data from the model, leaving only the CPUE indices as data 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 and depletion scaling and trends estimated by the ASPM and the full model gives an indication of the extent to which these estimates are informed by the CPUE indices only. A second version of the ASPM, in which recruitment deviations were estimated, was also run.

6.4 Catch Curve Analysis (CCA)

A CCA is essentially the reverse of the ASPM, whereby the CPUE index data are removed from the model and all parameter estimation retained. The CCA indicates the information on population trends and scaling provided by the LF and CAAL data.

6.5 Likelihood profile

A likelihood profile over a metric related to population scale is frequently used to evaluate conflict among data types. In this assessment, we profiled on the estimated population biomass averaged over the full model period (1954-2022) and examined the response of the likelihood of each data type (and their components). Ideally, we would like to see the negative log likelihood of each data component used in the model reaching a minimum at similar levels of average biomass.

6.6 Retrospective analysis

Retrospective analyses were undertaken as a general test of the stability of the model. A robust model, when rerun with data for the terminal year/s sequentially excluded (Cadigan and Farrell, 2005), should produce outputs that are variable across runs, and without a systematic pattern in either the scaling or time-series trends. The Mohn's rho statistic, a measure of the average relative bias of retrospective estimates, was computed for recruitment, SB and spawning biomass depletion ($SB_t/SB_{F=0(t)}$) to indicate whether significant retrospective bias was present in the model. When calculating Mohn's ρ for recruitment, the penultimate year was used for comparison since the terminal year recruitment value was set to the arithmetic mean of the recruitment time series, thereby having the effect of reducing any potential retrospective pattern.

7 Sensitivity analysis methods

Sensitivity analyses were undertaken to determine the sensitivity of important stock assessment results (recruitment, SB, $SB_{\text{recent}}/SB_{F=0}$, $F_{\text{recent}}/F_{\text{MSY}}$ and $SB_{\text{recent}}/SB_{\text{MSY}}$; see Table 4 for reference point definitions) to various structural assumptions, parameter settings and decisions made during model development. For the purpose of these comparisons, we used the diagnostic case model as the reference model. The results of these tests informed decisions regarding the composition of the multi-model ensemble used to characterise uncertainty in the assessment results.

The range of sensitivity tests was developed by the assessment team over the course of model development, taking account of suggestions by the 2024 SPC Pre-Assessment Workshop (Hamer, 2024) as well as discussions with external scientists. The tests, hopefully, capture the main sources of potential uncertainty in the key model results given the time available.

The sensitivity tests conducted included (1) use of an alternative longline CPUE index for the WCPFC-CA region; (2) exclusion of the NZ troll CPUE index; (3) exclusion of the NZ troll LF

data; (4) alternative weighting of the CAAL data in the likelihood; (5) alternative movement probabilities between the WCPFC-CA and EPO regions; (6) several longline effort creep scenarios; (7) alternative settings for the distribution of recruitment between the WCPFC-CA and EPO regions; (8) alternative numbers of age classes in the model; (9) use of selectivity and catchability time blocks in the model; (10) use of an alternative growth model (simple VB) in the assessment; (11) alternative settings for the natural mortality rate; and (12) alternative settings for the steepness parameter in the Beverton and Holt stock-recruitment relationship.

8 Monte–Carlo model ensemble uncertainty estimation methods

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 (described above). The second involves "model" uncertainty, which is the uncertainty in the structural and fixed-parameter assumptions underpinning individual models, e.g., fixed M, h, etc.... The third involves data inputs, such as alternative abundance indices or other data inputs. Stock assessments of tuna for the WCPFC have often included an approach to assess the model uncertainty in the assessment model by running a factorial "grid" of models to explore the interactions among selected "axes of uncertainty". The grid contains all combinations of two or more parameter settings or assumptions for each uncertainty axis and this was commonly referred to as the "structural uncertainty grid".

In the current assessment, the characterization of uncertainty in management reference points and quantities of interest was accomplished by applying a Monte Carlo model ensemble approach following the methods introduced by Ducharme-Barth and Vincent (2021), implemented in the 2021 stock assessment of south-west Pacific swordfish (Ducharme-Barth et al., 2021) and recommended as good practice by Neubauer et al. (2023). Building off the familiar "model uncertainty grid", the model ensemble approach continues to consider the effects of model uncertainty while extending it to also account for the statistical estimation uncertainty from each model in the ensemble. This allows for a more holistic and transparent description of the uncertainty in estimates of stock status. Another key difference between the model ensemble and the model uncertainty grid is the relaxation of the full factorial design. Instead of choosing set levels for certain fixed parameters (e.g. $h \in 0.65$, 0.8, 0.95), a random set of the fixed parameters is drawn from an assumed prior distribution for each model in the ensemble. This approach has the advantage of implicitly weighting the ensemble to the most likely parameter combinations given the shape of the prior.

For aspects of model uncertainty that cannot be parametrized using prior assumed distributions (e.g., fitting to alternate CPUE indices), a full factorial approach could still be used and then overlaid on the parameter draws from the priors in a hybrid factorial ensemble.

The estimation uncertainty for each model in the ensemble was determined as described in Parameter and uncertainty estimation However, for the ensemble models, we computed the estimation

uncertainty (i.e., standard deviations of the estimates) for only the key reference point variables ($SB_{\text{recent}}/SB_{F=0}$, $F_{\text{recent}}/F_{\text{MSY}}$ and $SB_{\text{recent}}/SB_{\text{MSY}}$) and not the full set of time-series estimates as an efficiency measure. This was done for all models in the model ensemble and the estimation uncertainty was combined across models in a parametric bootstrap similar to the approach used in stock assessments conducted by the International Pacific Halibut Commission (Stewart and Martell, 2014) and applied in the 2023 yellowfin and bigeye assessments (Day et al., 2023; Magnusson et al., 2023).

Natural mortality and steepness emerged from the sensitivity analyses as the key sources of model uncertainty impacting stock-assessment-related estimates (see Sensitivity analyses). Therefore, prior distributions for these parameters were constructed as described below.

Diagnostics applied to the model ensemble were not as extensive as those applied to the diagnostic case model as described in Diagnostics methods due to the high number of models in the ensemble. Therefore, ensemble model were considered converged if 1) a maximum parameter gradient of 1e-4 was achieved; 2) the Hessian matrix was positive definite; 3) the negative log-likelihood of the NZ troll fishery for samples 14, 15 and 16 (as these showed distinct age-class modes) was $\leq$-130.

8.0.1 Natural mortality (M) prior distribution

As described in Section 5.2.7, $\overline{M}$ of $0.36yr^{-1}$ was specified using the A_{max} approach assuming an A_{max} of 15yr. Hamel and Cope (2022) recommended a CV of 0.31 for $\overline{M}$. However, we found that sampling from this distribution produced a significant number of low (<0.25) and high (>0.7) values of average M that were considered outside the range. Therefore, the CV was reduced to 0.2 to focus the replicates on a more plausible range of ~0.25–0.55 (Figure 12). Each value of average M sampled from the prior was converted to M-at-age using the Lorenzen approach described in Natural mortality.

8.0.2 Steepness (h) prior distribution

In order to develop the h prior, the approach of Brodziak et al. (2011) was considered, which uses various life history parameters, including M-at-age, in characterising a prior for h. With the assistance of the author, this approach was applied using South Pacific albacore reproductive parameters and M-at-age from the diagnostic case model. The mode of the resulting distribution was at 0.99 and initial runs of the Monte-Carlo ensemble model approach were generating h values of $\geq$0.99 in ~20% of the replicates. This distribution also implied a non-trivial probability of very low h values, which represented a significant change from the previous approach of using 0.65, 0.80 and 0.95 as discrete values in a factorial grid with equal probability.

The MULTIFAN-CL model frequently did not converge due to inadequately fitting the SRR to the SB and recruitment estimates when using the very high values of h. Furthermore, there were several replicates for which h had been sampled to be very low, <0.5, which were producing what

were judged to be unreasonable estimates of population dynamics in the MULTIFAN-CL models. To proceed, approximately ∼30% of the ensemble models would be rejected, effectively truncating the h prior on the low and very high ends. Subsequently, a modification to the prior for h was made so that such high values >0.99 and low values <0.5 had much lower probability than as indicated in the distribution described above. The approach that was therefore adopted was similar to one used by Ducharme-Barth et al. (2021). Firstly, we wanted to recognise the application of Brodziak et al. (2011) that incorporated life-history criteria and indicated that h, on average, was likely to be considerably higher than what had been assumed in previous assessments. Secondly, we also wanted to respect the approach used in the previous assessment, including meta-analyses, that indicated a reasonable range of h was likely to be ∼0.65–0.95, as applied in the factorial grid in 2021. Using this rationale, a censored (0.2–1) beta prior with a mean of 0.87 and $\sigma = 0.063$ was developed resulting in the prior distribution shown in Figure 12.

9 Stock assessment interpretation methods

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

9.1 Yield analysis

The yield analysis consists of computing equilibrium catch (or yield) and SB, conditional on a specified basal level of age-specific fishing mortality (F_a) for the entire model domain and determined as the mean over a recent period of time (2018–2021), a series of fishing mortality multipliers (f_{mult}), the M-at-age, the mean weight-at-age (W_a) and the SRR parameters. All of these parameters, apart from f_{mult}, which is arbitrarily specified over a range of 0–50 (in increments of 0.1), are available from the parameter estimates (or specifications) of the model. The maximum yield with respect to f_{mult} can be determined using the formula given in Kleiber et al. (2019), and is equivalent to the maximum sustainable yield (MSY). The reciprocal of the F_{mult} resulting in the MSY is equivalent to $F_{\text{recent}}/F_{\text{MSY}}$. Similarly, the SB at MSY (SB_{MSY}) can also be determined. The ratios of the current (or recent average) levels of fishing mortality and SB to their respective levels at MSY are determined for all models of interest, including those in the model ensemble. Note that in this case MSY quantities can only be estimated at the scale of the entire model domain, and not for individual regions, or in this case, separately for the WCPFC-CA and EPO.

9.2 Depletion and fishery impact

Many assessments estimate the ratio of recent to equilibrium biomass (usually SB) 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 general high uncertainty with early time series

data, then the ratio may not measure fishery depletion reliably.

We assess fishery depletion by computing the unexploited SB time-series using the estimated model parameters, but assuming that fishing mortality was zero. Because both the estimated SB_t (with fishing), and the unexploited $SB_{F=0(t)}$, incorporate recruitment variability, their ratio at each time step (t) of the analysis, $SB_t/SB_{F=0(t)}$, can be interpreted as an index of fishery depletion. The computation of unexploited SB includes an adjustment in recruitment to acknowledge the possibility of a reduction in recruitment for exploited populations through stock-recruitment effects. To achieve this, the estimated recruitment deviations are multiplied by the ratio of the unexploited equilibrium recruitment and the equilibrium recruitment as predicted by the Beverton and Holt SRR at the year-specific level of spawning biomass.

A fishery impact analysis was used to estimate depletion associated with specific fisheries or groups of fisheries. Here, fishery groups of interest, tropical longline (sub-regions 1-ab), sub-tropical longline (sub-regions 1-cd), southern longline (sub-regions 1-ef), EPO fisheries (sub-regions 2-abc), troll and driftnet fisheries, are removed in-turn in separate simulations. The changes in depletion observed in these runs are then indicative of the depletion caused by each of the removed fisheries.

9.3 Reference points

The $SB_{F=0}$ 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. The WCPFC adopted $20\%SB_{F=0}$ as a limit reference point (LRP) for the albacore stock (and for all other key tuna), where $SB_{F=0}$ is calculated for this assessment ending in 2022 as the average over the period 2012–2021. The WCPFC also adopted an interim target reference point (iTRP) for South Pacific albacore specified as 0.96 $SB_{2017-2019}/SB_{F=0}$. This was computed as 0.96 times the mean of the following three ratios: $SB_{2017}/SB_{F=0,2007-2016}$, $SB_{2018}/SB_{F=0,2008-2017}$ and $SB_{2019}/SB_{F=0,2009-2018}$. Stock status was referenced against the iTRP by calculating $SB_{recent}/SB_{F=0,2012-2021}$ and $SB_{latest}/SB_{F=0,2012-2021}$, where "recent" is 2019-2022 and "latest" is 2022, and expressing these as ratios to the iTRP (Table 4). This was performed for the 100 models in the ensemble with the median value and 10 and 90 percentiles taken to define the central tendency of the stock status estimates and its uncertainty in relation to the iTRP.

The other key reference point, $F_{\text{recent}}/F_{\text{MSY}}$ (Table 4), is the estimated average fishing mortality over the full assessment area over a recent period of time (F_{recent}; 2018–2021 for this stock assessment) divided by the fishing mortality producing MSY (as produced by the yield analysis and detailed in Yield analysis)

9.4 Majuro and Kobe plots

For the standard yield analysis (Yield analysis), F_a, is determined as the average over some recent period of time (2018–2021 herein). In addition to this approach, the MSY-based reference points

($F_t = F_{MSY}$ and $SB_t = SB_{MSY}$) were also computed for each year included in the diagnostic case model (1954–2022, with no value calculated for the terminal year) by repeating the yield analysis for each year in turn. This enabled temporal trends in the reference point variables to be estimated taking account of the differences in MSY levels under varying historical patterns of age-specific exploitation. This analysis is presented in the form of dynamic Kobe and Majuro plots, which have been presented for all recent WCPFC stock assessments.

10 Results

10.1 CPUE trends

In both the WCPFC-CA and EPO, the standardised trends in relative abundance based on longline data showed an overall decline from the early part of the time series to relatively stable trends since approximately the 1980s with recent estimates of relative abundance remaining below the long-term mean (Figure 13). The WCPFC-CA index indicated an increasing trend prior to 1960 when the EPO index indicated a decline. For the NZ troll data, the index is more variable at the beginning of the time-series (from 1992) however, the overall trend is relatively stable throughout.

10.2 Consequences of key model developments

Aspects of the progression of model development from the 2021 reference case to the model used as the diagnostic case in 2024 are described below with brief notes on the implication of the developments for SB and $SB_{\text{recent}}/SB_{F=0}$, which are also displayed in Figure 14. We note that this is not intended as a full detailed stepwise analysis given our intention from the start was to go back to a much a simpler model, that would be very different to the previous assessment model, but is provided for information to those interested in implications of model development:

1. **Started with the 2021 reference case model** (Castillo Jordan et al., 2021).
2. **Application of the catch-conditioned approach and concentrated CPUE likelihood:** resulted in an overall lower $SB_{\text{recent}}/SB_{F=0}$ and SB but with similar trends with the exception of the early 1960s, compared to model 1.
3. **Reduction of the number of cubic spline selectivity nodes from 4 to 3:** had almost indistinguishable differences from the previous model 2.
4. **Simplification by collapsing the sub-regions and assigning the overlap area to the WCPFC-CA:** had almost indistinguishable differences from model 3.
5. **Simplification by collapsing WCPFC-CA into one region and assigning new fisheries definitions using areas-as-fleets approach resulting in two regions (WCPFC-CA and EPO) over the spatial extent:** resulted in $SB_{\text{recent}}/SB_{F=0}$ beginning and terminating at equivalent levels but exhibiting higher levels in the interim, compared to model

4. The SB showed similar results with the exception of higher SB at the beginning of the time-series compared to model 4.

6. **Removal of "contaminated" LF data:** resulted in $SB_{\text{recent}}/SB_{F=0}$ beginning and terminating at equivalent levels but exhibiting lower levels in the interim. The SB showed similar results with the exception of lower SB at the beginning of the time-series compared to model 5.

7. **Updating with the new MULTIFAN-CL executable (v2.2.5.1):** resulted in indistinguishable results from model 6.

8. **Converting recruitment frequency from quarterly (at the beginning of the calendar year) to annual (occurring at the beginning of month 10):** resulted in similar trends with slight differences in the fluctuations throughout the time-series for both $SB_{\text{recent}}/SB_{F=0}$ and SB compared to model 7.

9. **Converting the time-step for CPUE to annual, conversion from Maunder M to Lorenzen M, extending time-series back to 1954, updating length-weight parameters, updating with the new years of data, applying LF data weighting using Francis method, inclusion of CAAL, and implementation of the VB offsets growth:** resulted in similar trends with more variability and lower levels of both $SB_{\text{recent}}/SB_{F=0}$ and SB compared to model 8.

Overall, the 2024 South Pacific albacore diagnostic case model estimates very similar levels of $SB_{\text{recent}}/SB_{F=0}$ and SB at the beginning and terminal years compared to the 2021 diagnostic model. The temporal trends were also similar, although the 2024 model showed lower levels of $SB_{\text{recent}}/SB_{F=0}$ and SB during the middle period of the assessment. The increasing trend in recent years shown by the new diagnostic case model is notable. This recent increase results in 2022 SB (and less so for $SB_{\text{recent}}/SB_{F=0}$) that is higher than the level estimated in the final year (2019) of the 2021 WCPFC-CA only model (i.e. compare navy blue and grey lines; Figure 14).

10.3 Model parameter estimation

Estimates from the diagnostic case model are discussed in this section to explore model behaviour and parameter estimates.

10.3.1 Selectivity

Estimated selectivity functions were consistent with known operational characteristics of the different gear types; longline fisheries selecting larger, older individuals (Figure 15) and the driftnet and troll fisheries selecting smaller, younger fish that are more prevalent in the surface fisheries of the southern regions. These latter fisheries displayed dome-shaped selectivities which, increased rapidly from age 0 to maximum selectivity at, or below, 4 years of age and then declined back to 0. Slight differences among these surface fisheries existed, with the driftnet fisheries selecting a

slightly wider range of ages and a slightly higher modal age than the troll fisheries (age 4) whereas, the troll fisheries (including the EPO) were estimated to have a lower modal age (age 3).

Most of the longline fisheries were estimated to have dome-shaped selectivity with the exception of the following: the two longline indices (as constrained with high penalties to have asymptotic selectivity), the longline fisheries in sub-regions 1-ef, season 1 in some of the longline fisheries in northern areas of the WCPFC-CA, and season 1–3 in the AU/NZ in sub-region 1-abcd. Slight differences among longline fisheries were evident in the age at which fish began to be selected, and the slopes of the ascending limbs of the selectivity functions, which varied for different flag groupings and regions of operation of the fishery. Longline fisheries in the northern part of the WCPFC-CA exhibited significant seasonality in selectivity patterns, as would be expected due to seasonal movements of different sized fish.

10.3.2 Growth and natural mortality

Growth estimates from the diagnostic case model indicated a mostly linear shape for the first four age classes as estimated by the VB offsets parametrization (Figure 16). The estimated growth parameters were 45.54cm (fixed), 101.36cm, and 3.398 for $L1$, $L2$, and k, respectively with -6.191, -8.422, and -5.042 estimated offsets for ages 2–4, respectively. The added flexibility from estimating offsets in early stage growth provided some deviation from the traditional shape of the VB growth curve specified by MULTIFAN-CL which, has been indicated in previous research on albacore growth. Specifically, Farley et al. (2021) found that a logistic curve fit the data better than a VB curve. The M-at-age curve derived from the growth and maximum age of 15 (A_{max} method) estimates are shown in Figure 17.

10.3.3 Recruitment

The time series of estimated recruitment is displayed in Figure 26. High recruitment estimates in the early period of the assessment are likely related to the initial high CPUE that drops rapidly, typical of tuna longline indices in the WCPO, and are treated with some caution. Across the time series recruitment estimates are highly variable at the inter-annual time scale and show an increasing trend from around 1980. The estimated 95% confidence intervals are very wide in the early years of the time series, and contract moving forward in time. Note that the terminal recruitment is fixed at the average of the time series and therefore has lower uncertainty than would otherwise be the case.

The spawner recruit relationship for the diagnostic case is displayed in Figure 27. Recruitment deviation around the spawner recruitment relationship is similar across the SB levels estimated throughout the model time period.

10.3.4 Biomass and biomass depletion

SB declined rapidly from 1950s until the late 1970s after which SB stabilised and even has increased slightly in the WCPFC-CA until 2015 after which it declined again until the very recent years when an increase is estimated (Figure 28). The overall trend is relatively consistent in each model region despite the estimated levels of biomass associated with each region differing substantially. As for recruitment, 95% confidence intervals for SB are wider in the early years of the time series.

$SB_t/SB_{F=0(t)}$ has declined continuously since 1954 in both model regions and in 2022 is estimated to beat its lowest level in the EPO. In the WCPFC-CA $SB_t/SB_{F=0(t)}$ reached a minimum in 2020 and has since increased sharply. (Figure 29). The 95% confidence intervals for $SB_t/SB_{F=0(t)}$ begin at zero in 1954 (because of the unexploited population condition) but expand moving forward in time. At the end of the time series, the CV is approximately 0.05.

10.3.5 Fishing mortality and age-specific exploitation

An increase in fishing mortality of adult age-classes is estimated to have occurred over most of the assessment period (Figure 30), accelerating since the 2000s. Adult mortality has continued to increase from 2010, and although the last two years show a drop in adult fishing mortality, the recent year estimates are considered less reliable. Juvenile fishing mortality increased until ∼1990, with a peak in 1989 due to the driftnet fishery and has remained stable at a comparatively lower level since that time. A small peak in juvenile fishing mortality is estimated to have occurred in 2020.

The estimated relative frequency at age by decade (Figure 31) indicated the dominance of the first four age classes and the persistence of older fish throughout the temporal extent, although in lower proportions. The estimated instantaneous mortality at age and decade indicated that prior to 2000, effectively no fishing mortality on the youngest age classes until ∼age-3 when a gradual increase with age was indicated. After 2000, the fishing mortality increased more noticeably at age-5 until age-7 when fishing mortality decreased gradually. The estimated annual instantaneous fishing mortality by age and area (Figure 32) indicated a sharp increase in fishing mortality for age-6 and older in the WCPFC-CA. A similar pattern was exhibited in the EPO after 2010 most notably for ages 6–8. These age-specific characteristics are reflected in the estimates of 2019-2022 depletion by age class (Figure 33).

10.3.6 Fishing impact

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 34). Fishing impacts were estimated to increase gradually through the 1960s in the WCPFC-CA primarily due to the longline fishery with the tropical, sub-tropical, and southern area fisheries all having approximately equal contributions.

During the late 1960s, the longline in the EPO began to have an impact, although to a lesser extent. During the 1970s, impacts stabilized and even declined until the mid 1990s with the southern longline having the largest impact. In the 1980s, the troll fishery (with some contribution from the drift-net fishery in the 1980s) began to have a substantial impact until the early 2000s when the impact decreased to a stable level as a result of the banning of driftnet fishing. From the mid 1990s the impact for the sub-tropical longline increased notably until approximately 2018 when it began to decline while the EPO longline began to have a larger impact than previously.

10.4 Diagnostics results

In the following section, we present diagnostics for the diagnostic case model.

10.4.1 Convergence diagnostics

The following indicators of model convergence were recorded:

- The maximum parameter gradient for the diagnostic case model was 5.52e-07.
- The Hessian was positive definite.
- A jitter analysis was not able to improve on this solution, with all 25 jitters having total negative log likelihoods $\geq$ the diagnostic case model (Figure 18).
- Of the 165 estimated parameters, seven had converged at their lower bound. In all cases, these were the first of the three selectivity spline coefficients for longline fisheries 1/3 (shared selectivity) for quarter (Q) 1 and Q2, fisheries 2/4 for all quarters and fishery 14 for semester 1. It appeared that this was related to the model trying to estimate zero selectivity for young age classes for these fisheries/seasons. Potentially this could be remedied by specifying the required number of age classes to have zero selectivity. Unfortunately, this specification in MULTIFAN-CL can currently be made only generically (for all seasons). This is an area where the model code can be improved.
- We examined the correlation matrix for the estimated parameters, and found 16 cases where correlations <-0.9 occurred. In all cases these occurred for the 2^{nd} and 3^{rd} (for fisheries with 3 selectivity nodes) or for the 3^{rd} and 4^{th} nodes (for fisheries with 4 selectivity nodes).

Overall, we are confident that the converged diagnostic model represents a global solution on the basis of the small parameter gradients, a positive definite Hessian, and the jitter analysis. However, there are areas where the model could be improved in the selectivity parameterisation, which is an area of follow-up research.

10.4.2 Model fit

The Diagnostic case model showed acceptable fits to the CPUE indices (Figure 19), with predicted relative abundance being within the 95% confidence intervals of the longline indices throughout

the time series. The longline residuals plots show some patterns for the final decade, with mainly negative residuals (indicating some over-estimation of relative abundance) for the WCPFC-CA index and mainly negative residuals for the EPO index.

The diagnostic model provided reasonable fits to the time-aggregated length frequency data for fisheries in the WCPO, particularly the longline fisheries (Figure 20). Poorer fits occurred for fisheries with low sample sizes (adjusted and/or observed) and temporally patchy and variable compositions (e.g. troll and driftnet fisheries). Furthermore, there were poorer fits for several longline fisheries, more specifically, for sub-regions 1-ef where the model had difficulty fitting the multimodal observations. For the longline fishery in the EPO, the data were fit reasonably well, while, similar to the WCPO, a poorer fit was obtained for the troll fishery in the EPO. The time series of observed and predicted median lengths show some lack of fit, e.g. an increasing trend in median size in fishery 4 (Figure 21). Note that the estimates of effective sample size provided by the Francis weighting method strongly down-weighted the LF data for most fisheries, allowing the model to fit these data less well than would otherwise be the case.

The diagnostic case model provided a reasonable fit to the CAAL data, although there is a tendency for the diagnostic model to underestimate age composition for the largest length classes in the sample, as evidenced by a group of positive residuals for 90-110cm sizes (Figure 22). This is an area of follow-up research with the fish ageing scientists.

10.4.3 Age-structured production model (ASPM)

Two versions of the ASPM were fitted, one without and one with estimated recruitment deviations. In the first only population scale is being estimated, informed only by the CPUE data. The mean recruitment, SB and $SB_{F=0}$ are scaled reasonably well compared to the diagnostic case model, indicating that the CPUE indices are informing population scale (Figure 23). Some of the time series changes in SB and $SB_{F=0}$ are also captured, although SB is over-estimated through the middle part of the time series. When the estimation of recruitment deviations is included, the estimates of recruitment, SB and $SB_{F=0}$ track the diagnostic case model estimates reasonably closely. This indicates the desirable property that population scaling are informed by the CPUE indices.

10.4.4 Catch curve analysis (CCA)

The catch curve analysis, in which the model is fitted to the size composition and CAAL data only, indicates somewhat lower population scaling and divergent trends, particularly in the early part of the time series, compared to the diagnostic case model (Figure 23). However, in the more recent period since 1990 (where better sampling occurs for most fisheries), the trends and scaling tend to converge to some extent. Interestingly, some of the recent recruitment variability lines up quite well between the CCA, diagnostic case and the ASPM with recruitment deviations. This suggests that the more recent population signals are consistent among the CPUE and LF/CAAL data.

10.4.5 Likelihood profiles

Likelihood profiles by data component and by index and fishery for the CPUE indices and LF data, respectively, are shown in Figure 24. The profile by data component indicates that in aggregate, the CPUE indices, LF and CAAL data are all providing a consistent signal with respect to population scale.

For the CPUE indices, both longline indices are very consistent while the NZ troll index indicates higher population scaling. The extent to which this impacts the results of the assessment is explored further in the Sensitivity Analysis section.

For the LF data, there is considerable variability in population scale information across the range of fisheries. This is perhaps not unexpected as the data are broken down into smaller components, but nevertheless indicates an area where follow-up research is required.

10.4.6 Retrospectives

The retrospective analysis indicated the diagnostic model showed no significant retrospective patterns (Figure 25) as evidenced by $|\rho| \leq 0.1$ for SB and $SB_t/SB_{F=0(t)}$ (for each region and combined regions) however, recruitment indicated some retrospective pattern $\rho = -0.31$ for each region and combined regions. This might be related to the assumption that recruitment in the final year is fixed at the historical average.

10.5 Sensitivity analyses

10.5.1 CPUE indices

Two longline CPUE indices for region 1 (WCPFC-CA) have been explored in this assessment (see Teears et al. (2024)). In the diagnostic case model, the region 1 longline index was constructed for the full tropical region (covering sub-regions 1-a, 1-b, 1-c, and 1-d) and is an annual average index using data for all months of the year. This CPUE index is denoted WCPFC-CA.NORTH. As an alternative, we applied a more restricted index, denoted WCPFC-CA.SPAWN, that was designed to focus on the South Pacific albacore spawning area (10-25°S) and season (October–January). Time-series of the two indices, and their 95% confidence intervals, are shown in Figure 35. Both indices show sharp initial declines to the mid-1970s, although estimates in the 1950s and early 1960s are associated with high uncertainty. The WCPFC-CA.NORTH index is then fairly stable through to the late 1990s after which there is a slow decline. By contrast, the WCPFC-CA.SPAWN index shows a more continuous decline from the late 1970s to the present.

Models incorporating these alternative indices for the WCPFC-CA region showed good convergence, and with positive definite Hessians. Fits to data were similar and acceptable. In particular, in both cases the assessment model fits to the indices were good (Figure 36). The key stock assessment results are shown in Figure 37. The trends and inter-annual variability in recruitment are similar for both models, with small differences in scaling of the three metrics (recruitment,

SB, and $SB_t/SB_{F=0(t)}$). The three reference point variables ($SB_{\text{recent}}/SB_{F=0}$, $SB_{\text{recent}}/SB_{\text{MSY}}$, and $F_{\text{recent}}/F_{\text{MSY}}$) were also very similar for the two models (Table 5). Given the similarity of the results, it was decided to include only the model based on the WCPFC-CA.NORTH index in the final multi-model ensemble.

10.5.2 Troll CPUE index

A CPUE index based on the NZ troll fishery (Neubauer and Hill-Moana, 2024) was incorporated into the diagnostic case model. The motivation for this was to provide the model with information on the younger age-classes of albacore that are targeted by this fishery. However, we acknowledge concerns around whether the troll CPUE responds to stock-wide recruitment processes, or more reflects environmentally driven availability of juvenile albacore to the local NZ fishery. Therefore, we undertook a sensitivity analysis to examine the impacts of effectively removing the influence of the CPUE index by increasing the CV to 10,000, thereby down weighting the data to a trivial level. We adopted this approach rather than simply removing the data to enable the model to predict the CPUE even though it was not being fitted. The length frequency data associated with the troll index was also removed for this test, and the selectivity fixed according to that estimated in the diagnostic case model.

The main effect of including the troll CPUE index is to provide some constraints on recruitment variability. This is evident in the recruitment estimates for the models including and excluding the troll index (Figure 38). In particular when the troll index is excluded, recruitment estimates for the recent period (2018-2021) are much higher; and probably unrealistically so. This is also reflected in the model predictions of the troll CPUE (Figure 39). While the exclusion of the troll index does not unduly impact the estimates of pre-2018 recruitment, SB, and $SB_t/SB_{F=0(t)}$, such high terminal recruitments will have a large impact on projections. There is also some minor impact of excluding the troll CPUE on the reference point variables (Table 5) probably related to the inflated terminal recruitment. In view of this stabilising influence and the likelihood that the index is at least partly driven by recruitment dynamics, it was decided that the troll CPUE index should be retained in the diagnostic and ensemble models.

10.5.3 NZ troll length frequency data

For the same reasons alluded to earlier regarding the representativeness of the NZ troll fishery CPUE, similar concerns have been raised regarding the length frequency data for this fishery. To investigate the impact of the troll length frequency data on key stock status results, we undertook a sensitivity in which the length frequency data for the three troll fisheries were removed from the data. In doing so, selectivity for the troll fisheries was fixed at their estimates from the diagnostic case model. We adopted this approach because it would not have been possible to estimate selectivity without any size composition data, and the interest is in the impact of the length data on recruitment and related population estimates.

The removal of the troll length frequency data resulted in some minor differences in scaling and the timing of recruitment troughs and peaks (Figure 40a) and minor differences in the scaling of SB (Figure 40b) and $SB_t/SB_{F=0(t)}$ (Figure 40c). The reference point variables were slightly more optimistic with the exclusion of the troll length frequencies Table 5. In view of these results, we were comfortable in including these data in all models for the assessment.

10.5.4 Weighting of the conditional age-at-length data

All models considered in this assessment incorporate conditional age-at-length (CAAL) data. These data were compiled from 654 age estimates obtained from reading the otoliths of albacore of known length (Farley et al., 2021). Rather than use estimates of growth parameters obtained from external analyses of these data as fixed parameters in the assessment, we opted, following the recommendation of Punt et al. (2023), to incorporate the data as CAAL into the assessment model. The likelihood for the CAAL data was multinomial, and the weighting that the data receive in the overall likelihood depends on the assumed ESS of individual samples (categorised by sampled fishery and time period). While the nature of the sampling design is consistent with the multinomial structure, we recognise that it is likely that the multinomial ESS is less than the observed sample size (OSS) because of unaccounted for process error in the model. In the diagnostic and ensemble models we made the arbitrary assumption that ESS is 0.75*OSS. In sensitivity analysis, we tested the impact of this assumption by comparing the diagnostic model with models in which ESS=0.5*OSS and ESS=1.0*OSS.

The key stock assessment results are almost identical regardless of the multiplier used to specify the ESS (Figure 41, trajectories of recruitment, SB, and $SB_t/SB_{F=0(t)}$ are effectively identical). The reference point variables are also unaffected (Table 5). Therefore, it was not necessary to include the weighting of the CAAL data in the ensemble of models for characterising uncertainty.

10.5.5 Movement

Initial attempts to estimate movement internally in the assessment produced unrealistic results thus, this approach was not pursued. Instead, movement probabilities were determined from a SEAPODYM model of South Pacific albacore fitted to high resolution spatial data (Senina et al., 2020). The movement probabilities were averaged across years and specified by age class and season (see Movement). To test the assessment model's sensitivity to the movement specification, we ran four additional models for comparison: a) no movement between the two regions, b) movement probabilities half those indicated by the SEAPODYM model, c) movement probabilities twice those indicated by the SEAPODYM model and d) movement probabilities in both directions of 0.9 per period approximating full and instantaneous mixing of the stock. An example of the range of movement probabilities considered, for the first quarter, is shown in Figure 42. We did not consider scenarios a) and d) to be necessarily biologically plausible; they are included here as extreme scenarios at the lower and upper limits and were requested by participants at the Pre-assessment

Workshop.

Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ for the five movement scenarios are shown in Figure 43. With the exception of the no movement scenario, the results for the other movement scenarios provide reasonably consistent results. There was a slight decrease in population scaling with increased movement, but the results for the 2 x SEAPODYM movement and full mixing are practically identical, indicating that the results are relatively insensitive to movement higher than 2 x SEAPODYM. Impacts of movement on the reference point variables (Table 5) are consistent with the impacts on scaling. On the basis of these results and noting that we do not consider "no movement" to be a biologically plausible scenario, it appears that the stock assessment results are robust to the movement assumption and that it is not necessary to consider alternative movement scenarios in the final model ensemble for uncertainty characterisation.

10.5.6 Effort creep

In stock assessments that use commercial CPUE as the basis for abundance indices, there is always a concern that the CPUE series does not account for increasing fishing power over time resulting from, inter alia, the adoption of new technology, increasing skill of fishers and information sharing. In combination, these factors are often referred to as "effort creep". The CPUE analysis undertaken to support this assessment (Teears et al., 2024) may have captured some aspects of effort creep, e.g., by including changes in longline gear materials and configuration in the statistical analysis. However, it is probable that other sources of effort creep have not been accounted for due to absence of information. We have therefore undertaken a sensitivity analysis to investigate the possible impacts on assessment results of unaccounted for effort creep. We did this by modifying the CPUE time series of the two longline indices in such a way that effective effort increased by 0.5%, 1% or 1.5% annually (Figure 44). These changes were applied consistently over the full time series. Changes were assumed to be additive rather than multiplicative. For example, under the 1.5% scenario, effective effort (CPUE) in 2022 is approximately two times (half) the level in the absence of effort creep. These levels of effort creep are similar to those proposed for longline fisheries in Hoyle (2024).

Somewhat surprisingly, the key stock assessment results are consistent across the effort creep scenarios (Figure 45, Table 5). Most of the impact on recruitment and SB occurs near the start of the time series, making initial declines steeper. However, there appears to be little impact on the estimates subsequent to the 1990s. The estimates of SB depletion are impacted mainly over the early-to middle portion of the time series. There is minor impact of effort creep scenarios on SB depletion subsequent to about 2000. For this reason, we considered it unnecessary to incorporate uncertainty due to effort creep into the model ensemble.

This possibly counter-intuitive result is also consistent with the minor impact seen when running the model with the alternative WCPFC-CA.SPAWN CPUE index. This index also displayed a stronger decline in CPUE (Figure 35, Figure 36) but also had little impact on the assessment

outcomes (Figure 37). This apparent insensitivity may be related to the relatively old portion of the population that is estimated to be selected by longline fisheries compared to, for example, the portion making up the SB (Figure 46). This makes it possible for the model to reconcile different trends in longline CPUE and SB. Further investigation of this issue is required.

10.5.7 Recruitment distribution

In this assessment, we specified the distribution of recruitment, assumed constant over time, between regions 1 (WCPFC-CA) and 2 (EPO). The assumed allocation of (0.82, 0.18) between regions 1 and 2, respectively, was informed by the SEAPODYM model for South Pacific albacore (Senina et al., 2020). Initial attempts to estimate the recruitment distribution internally in the assessment model produced results considered unrealistic (most of the recruitment in region 2). Given that we are using a fixed allocation, we examined alternative plausible specifications (90:10 and 70:30) to see if the assumption impacted the key assessment results. The results indicated very little impact (Figure 47, Table 5); therefore, alternative recruitment distributions were not included in the model ensemble for uncertainty characterisation.

10.5.8 Number of age classes

It is necessary to make an assumption regarding the number of age classes in the model. This assumption can be informed by the known maximum age, although fewer age classes might be appropriate if there is little change in the mean size at age with increasing age for the older age classes, noting that the final age class is treated as a "plus group" in the catch equations of the model. In this assessment, consistent with previous assessments of this stock, we assumed 12 age classes. In this sensitivity analysis we ran models with 10 age classes and 15 age class for comparison. The results for recruitment, SB and $SB/SB_{F=0}$ are shown in Figure 48. The estimates of recruitment are consistent across the three models, but there are small differences in scaling of SB and to a lesser extent $SB_t/SB_{F=0(t)}$, with the 10 age class model having slightly higher scaling that the 12 and 15 age class models, which are practically identical. The reference point variables show similar lack of sensitivity (Table 5). In view of these results, we were comfortable in going forward with the 12 age class model in the model ensemble and not including uncertainty in this assumption.

10.5.9 Selectivity and catchability time-blocks

Our standard assumption in the diagnostic case model is that selectivity is constant over time for all fisheries, including index fisheries, and that catchability is constant for the index fisheries. There is the possibility, at least for the long-standing distant-water longline fisheries, that selectivity and catchability could have changed in response to changes in hooks-between-floats and mainline material. Inspection of available data suggested that 1977 and 1994 were years in which major changes could have occurred in DWFN longline fleets. We therefore constructed two models in

which selectivity for fisheries 1–4, 8, 14, 16, 18 and 20 was allowed to vary across three time periods – 1954–1976, 1977–1993 and 1994–2022. In the second of these models, the two longline-based index fisheries (18 and 20) were also split into these periods to allow catchability to change.

Adding time blocks for selectivity and catchability added considerable parameter complexity increasing the degrees of freedom of the model. The fit statistics of the models are in Table 6. The Akaike Information Criterion (AIC) statistic indicates that the improvement in fit (smaller negative log likelihood) achieved by the two models with time blocks does not justify the additional parameter complexity. On these grounds, we would reject the more complex models and retain the diagnostic case. For completeness, we show the comparison of model results in Figure 49. The time-blocked models estimate higher recruitment and SB, particularly early in the time series, and slightly more optimistic $SB_t/SB_{F=0(t)}$. Accordingly, the reference point variables become more optimistic under the time-blocked scenarios (Table 5). We did not diagnose these results in detail, but it is possible that some of the reduction in size of fish in the catch is attributed to changed selectivity rather than a change in the size structure of the underlying population.

10.5.10 Growth

During model development, we found that a von Bertalanffy growth curve struggled to reconcile both the length frequency data, in particular the positioning of the three modes in the NZ troll fishery data, and the conditional age-at-length data. However, after a thorough investigation, we found that allowing "VB offsets" to be estimated, i.e., deviations from von Bertalanffy growth, for age classes 2, 3 and 4 did a reasonable job in fitting both data sets. The growth model approach adopted for the diagnostic case is summarised in Table 7.

Table 8 compares the likelihood components of the diagnostic case model with a standard VB model without estimated offsets. In the latter, all growth parameters ($L1$, $L12$, k, $V1$ and $V2$) were estimated. The standard VB model has better likelihood components for CPUE and CAAL but is inferior for length-frequency data compared to the VB offsets model. Overall, the VB offsets model has a superior likelihood by 7 units. The VB model estimates a higher variance for mean-length at age for the young age classes (Figure 50) and cannot model the modal structure of the NZ troll fishery size data (Figure 51). For these reasons, we have retained the VB offsets model as the basis for our diagnostic case and the model ensemble. For completeness, the key stock assessment results for the VB model are presented in Figure 52 and Table 5.

10.5.11 Natural mortality

The rate of natural mortality is uncertain and cannot be estimated from data typically included in stock assessments. For this reason, it is essential to consider the uncertainty in M in any assessment. Here we have compared the key stock assessment results from three levels of $\overline{M}$ (averaged across age classes 4-12) – 0.36, 0.24, and 0.54 yr^{-1}. These values represent, respectively, the mean of the A_{max}–based prior (assuming A_{max} = 15 yr), and approximate lower and upper 95% confidence

intervals on the mean assuming a CV of 0.2.

The different M settings have a profound impact on the key stock assessment results (Figure 53). Between the lower and upper values tested, recruitment varied by about 20 times, SB by more than 2 times and $SB_t/SB_{F=0(t)}$ in recent years ranged from about 0.3 to 0.7. The impact of M on the reference point variables (Table 5) is similarly high. Clearly, M is a major factor contributing to uncertainty in the assessment results and is therefore included in the framework developed for the multi-model ensemble for charactering uncertainty.

10.5.12 Steepness

Like natural mortality, the steepness parameter (h) of the Beverton and Hold stock-recruitment relationship is almost always uncertain and cannot be estimated internally in the assessment. Here we have used three settings for h, 0.65, 0.8 and 0.95, to compare the key stock assessment results. This is the range of steepness values that have been typically used in WCPFC-CA tuna assessments. The different steepness settings have virtually no effect on the estimates of recruitment and SB, and have slight effect on the $SB_t/SB_{F=0(t)}$ at lower levels of SB (Figure 54). However, there is a strong effect on the MSY-related reference point variables; $F_{\text{recent}}/F_{\text{MSY}}$ and $SB_{\text{recent}}/SB_{\text{MSY}}$ (Table 5) with lower (higher) steepness producing more pessimistic (optimistic) outcomes. We have therefore included steepness in the multi-model ensemble to capture these effects.

10.6 Monte–Carlo model ensemble uncertainty estimation

Diagnostics for the 100 models in the ensemble indicated that all models converged based on the selected criteria and were, therefore, included in the final results. Tables of diagnostics and likelihoods are available in Appendix 2.

The results of the ensemble uncertainty analysis are summarised in several forms: 1) histograms of model uncertainty estimated from Monte-Carlo draws from the model ensemble of $SB_{\text{recent}}/SB_{F=0}$, $SB_{\text{recent}}/SB_{\text{MSY}}$, and $F_{\text{recent}}/F_{\text{MSY}}$ coloured by values of $\overline{M}$ and h (Figure 55, Figure 56, and Figure 57) combined with 2) estimates of $SB_{\text{recent}}/SB_{F=0}$, $SB_{\text{recent}}/SB_{\text{MSY}}$, and $F_{\text{recent}}/F_{\text{MSY}}$ by values of $\overline{M}$ and h for each model in the ensemble; 3) quantile trajectories (90% and 75%) are provided for $SB_{\text{recent}}/SB_{F=0}$ and SB in Figure 58 and Figure 59; 4) a table of summary statistics of reference points for the model ensemble is included in Table 9; and 5) Majuro and Kobe plots are shown for estimates from the model ensemble and the dynamic MSY analysis (Figure 61).

The models from the ensemble indicated the probability that $SB_{\text{recent}}/SB_{F=0}$ <0.2 was 0, the probability that $F_{\text{recent}}/F_{\text{MSY}}$ >1 was 0, and the probability of $SB_{\text{recent}}/SB_{\text{MSY}} > 1$ was 0.999. $\overline{M}$ had a high influence on $SB_{\text{recent}}/SB_{F=0}$, $SB_{\text{recent}}/SB_{\text{MSY}}$, and $F_{\text{recent}}/F_{\text{MSY}}$ as evidenced by the linear (or exponential for $F_{\text{recent}}/F_{\text{MSY}}$) relationship indicated in Figure 55, Figure 56, and Figure 57 whereas, the relationship between h and these reference points were less evident.

Annual quantiles indicated 90% of ensemble had terminal $SB_{\text{recent}}/SB_{F=0} \geq 0.358$, 0.353, and 0.364

for the WCPFC-CA, the EPO, and combined regions, respectively (Figure 58). Annual quantiles indicated 90% of ensemble models had terminal $SB \geq$ 131.0, 59.9, and 192.0 (1,000s of tons) for the WCPFC-CA, the EPO, and combined regions, respectively (Figure 59). Annual quantiles indicated 90% of ensemble models had terminal fishing mortalities $\leq$ 0.454, 0.323, and 0.418 for the WCPFC-CA, the EPO, and combined regions, respectively (Figure 60)

The dynamic MSY analysis indicated that for all time periods, the $SB_{\text{recent}}/SB_{F=0}$ was > 0.2, $SB_{\text{recent}}/SB_{\text{MSY}}$ was > 1 and the $F_{\text{recent}}/F_{\text{MSY}}$ was < 1 (Figure 61). Similarly, all models in the ensemble for the recent period (2019–2022) indicated the $SB_{\text{recent}}/SB_{F=0}$ was > 0.2, $SB_{\text{recent}}/SB_{\text{MSY}}$ was > 1 and the $F_{\text{recent}}/F_{\text{MSY}}$ was < 1.

10.6.1 'Status quo' stochastic projections

Preliminary "status quo" stochastic projections were performed using the 100 models developed in the model ensemble. Projections were run for 40 years, and therefore ran from 2023 through to 2062. Future catch levels in longline and troll fleets across the South Pacific were assumed to be the average catches across the period 2020-2022. Catchability of each fishery was assumed to remain constant in the projection period at the level estimated in the terminal period of the assessment model. To be consistent with other analyses and noting recent increases in catch levels within the EPO region, future catch of fisheries within the "remainder of the EPO" (EPO excluding the overlap area) were scaled up to an equivalent of 22,500 mt, the levels reported in 2021 and 2022.

50 stochastic projections were performed from each of the 100 assessment models. Future recruitment was defined by the estimated stock recruitment relationship, with variability around it defined by recruitment deviation estimates from the stock assessment over the period 1972 to 2020.

Figure 62 presents the resulting South Pacific albacore depletion level of the stock within the WCPFC Convention Area. Depletion is calculated consistent with the guidance provided by WCPFC20 in terms of the calculation of the iTRP[6], reflecting the dynamic nature of this metric. Also included is a point denoting the calculation of $SB_{\text{recent}}/SB_{F=0}$ as defined for the stock assessment (see definition in Table 4).

[6] "Spawning potential depletion" refers to the estimated South Pacific albacore spawning potential as a percentage of the estimated spawning potential in the absence of fishing (i.e., the unfished spawning potential). The metric is dynamic and is estimated for each model time step.
The method to be used in calculating spawning potential in the absence of fishing ($SB_{F=0}$) shall be:
a. $SB_{F=0,t_1-t_2}$, is the average of the estimated spawning potential in the absence of fishing for a time window of ten years based on the most recent South Pacific albacore stock assessment, where $t_1 = y - 10$ to $t_2 = y - 1$ where y is the year under consideration; and
b. The estimation shall be based on the relevant estimates of recruitment that have been adjusted to reflect conditions without fishing according to the stock recruitment relationship.

11 Discussion

11.1 General remarks on the stock assessment

South Pacific albacore present multiple challenges for stock assessment because much of the spatial population structure and connectivity dynamics have yet to be clearly defined. As such, Goethel and Berger (2017) suggested that this could lead to biased estimates of stock status indicators and a simpler spatial structure would be preferable to one of higher complexity. The current assessment has been substantially spatially simplified to two regions (WCPFC-CA and EPO), temporally simplified to annual recruitment (occurring in October), and fisheries restructured in an areas-as-fleets approach based on a regression tree methodology using LF data (Lennert-Cody et al., 2010, 2013; Potts et al., 2024). Furthermore, the current assessment continued the application of recruitment distribution and movement information from SEAPODYM (Senina et al., 2020), which combine biological, fisheries, and environmental data to provide more informed estimates of these important parameters given the lack of informative movement and recruitment data.

This assessment has incorporated several changes that we consider improves the quality of the assessment, including:

- Incorporating the CAAL data into the assessment has, in combination with the use of the VB offsets growth model, resolved some of the previous uncertainties in growth that were had been problematic. As a result, there is now more consistency in growth estimates and reduced data conflict between the CAAL data and the LF data than in the past. The allowance for non-VB growth and possibly a more appropriate weighting of the LF data based on the Francis method appears to have been critical in resolving this issue. An area of further improvement would be to allow for a logistic growth curve in MULTIFAN-CL, which might provide a more parsimonious method of accommodating the near linear growth of juveniles.
- Conversion of the model to a catch-conditioned approach and a simplification of the regional structure of the model has resulted in better model diagnostics and an overall improvement in the statistical properties of the model. The changes allowed for rapid model convergence, positive definite Hessians and reasonable estimates of most parameters away from bounds and without high parameter correlations.
- The simplified model structure and parameterisation resulted in a considerable reduction in run time for models. Consequently, a greater range of sensitivity analyses were able to be undertaken in the time available, resulting in a better understanding of how various structural, parameter and data decisions impact the key model results.
- We adopted a specification for M-at-age that was consistent with emerging best practice, as identified by the 2023 CAPAM "Tuna Stock Assessment Good Practices Workshop". Accordingly, natural mortality was specified as age-specific based on Lorenzen (1996) and M was scaled using the Hamel and Cope (2022) approach. A prior distribution was developed

following this approach with a somewhat lower than recommended CV of 0.2 to constrain M to plausible values.

- The treatment of uncertainty in this assessment was greatly improved through the use of a model ensemble Monte Carlo bootstrap approach (Ducharme-Barth and Vincent, 2021; Ducharme-Barth et al., 2021) applied to uncertainty due to natural mortality and steepness. The approach provided a more natural weighting of models in the ensemble than had been the case with a fully factorised grid across multiple dimensions of uncertainty. This, and the incorporation of estimation uncertainty for individual models of the ensemble into the estimation of distributions of key stock status reference points results in a more comprehensive and balanced treatment of uncertainty than in previous assessments.

11.2 Examining other key data inputs

The approach to standardise the longline CPUE data for the indices was similar to the previous assessment in that both used a spatiotemporal delta-GLMM modelling approach. However, the synthetic targeting cluster variable as a catchability covariate was replaced with hooks-between-floats due to confounding of the former with abundance. Vessel flag as a catchability covariate was maintained from the 2021 assessment; however the addition of season as a catchability covariate informed the model of seasonal variation in fishery dynamics caused by non-biological drivers.

Preliminary analyses explored splitting the CPUE time series to allow for changes in gear configuration over time. However, there was concern that important information would be lost with this approach. Ideally, it would be better to model the entire time-series and include gear information to account for changes in selectivity and/or catchability due to gear changes. This was further explored with sensitivities in selectivity and catchability time-blocks. However, important gear configuration information is lacking from the data to appropriately model these dynamics. There remains a need for further exploration.

Spatial stratification of the albacore stock by size has been observed, with smaller/younger fish more available in the southern regions of the assessment domain and the larger individuals more distributed throughout the sub-tropical waters. Standardizing the catch data by size class, or life-stage (juveniles versus adults), could improve the estimation of selectivity for the index fisheries, which may differ from the selectivity of the capture fisheries. These alternative size-based approaches to CPUE index standardization (Maunder et al., 2020) could prove valuable for the next South Pacific albacore assessment.

There is some limited data from archival tagging on vertical habitat utilisation of albacore, both from the south and north Pacific. This may provide an opportunity to revisit the habitat-based standardisation approach (Bigelow et al., 2002) to potentially better estimate effective longline fishing effort for albacore.

While developments related to the CPUE standardisation methodology are important, perhaps

more important is the need for greater focus on improving the scope of the data to support these analyses so indices of relative abundance are more representative of the true stock dynamics. Scientific observers collect detailed vessel, gear, and fishing strategy related information that could improve these analyses, but coverage levels are insufficient (generally less than 5%) to provide the spatial and temporal coverage needed for these analyses. Electronic monitoring for longline fisheries has shown great promise (Brown et al., 2021), and steps to expand and enhance these tools to better monitor the longline sector are encouraged. Lastly, developing collaborative partnerships and cooperative research programmes with the fishing industry is important to; 1) explore questions related to fishing strategy, gears, decision making, and other operational factors influencing catch rates, and 2) improve the understanding of fishery operations by the scientists tasked with conducting these analyses. Both of these would no doubt lead to improvements in the approaches for developing abundance indices from fishery dependent catch and effort data.

11.3 Main assessment conclusions

The 2024 stock assessment provided results that were broadly consistent with the previous assessment. Spawning potential is estimated to decrease until 1975 and then stabilise or gradually increase until 2015. A decline occurred following this, until a slight increase occurred in the WCPFC-CA in the last 2 years. Dynamic depletion shows a more consistent decrease over time, accelerating post-2000. Some of the unusual features of the previous assessment concerning a recruitment "dip" and sharp drop in $SB_t/SB_{F=0(t)}$ towards the end of the assessment period appear to have been resolved, or at least moderated, through the extension of the data series for the assessment. Estimates of uncertainty for the diagnostic case model as far back as the 1970s are relatively low and the various diagnostics indicated the model to be reasonably well fit, stable, without major data conflicts or retrospective patterns, and provide similar results when various data types are removed.

The main conclusions of this assessment are summarised as follows:

- The assessment indicates the stock is not overfished, and there was zero probability of the stock being below $20\% SB_{F=0}$. Estimates of $SB_{\text{recent}}/SB_{F=0}$ (with estimation uncertainty included) were 0.48, 0.36, and 0.62 for median, 10^{th}, and 90^{th} percentiles, respectively.
- The assessment indicates that the stock was not subject to overfishing and there was zero probability of $F_{\text{recent}}/F_{\text{MSY}}$ being above 1 with $F_{\text{recent}}/F_{\text{MSY}}$ estimates (with estimation uncertainty included) of 0.18, 0.06, and 0.44 for median, 10^{th}, and 90^{th} percentiles, respectively.
- The ratio of $SB_{\text{recent}}/SB_{F=0}$:iTRP was approximately 1 (median 0.952; 0.924 – 0.986, 10th and 90th percentiles).

11.4 Recommendations for further research and software development

- Continue and refine as necessary the CKMR study for albacore, and develop the necessary stock assessment model software module to incorporate CKMR data into the integrated as-

sessment model.

- Additional collection of otolith and/or other age samples to better characterise the distribution of age-at-length across different source fisheries. Further evaluation of the existing otolith sample to examine the large positive age residuals at larger lengths.
- Develop a logistic growth model for MULTIFAN-CL.
- Investigate the potential importance of sex structure, and identify the potential bias of ignoring sex structure, in the South Pacific albacore assessment.
- Consider and test alternative parameterisations of selectivity for the assessment, to address issues noted in this assessment regarding selectivity parameter correlation and convergence at bounds.
- Continue the development of the SEAPODYM model for South Pacific albacore and its alignment with the stock assessment.
- Further develop the longline CPUE standardisation model, in particular to investigate a habitat-based model for better estimation of effective longline effort.
- Further work on the quality of longline LF data to address issues of data conflict within the various fisheries data sets identified in this assessment.

12 Acknowledgements

We thank the various fisheries agencies and CCMs for the provision of the catch, effort and size frequency data used in this analyses, and gratefully acknowledge the hard work of the observers working throughout the region to gather data and biological samples. We also acknowledge the support from the Pacific European Union Marine Partnership (PEUMP) for various areas of work on tuna data that contributed to this assessment. We thank Jon Brodziak for his assistance in implementing the steepness prior. We thank Inna Senina for providing the movement probabilities from SEAPODYM. We thank Fabrice Bouye for ensuring our Condor flock was up and flying when we needed it most. Sam McKechnie, as always provided valuable advice, reviews and reality checks along the way. We thank contributors for their recommendations of the 2022 peer review of the WCPFC yellowfin tuna assessment, comments received on the previous stock assessment, recommendations of the CAPAM "Tuna Stock Assessment Good Practices Workshop", recommendations received from the 2024 SPC Pre-Assessment Workshop, and follow-up inputs from PAW participants. The assessment was supported by the WCPFC secretariat through the funding contributions of CCMs.

13 References

Anderson, G., Hampton, J., Smith, N., and Rico, C. (2019). Indications of strong adaptive population genetic structure in albacore tuna (*Thunnus* alalunga) in the southwest and central Pacific Ocean. *Ecol Evol*, 9(18):10354–10364.

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

Bigelow, K., Musyl, M. K., Poisson, F., and Kleiber, P. (2006). Pelagic longline gear depth and shoaling. *Fisheries Research*, 77(2):173–183.

Bigelow, K. A., Hampton, J., and Miyabe, N. (2002). Application of a habitat-based model to estimate effective longline fishing effort and relative abundance of pacific bigeye tuna *Thunnus obesus*. *Fisheries Oceanography*, 11(3):143–155.

Brodziak, J., Lee, H.-H., and Mangel, M. (2011). Probable values of stock-recruitment steepness for north pacific albacore tuna. Technical report, ISC/11/BILLWG-2/11 Probable.

Brown, C. J., Desbiens, A., Campbell, M. D., Game, E. T., Gilman, E., Hamilton, R. J., Heberer, C., Itano, D., and Pollock, K. (2021). Electronic monitoring for improved accountability in western pacific tuna longline fisheries. *Marine Policy*, 132:104664.

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., Punt, A. E., Chang, Y.-J., Maunder, M. N., and Piner, K. R. (2017). Can diagnostic tests help identify model misspecification in integrated stock assessments? *Fisheries Research*, 192:28–40.

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K. R., et al. (2021). A cookbook for using model diagnostics in integrated stock assessments. *Fisheries Research*, 240:105959.

Castillo Jordan, C., Hampton, J., Ducharme-Barth, N., Xu, H., Vidal, T., Williams, P., Scott, F., Pilling, G., and Hamer, P. (2021). Stock assessment of South Pacific albacore tuna. Technical Report WCPFC-SC17-2021/SA-WP-02.

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.

Davies, N., Fournier, D., Bouye, F., and Hampton, J. (2022). Developments in the Multifan-CL software 2021-2022. Technical Report WCPFC-SC18-2022/SA-IP-03.

Day, J., Magnusson, A., Teears, T., Hampton, J., Davies, N., Castillo Jordan, C., Peatman, T., Scott, R., Scutt-Phillips, J., McKechnie, S., Scott, F., Yao, N., 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.

Ducharme-Barth, N., Castillo-Jordan, C., Hampton, J., Williams, P., Pilling, G., and Hamer, P. (2021). Stock assessment of swordfish in the southwest Pacific Ocean. Technical Report WCPFC-SC17-2021/SA-WP-04.

Ducharme-Barth, N. and Vincent, M. (2021). Focusing on the front end: A framework for incorporating in uncertainty in biological parameters in model ensembles of integrated stock assessments. Technical Report WCPFC-SC17-2021-SA-WP-05.

Farley, J., Krusic-Golub, K., and Eveson, P. (2021). Ageing of South Pacific Albacore: (Project 106). Technical Report WCPFC-SC17-SA-IP-10.

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.

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

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

Farley, J. H., Williams, A. J., Hoyle, S. D., Davies, C. R., and Nicol, S. J. (2013b). Reproductive dynamics and potential annual fecundity of south pacific albacore tuna (*Thunnus alalunga*). *PLoS One*, 8(4):e60577.

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. A., Hampton, J., and Sibert, J. R. (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(9):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. C. (2017). Revisiting data weighting in fisheries stock assessment models. *Fisheries Research*, 192:5–15.

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.

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

Goethel, D. R. and Berger, A. M. (2017). Accounting for spatial complexities in the calculation of biological reference points: effects of misdiagnosing population structure for stock status indicators. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(11):1878–1894.

Hamel, O. S. and Cope, J. M. (2022). Development and considerations for application of a longevity-based prior for the natural mortality rate. *Fisheries Research*, 256:106477.

Hamer, P. (2024). Summary report from the 2024 spc pre-assessment workshop. Technical Report WCPFC-SC20-2024/SA-IP-01, Manilla, Phillippines, 14–21 August 2024.

Hampton, J. and Fournier, D. A. (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(7):937–963.

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

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

Hoyle, S. (2024). Effort creep in longline and purse seine cpue and its application in tropical tuna assessments. Technical Report WCPFC-SC20-2024/SA-IP-19, Manilla, Phillippines, 14–21 August 2024.

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

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

Hoyle, S. D., Williams, A. J., Minte-Vera, C. V., and Maunder, M. N. (2023). 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, Busan, Republic of Korea, 7–15 August 2012.

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

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

Kinney, M. and Teo, S. (2016). Meta-analysis of north Pacific albacore tuna natural mortality. Technical Report ISC/16/ALBWG-02/07.

Kleiber, P., Fournier, D., Hampton, J., Davies, N., Bouye, F., and Hoyle, S. (2019). MULTIFAN-CL User's Guide. Technical report.

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

Langley, A. D. and Hampton, J. (2005). Stock assessment of albacore tuna in the South Pacific Ocean. Technical Report WCPFC-SC1-2005/SA-WP-03, Noumea, New Caledonia, 8–19 August 2005.

Lehodey, P., Senina, I., Nicol, S., and Hampton, J. (2015). Modelling the impact of climate change on South Pacific albacore tuna. *Deep Sea Research Part II: Topical Studies in Oceanography*, 113:246–259.

Lennert-Cody, C. E., Maunder, M. N., Aires-da Silva, A., and Minami, M. (2013). Defining population spatial units: Simultaneous analysis of frequency distributions and time series. *Fisheries Research*, 139:85–92.

Lennert-Cody, C. E., Minami, M., Tomlinson, P. K., and Maunder, M. N. (2010). Exploratory analysis of spatial–temporal patterns in length–frequency data: An example of distributional regression trees. *Fisheries Research*, 102(3):323–326.

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

Lorenzen, K. (2022). Size-and age-dependent natural mortality in fish populations: Biology, models, implications, and a generalized length-inverse mortality paradigm. *Fisheries Research*, 255:106454.

Macdonald, J., Anderson, G., Krusic-Golub, K., Prioul, F., Cuewapuru, C., Hosken, M., Barthelemy, V., Raoulx, T., Grewe, P., Nicol, S., and Panapa, A. (2024a). Spatial structure and regional connectivity of sp-alb tuna stocks in the wcpo and epo. Technical Report WCPFC-SC20-2024/SA-IP-04, Manilla, Phillippines, 14–21 August 2024.

Macdonald, J. I., Farley, J. H., Clear, N. P., Williams, A. J., Carter, T. I., Davies, C. R., and Nicol, S. J. (2013). Insights into mixing and movement of south pacific albacore derived from trace elements in otoliths. *Fisheries Research*, 148:56–63.

Macdonald, P., Bell, L., Scheiter, M., Ghergariu, C., Sanchez, R., Contreras, P., Judd, M., J., H., Potts, J., Day, J., Vidal, T., and Williams, P. (2024b). Project 90 update: Better data on fish weights and lengths for scientific analyses. Technical Report WCPFC-SC20-2024/ST-IP-04, Manilla, Phillippines, 14–21 August 2024.

Magnusson, A., Day, J., Teears, T., Hampton, J., Davies, N., Castillo Jordan, C., Peatman, T., Scott, R., Scutt-Phillips, J., McKechnie, S., Scott, F., Yao, N., 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.

Matsubara, N., Aoki, Y., Aoki, A., and Kiyofuji, H. (2024). Lower thermal tolerance restricts vertical distributions for juvenile albacore tuna *Thunnus alalunga* in the northern limit of their habitats. *Frontiers in Marine Science*, 11:1353918.

Maunder, M. N. and Piner, K. R. (2015). Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science*, 72(1):7–18.

Maunder, M. N., Thorson, J. T., Xu, H., Oliveros-Ramos, R., Hoyle, S. D., Tremblay-Boyer, L., Lee, H. H., Kai, M., Chang, S.-K., Kitakado, T., Albertsen, C. M., Minte-Vera, C. V., Lennert-Cody, C. E., Aires-da Silva, A. M., and Piner, K. R. (2020). The need for spatio-temporal modeling to determine catch-per-unit effort based indices of abundance and associated composition data for inclusion in stock assessment models. *Fisheries Research*, 229:105594.

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

Minte-Vera, C. V., Maunder, M. N., Aires-da Silva, A. M., Satoh, K., and Uosaki, K. (2017). Get the biology right, or use size-composition data at your own risk. *Fisheries research*, 192:114–125.

Neubauer, P. and Hill-Moana, T. (2024). Characterisation, cpue and length-composition analyses of the new zealand albacore fishery. Technical Report WCPFC-SC20-2024/SA-IP-08, Manilla, Phillippines, 14–21 August 2024.

Neubauer, P., Kim, K., A'mar, T., and Large, K. (2023). Addressing uncertainty in wcpfc stock assessments: Review and recommendations from WCPFC project 113. Technical Report WCPFC/SC19-SA-WP-10, Koror, Palau, 16–24 August 2023.

Nikolic, N., Montes, I., Lalire, M., Puech, A., Bodin, N., Arnaud-Haond, S., Kerwath, S., Corse, E., Gaspar, P., Hollanda, S., Bourjea, J., West, W., and Bonhommeau, S. (2020). Connectivity and population structure of albacore tuna across southeast Atlantic and southwest Indian Oceans inferred from multidisciplinary methodology. *Sci Rep*, 10(1):15657.

Nikolic, N., Morandeau, G., Hoarau, L., West, W., Arrizabalaga, H., Hoyle, S., Nicol, S. J., Bourjea, J., Puech, A., Farley, J. H., Williams, A. J., and Fonteneau, A. (2017). Review of albacore tuna, *Thunnus alalunga*, biology, fisheries and management. *Rev Fish Biol Fisheries*, 27(4):775–810.

Pilling, G., Scott, R., Hamer, P., Scott, F., and Yao, N. (2024). Recalibration of the adopted south pacific albacore interim target reference point and review of wcpfc20 requested options. Technical Report WCPFC-SC20-2024/MI-WP-03, Manilla, Phillippines, 14–21 August 2024.

Potts, J., Castillo-Jordan, C., Day, J., Hamer, P., and Teears, T. (2024). Analysis of longline size frequency data for the 2024 south pacific albacore and wcpo striped marlin assessments. Technical Report WCPFC-SC20-2024/SA-IP-03, Manilla, Phillippines, 14–21 August 2024.

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

Renck, C., Talley, D., Wells, D., and Dewar, H. (2014). Regional growth patterns of juvenile albacore (*Thunnus alalunga*) in the eastern North Pacific. Technical Report CalCOFI Rep., Vol. 55.

Senina, I. N., Lehodey, P., Hampton, J., and Sibert, J. (2020). Quantitative modelling of the spatial dynamics of South Pacific and Atlantic albacore tuna populations. *Deep Sea Research Part II: Topical Studies in Oceanography*, 175:104667.

Stewart, I. J. and Martell, S. J. (2014). A historical review of selectivity approaches and retrospective patterns in the Pacific halibut stock assessment. *Fisheries Research*, 158:40–49.

Teears, T., Castillo Jordon, C Day, J., Hampton, J., Magnusson, Peatman, T., Vidal, T., Williams, P., and Hamer, P. (2024). Background analyses and data inputs for the 2024 south pacific albacore tuna stock assessment. Technical Report WCPFC-SC20-2024/SA-IP-05.

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 SC19-SA-WP-03.

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., Hampton, J., McKechnie, S., and Pilling, G. (2018). Stock assessment of South Pacific albacore tuna. Technical Report WCPFC-SC-14-2018/SA-WP-05, Busan, South Korea, 8-16 August 2018.

Vidal, T., Castillo Jordan, C., Peatman, T., Ducharme-Barth, N., Xu, H., Williams, P., Lennert-Cody, C., and Hamer, P. (2021). Background analysis and data inputs for the 2021 South Pacific albacore tuna stock assessment. Technical Report WCPFC-SC17-SA-IP-03.

Williams, A. J., Allain, V., Nicol, S. J., Evans, K. J., Hoyle, S. D., Dupoux, C., Vourey, E., and Dubosc, J. (2015). Vertical behavior and diet of albacore tuna (*Thunnus alalunga*) vary with latitude in the South Pacific Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 113:154–169.

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

14 Tables

Table 1: Summary of data available for the 2024 South Pacific albacore assessment by fisheries and year defined in Table 2.

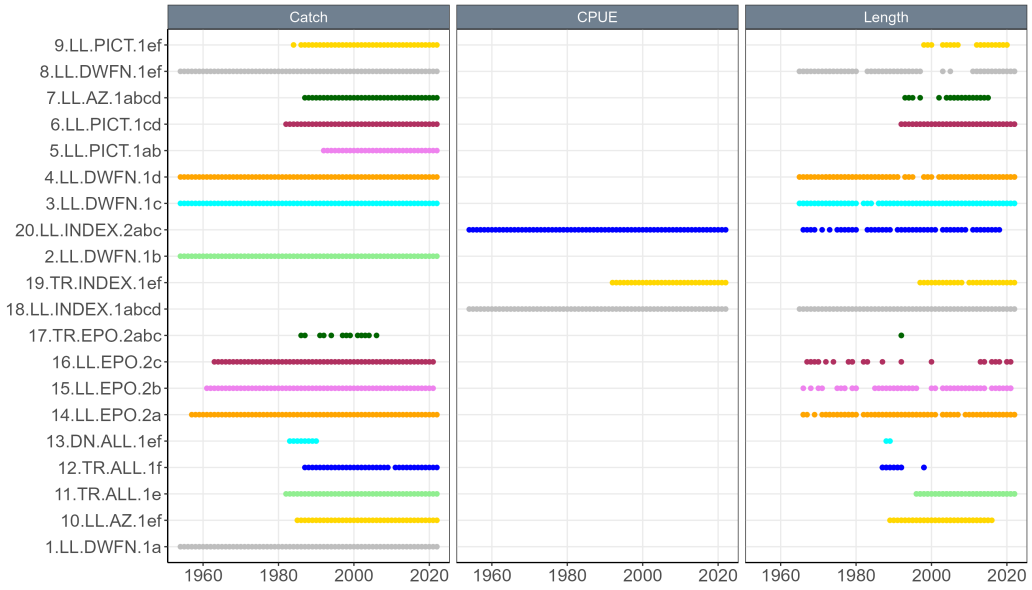


Table 2: Definition of fisheries for the 2024 MULTIFAN-CL South Pacific albacore tuna stock assessment, refer to Figure 3.

Fishery Number	Gear	Model Code-Fleets	Flags	Model Region	Fleet area
1	LL	1.LL.DWFN.1a	DWFN	1	a
2	LL	2.LL.DWFN.1b	DWFN	1	b
3	LL	3.LL.DWFN.1c	DWFN	1	c
4	LL	4.LL.DWFN.1d	DWFN	1	d
5	LL	5.LL.PICT.1ab	PICT	1	a, b
6	LL	6.LL.PICT.1cd	PICT	1	c, d
7	LL	7.LL.AZ.1abcd	AU/NZ	1	a, b, c, d
8	LL	8.LL.DWFN.1ef	DWFN	1	e, f
9	LL	9.LL.PICT.1ef	PICT	1	e, f
10	LL	10.LL.AZ.1ef	AU/NZ	1	e, f
11	TR	11.TR-ALL.1e	ALL	1	e
12	TR	12.TR.ALL.1f	ALL	1	f
13	DN	13DN.ALL.1ef	ALL	1	e, f
14	LL	14.LL.EPO.2a	ALL	2	a
15	LL	15.LL.EPO.2b	ALL	2	b
16	LL	16.LL.EPO.2c	ALL	2	c
17	TR	17.TR.EPO.2abc	ALL	2	a, b, c
18	LL	18.LL.INDEX.1abcd	INDEX	1	a,b,c,d
19	TR	19.TR.INDEX.1ef	INDEX	1	e,f
20	LL	20.LL.INDEX.2abc	INDEX	2	a,b,c

DWFN: BZ, CN, ES, JP, JP-DW, KR, SU, TW, US, VN, VU
PICT: CK, FJ, FM, ID, KI, NC, NU, PF, PG, PH, SB, TO, TV, WS

Table 3: Parameter structure for the diagnostic case model. Fisheries are numbered as per Table 2. Region 1 = WCPFC-CA, region 2 = EPO.

Parameter	Region	Fishery	Group	Age- related	No. seasons	No. of parameters	Fixed/ Estimated	Constraints	Comments
Selectivity									All selectivities are splines.
	1	1	1	3	4	12	Est		
	1	2	2	3	4	12	Est		
	1	3	1	3	4				
	1	4	2	3	4				
	1	5	3	3	4	12	Est		
	1	6	3	3	4				
	1	7	4	3	4	12	Est		
	1	8	5	4	1	4	Est		
	1	9	6	4	1	4	Est		
	1	10	7	4	1	4	Est		
	1	11	8	3	1	3	Est		
	1	12	9	3	1	3	Est		
	1	13	10	3	1	3	Est		
	2	14	11	3	2	6	Est		
	2	15	12	3	1	3	Est		
	2	16	13	3	1	3	Est		
	2	17	9	3	1				
	1	18	14	3	1	3	Est	Non-decreasing with age	
	1	19	15	3	1	3	Est		
	2	20	16	3	1	3	Est	Non-decreasing with age	
Recruitment									
Deviations	1		1			67	Est		Recruitment in 2022 set to historical average. $\sigma = 0.7$.
	2		1						Time-series variation shared across regions.
Regional proportions						1	Fixed		0.82, 0.18 as per SEAPODYM estimates.
Natural mortality						2	Fixed		Lorenzen shape (inverse mean length) and scale.
Growth									
L1						1	Fixed		
L12						1	Est		
k						1	Est		
Offsets						3	Est		Offsets for age classes 2, 3 and 4.
V1						1	Est		
V2						1	Fixed		
Movement	1		12	4		48	Fixed		SEAPODYM estimates.
	2		12	4		48	Fixed		
Population scale						1	Est		
Stock-recruitment									Beverton and Holt stock-recruitment relationship
Scale						1	Est		
Steepness						1	Fixed		
TOTAL						102	Fixed		
						165	Est		

Table 4: Description of symbols used in the yield and stock status analyses. For the purpose of this assessment, "recent" for F is the average over the period 2018–2021 and for SB is the average over the period 2019–2022 and "latest" is 2022.

Symbol	Description
F_{recent}	Average fishing mortality-at-age for a recent period (2018–2021)
$Y_{F_{\text{recent}}}$	Equilibrium yield at average fishing mortality for a recent period (2018–2021)
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 (2018–2021) relative to F_{MSY}
SB_{latest}	SB in the latest time period (2022)
SB_{recent}	SB for a recent period (2019–2022)
$SB_{F=0}$	Average SB predicted in the absence of fishing for the period 2012–2021
SB_{MSY}	SB that will produce the maximum sustainable yield (MSY)
$SB_{\text{MSY}}/SB_{F=0}$	SB that produces maximum sustainable yield (MSY) relative to the average SB predicted to occur in the absence of fishing for the period 2012–2021
$SB_{\text{latest}}/SB_{F=0}$	SB in the latest time period (2022) relative to the average SB predicted to occur in the absence of fishing for the period 2012–2021
$SB_{\text{latest}}/SB_{\text{MSY}}$	SB in the latest time period (2022) relative to that which will produce the maximum sustainable yield (MSY)
$SB_{\text{recent}}/SB_{F=0}$	SB for a recent period (2019–2022) relative to the average spawning biomass predicted to occur in the absence of fishing for the period 2012–2021
$SB_{\text{recent}}/SB_{\text{MSY}}$	SB for a recent period (2019–2022) relative to the SB that produces maximum sustainable yield (MSY)
$20\%SB_{F=0}$	WCPFC adopted limit reference point – 20% of SB in the absence of fishing average over years $t-10$ to $t-1$ (2012–2021)
$0.96 * SB_{2017-2019}/SB_{F=0}$	WCPFC adopted interim target reference point (iTRP) – 0.96 times the mean of: $SB_{2017}/SB_{F=0,2007-2016}$, $SB_{2018}/SB_{F=0,2008-2017}$ and $SB_{2019}/SB_{F=0,2009-2018}$

Table 5: Sensitivity of important reference point variables to the different model variants examined in the sensitivity analyses.

Sensitivity	Variant	$SB_{\text{recent}}/SB_{F=0}$	$F_{\text{recent}}/F_{\text{MSY}}$	$SB_{\text{recent}}/SB_{\text{MSY}}$
LL indices	WCPO.NORTH	0.421	0.366	2.21
	WCPO.SPAWN	0.416	0.368	2.16
Troll CPUE	Included	0.421	0.366	2.21
	Excluded	0.449	0.350	2.06
Troll LF	Included	0.421	0.366	2.21
	Excluded	0.449	0.281	2.48
CAAL weight	0.50	0.422	0.364	2.21
	0.75	0.421	0.366	2.21
	1.00	0.420	0.368	2.20
Movement	None	0.502	0.275	2.50
	0.5 x SEAPODYM	0.437	0.341	2.28
	SEAPODYM	0.421	0.366	2.21
	2 x SEAPODYM	0.408	0.388	2.15
	Full mixing	0.409	0.397	2.14
Effort creep	0	0.421	0.366	2.21
	0.50%	0.418	0.372	2.14
	1%	0.416	0.376	2.09
	1.50%	0.415	0.379	2.06
Recruit distribution R1:R2	70:30	0.420	0.368	2.20
	82:18	0.421	0.366	2.21
	90:10	0.422	0.364	2.21
No of age classes	10	0.442	0.346	2.31
	12	0.421	0.366	2.21
	15	0.426	0.355	2.23
Non-decreasing selectivity	F18, F20	0.421	0.366	2.21
	F7, F14	0.462	0.318	2.20
Time–block	None	0.421	0.366	2.21
	Selectivity	0.488	0.251	2.38
	Selectivity+catchability	0.480	0.260	2.25
Growth	VB with offsets	0.421	0.366	2.21
	Standard VB	0.430	0.310	2.58
Natural mortality	0.24	0.301	0.769	1.61
	0.36	0.458	0.283	2.40
	0.54	0.649	0.065	3.68
Steepness	0.65	0.396	0.637	1.71
	0.80	0.421	0.366	2.21
	0.95	0.436	0.172	3.36

Table 6: Fit statistics for sensitivity of selectivity and catchability models. AIC refers to Akaike Information Criterion.

Model	-Log likelihood	No. of parameters	AIC
Diagnostic model	-41,824.6	165	- 83,319.20
Selectivity TB	-41,868.3	251	- 83,234.60
Selectivity + catchability TB	-41,873.8	251	- 83,245.60

Table 7: The growth characteristics for the model adopted for the diagnostic case.

Parameter	Description	Treatment
$L1$	Mean length of age class 1	Fixed at a value (45.5cm) consistent with a mean age at recruitment of 9 mon.
$L12$	Mean length of age class 12	Estimated
k	Growth coefficient	Estimated
Offsets (3)	Deviations from VB growth for age classes 2, 3 and 4	Estimated
$V1$	Generic standard deviation of mean length-at-age	Estimated
$V2$	Change in mean length-at-age with age	Fixed at a value (0.547) that sets the SD of length for age class 1 to a value consistent with the first mode in the NZ troll length frequency data ($\sim$2.5cm)

Table 8: Comparison of the likelihood components of the diagnostic case model with a standard VB model without estimated offsets.

Likelihood component	Diagnostic (VB offsets)	Standard VB
CPUE indices	-170.0	-186.3
Length frequency	-42,457.3	-42,374.9
Conditional age at length	786.4	731.0
Penalties	16.3	13.2
Total	-41,824.6	-41,817.64
No. parameters	165	164

Table 9: Summary of reference points over the model ensemble, along with results incorporating estimation uncertainty. Note that these values do not include estimation uncertainty, unless otherwise indicated. Estimates for WCPFC-CA only are also provided.

	Mean	Median	Min	10%	90%	Max
F_{MSY}	0.15	0.16	0.10	0.12	0.18	0.20
f_{mult}	7.95	5.61	1.21	2.27	17.18	27.66
$F_{\text{recent}}/F_{\text{MSY}}$	0.22	0.18	0.04	0.06	0.44	0.82
MSY	113,308	101,100	62,120	74,018	176,330	202,400
SB_0	587,089	566,950	529,100	537,100	662,500	749,700
$SB_{F=0}$	724,200	711,059	665,389	674,633	788,312	857,071
SB_{latest}/SB_0	0.66	0.67	0.38	0.53	0.81	0.90
$SB_{\text{latest}}/SB_{F=0}$	0.54	0.54	0.29	0.41	0.70	0.78
$SB_{\text{latest}}/SB_{\text{MSY}}$	3.71	3.40	1.65	2.32	5.77	7.45
SB_{MSY}	111,738	110,950	65,140	80,350	142,690	172,600
SB_{MSY}/SB_0	0.19	0.20	0.11	0.13	0.24	0.27
$SB_{\text{MSY}}/SB_{F=0}$	0.15	0.16	0.10	0.11	0.19	0.22
$SB_{\text{recent}}/SB_{F=0}$	0.48	0.48	0.27	0.37	0.62	0.65
$SB_{\text{recent}}/SB_{\text{MSY}}$	3.30	3.06	1.54	2.10	5.23	6.34
$Y_{F_{\text{recent}}}$	74,531	74,375	61,760	67,731	83,023	86,180
$SB_{\text{latest}}/SB_{F=0}$:iTRP	1.065	1.051	0.961	1.015	1.139	1.213
$SB_{\text{recent}}/SB_{F=0}$:iTRP	0.952	0.952	0.899	0.924	0.986	1.016
WCPFC-CA only						
$SB_{\text{recent}}/SB_{F=0}$	0.49	0.48	0.27	0.37	0.62	0.66
Including estimation uncertainty						
$F_{\text{recent}}/F_{\text{MSY}}$	0.23	0.18	0.03	0.06	0.44	1.00
$SB_{\text{recent}}/SB_{F=0}$	0.48	0.48	0.23	0.36	0.62	0.77
$SB_{\text{recent}}/SB_{\text{MSY}}$	3.32	3.02	1.20	2.04	5.21	8.96

Note: Recalibrated value for iTRP = 0.50 (Pilling et al., 2024)

15 Figures

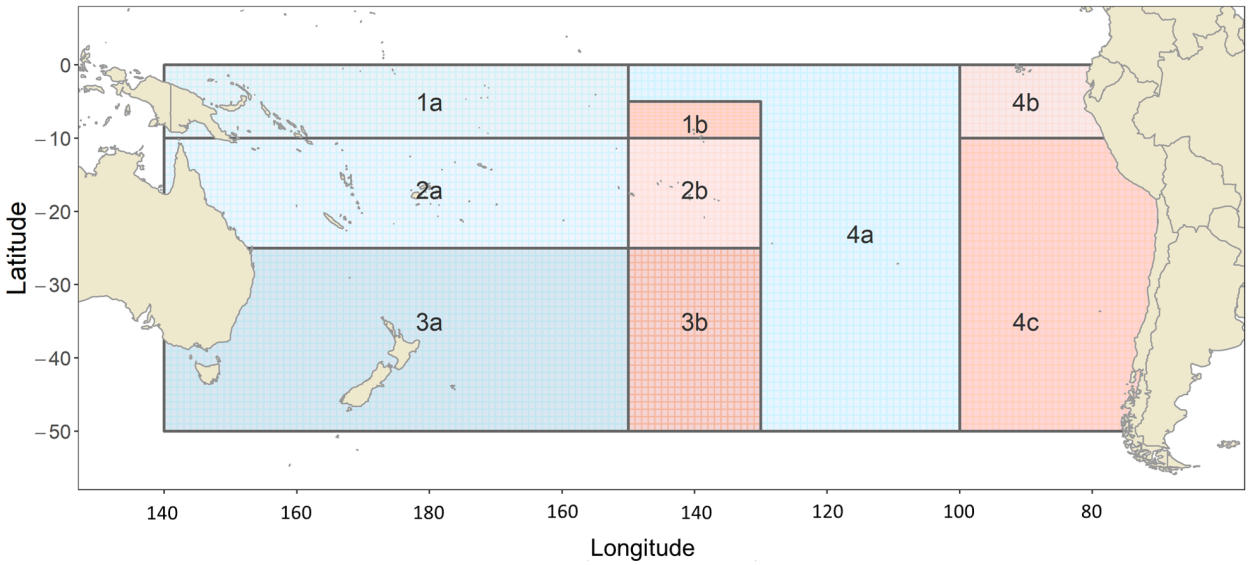


Figure 1: The geographical area covered by the 2021 stock assessment and the boundaries of the four model regions with addition of the sub-region applied for the EPO and overlap regions used for South Pacific-wide albacore assessment. The overlap region is the area between 130°–150° west demarcated by the dashed line.

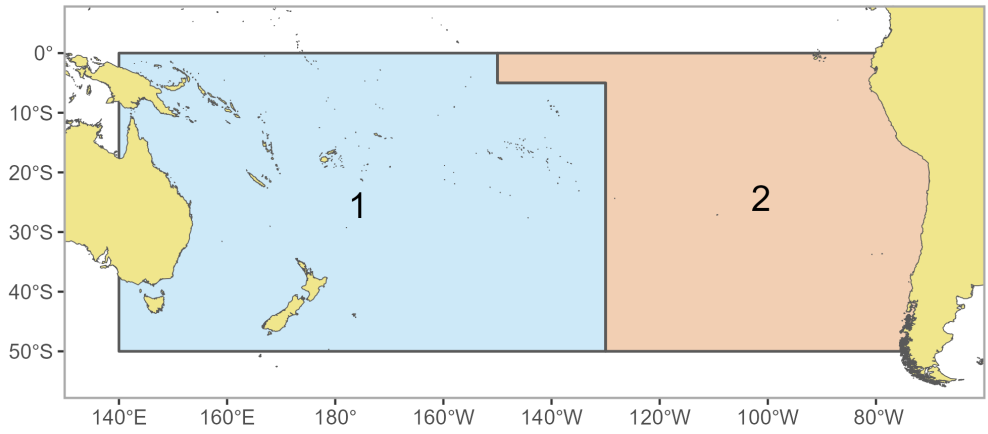


Figure 2: The geographical area covered by the stock assessment and the boundaries of the two model regions used for the South Pacific-wide 2024 albacore assessment.

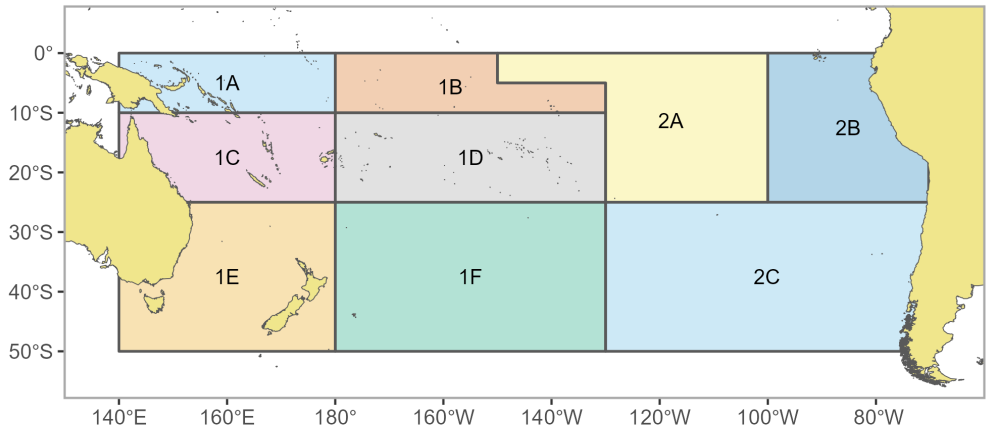
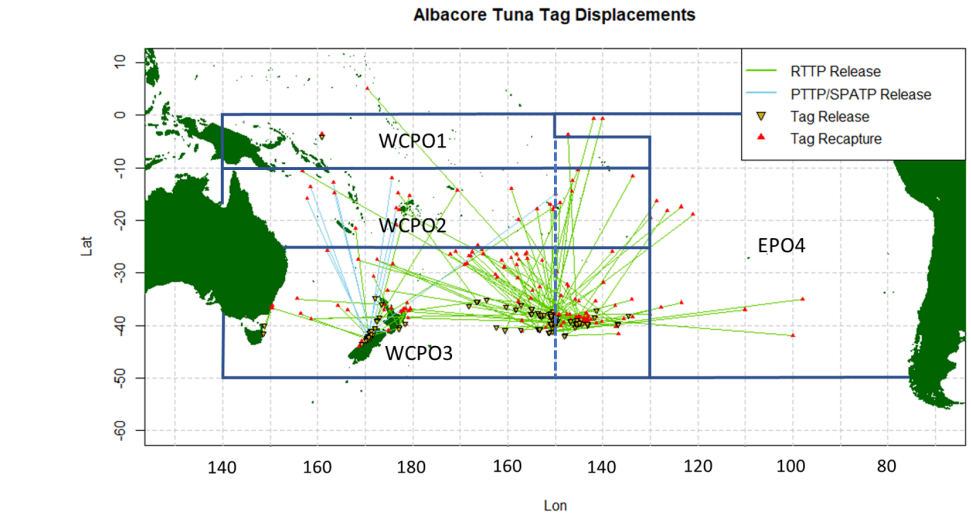


Figure 3: The geographical area boundaries of the nine fisheries areas used for the South Pacific-wide 2024 albacore assessment.



Summary

	RTTP-era											PTTP-era		
Year	1979	1980	1982	1986	1987	1988	1989	1990	1991	1992	→	2009	2010	2011
Released	19	1	4	861	1561	1274	1806	989	3701	5384		2766	111	5
Recaptured	0	0	0	3	11	3	6	7	56	83		32	1	1
Return rate				0.003	0.007	0.002	0.003	0.007	0.015	0.015		0.012	0.009	0.200

Figure 4: Map of tag displacements for South Pacific albacore tagged under the different tagging programs (top), and (bottom) table of tag-releases and recaptures for the various tagging programs 1979 to 1992.

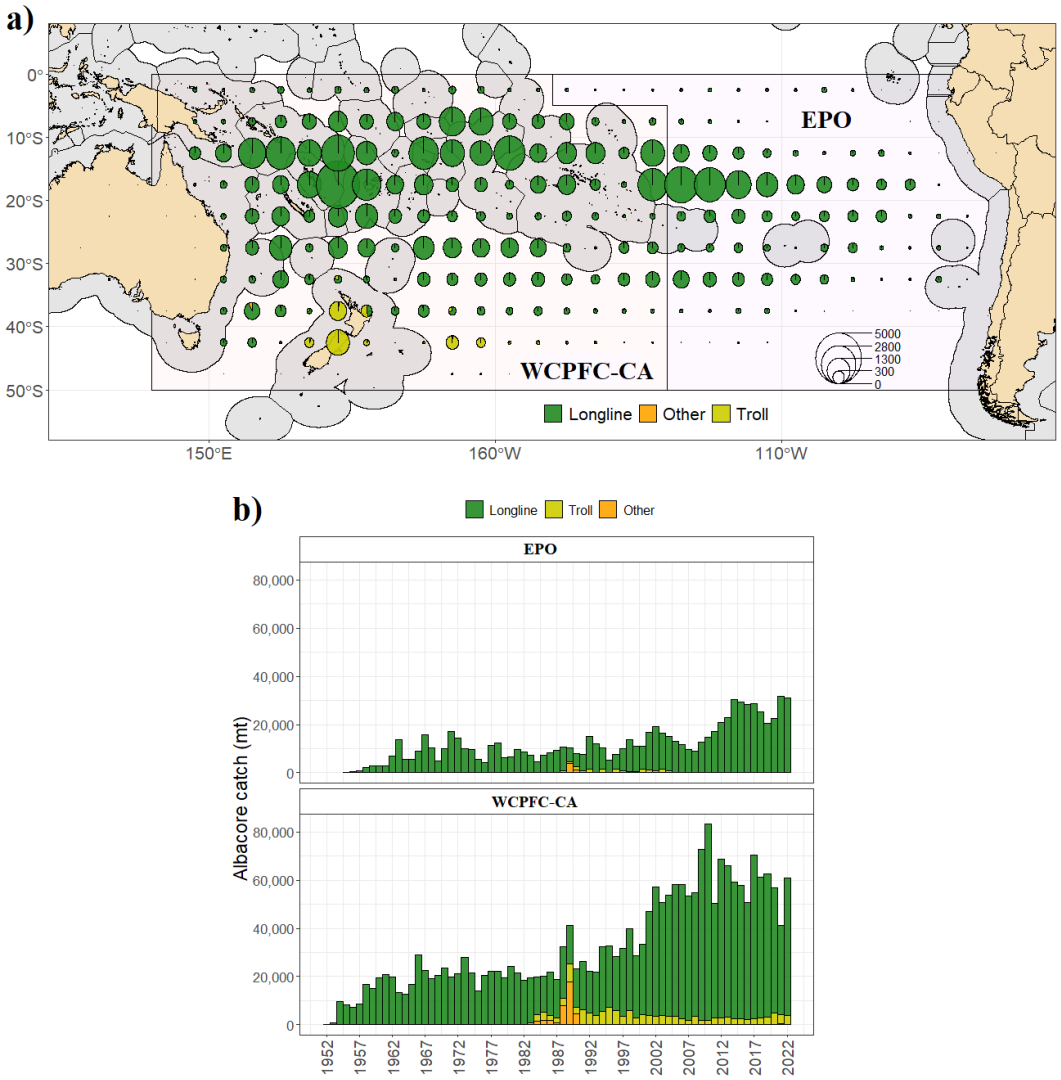


Figure 5: a) Spatial pattern of albacore catch by gear type over the last decade, and b) historical catches of albacore across the spatial extent from 1954-2022 by gear type.

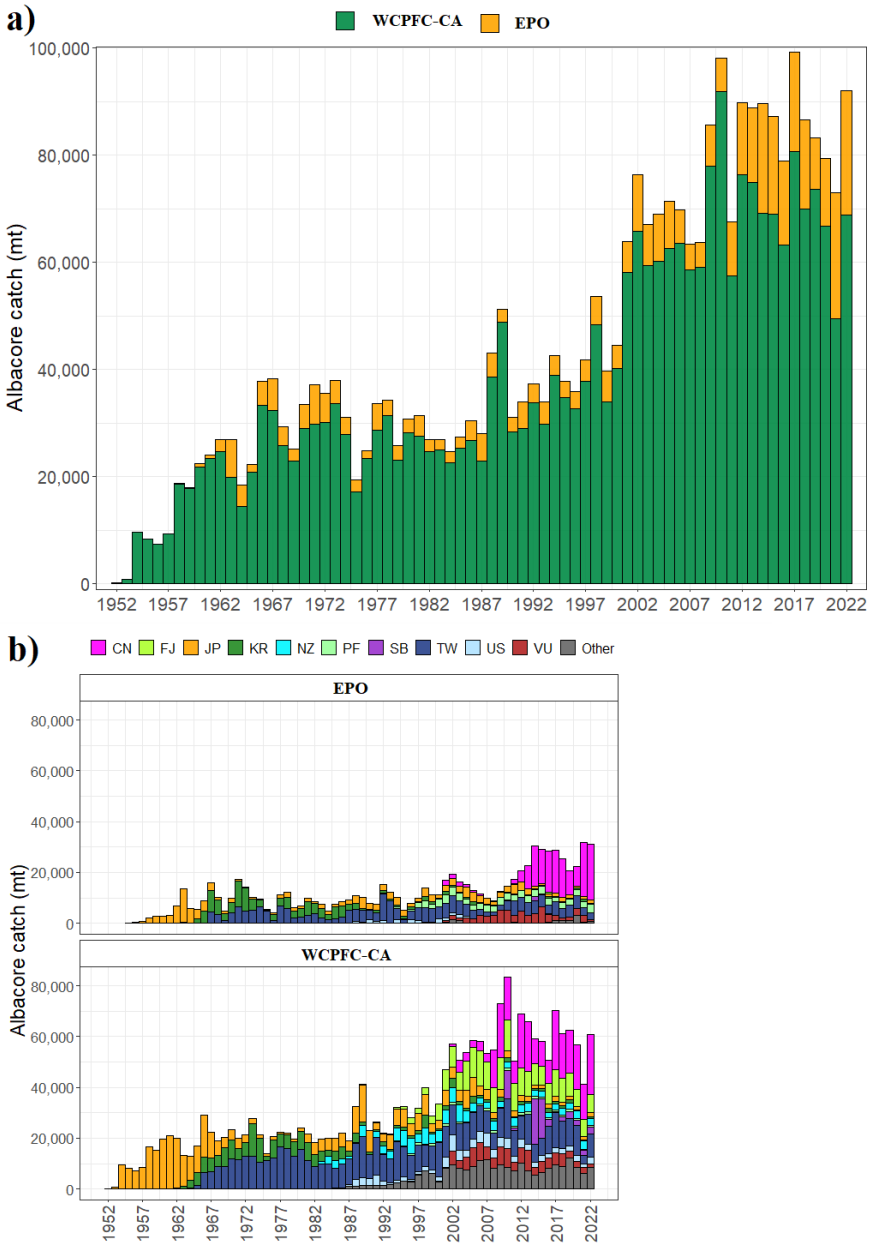


Figure 6: a) Annual catches of albacore from 1952-2022 separated by the WCPFC-CA and the IATTC (EPO) region, b) annual catches of albacore from 1952-2022 separated by flag for the WCPFC-CA and the IATTC (EPO) regions.

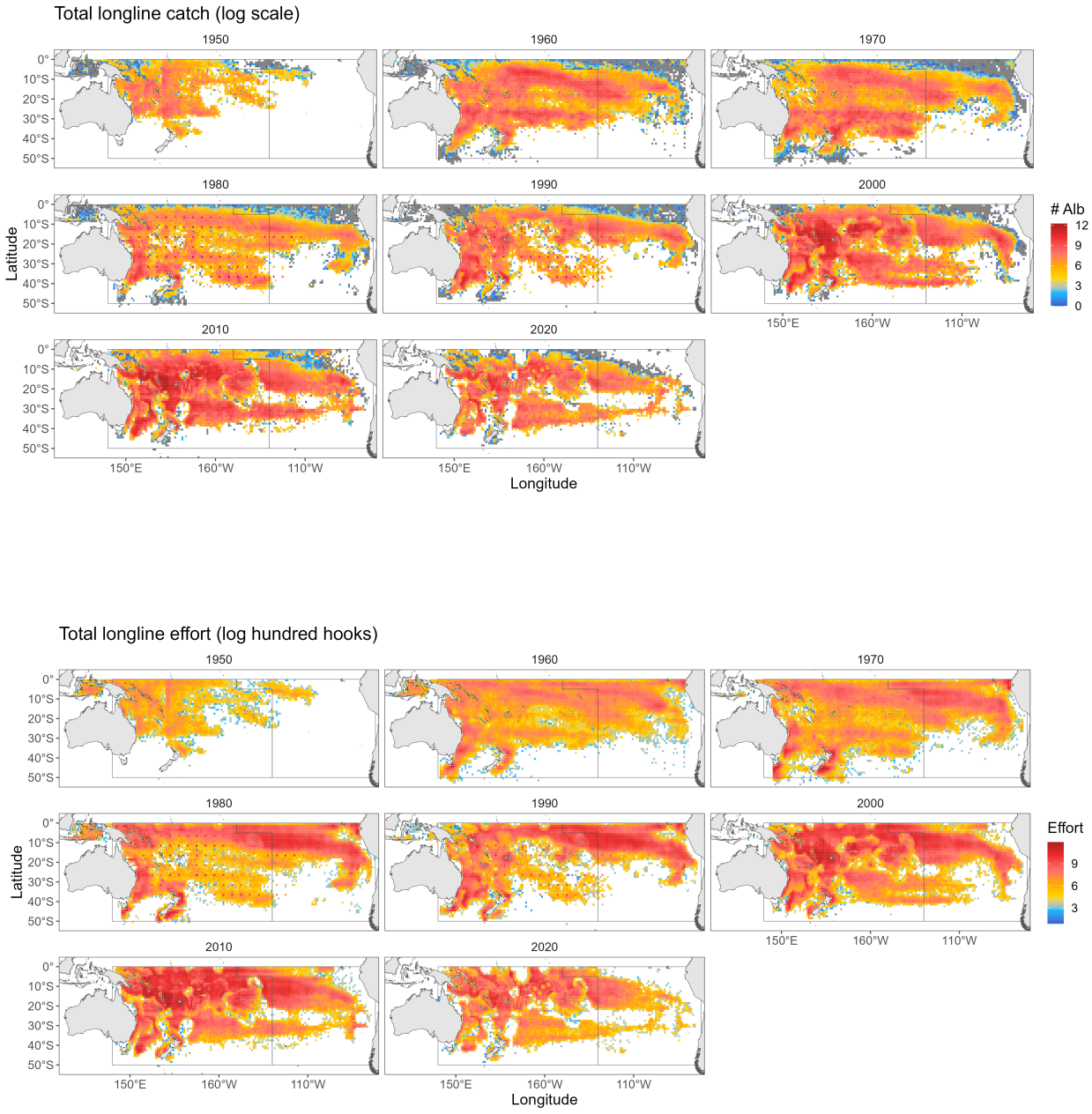


Figure 7: (top) Spatial distribution of albacore catches and (bottom) longline effort in the South Pacific by decade. Model regions indicated by black lines.

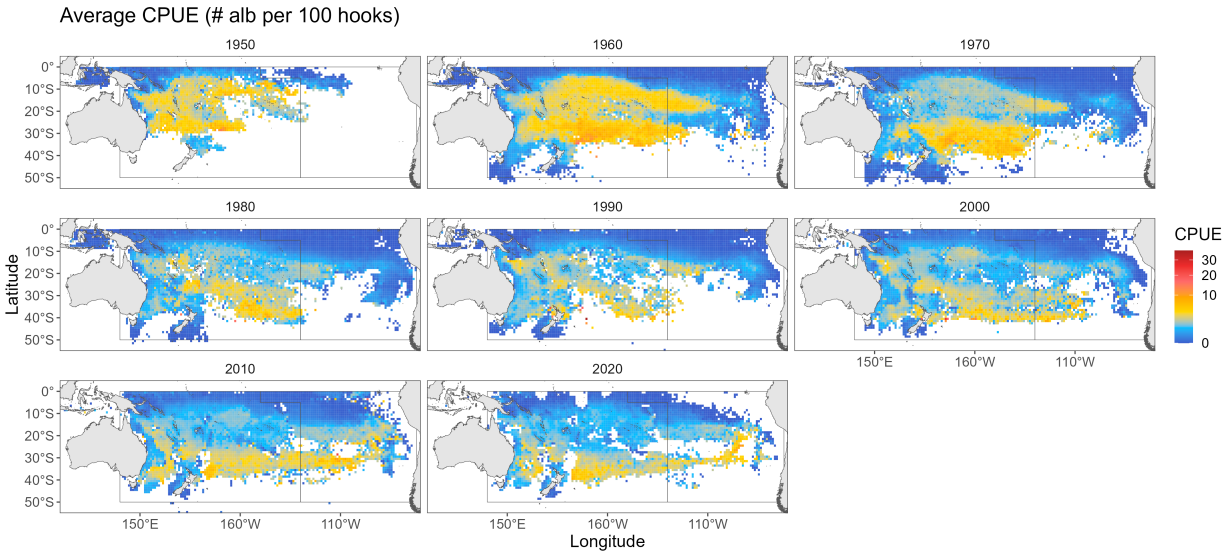


Figure 8: Spatial distribution of longline nominal albacore CPUE in the South Pacific by decade. Model regions indicated by black lines.

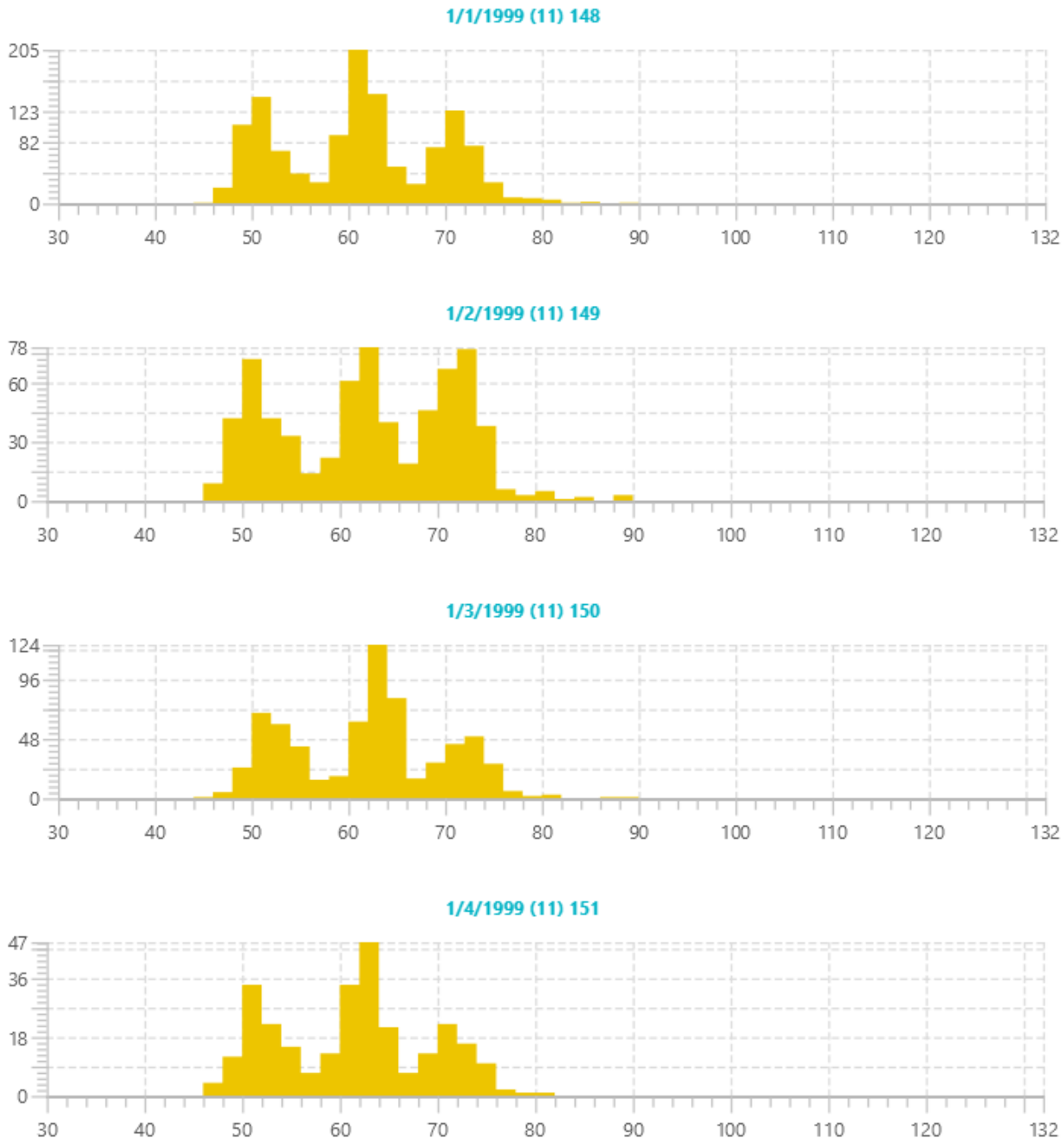


Figure 9: A segment of NZ troll fishery length-frequency data demonstrating strong modal structure representing annual age classes.

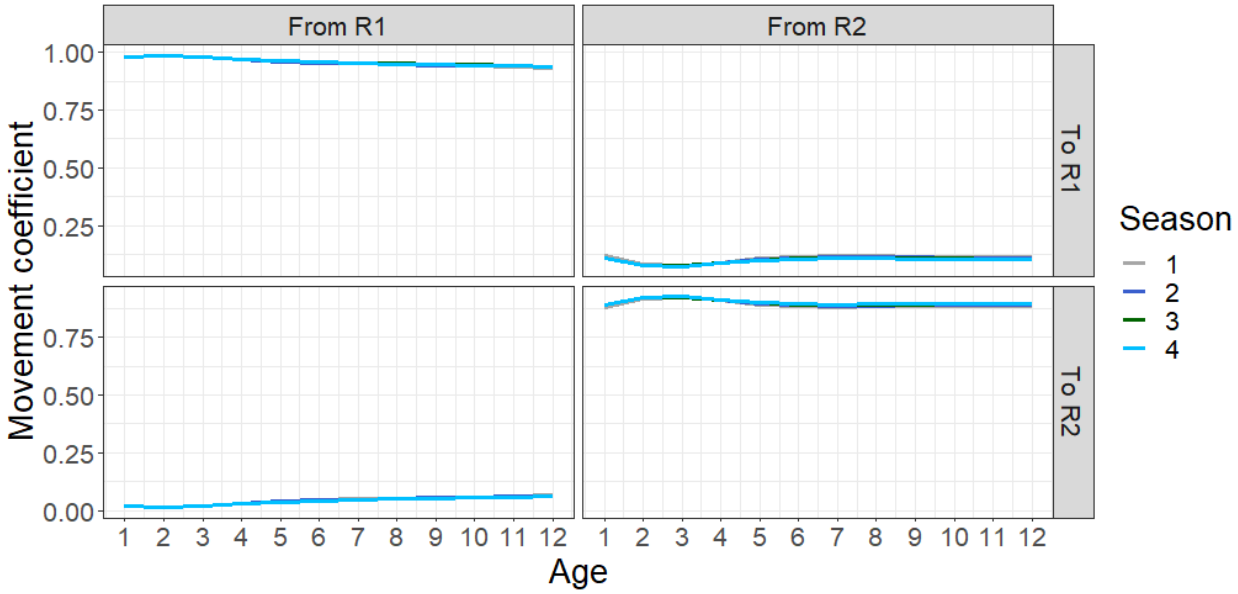


Figure 10: Regional movement parametrized by SEAPODYM estimates by age and season for the diagnostic case model.

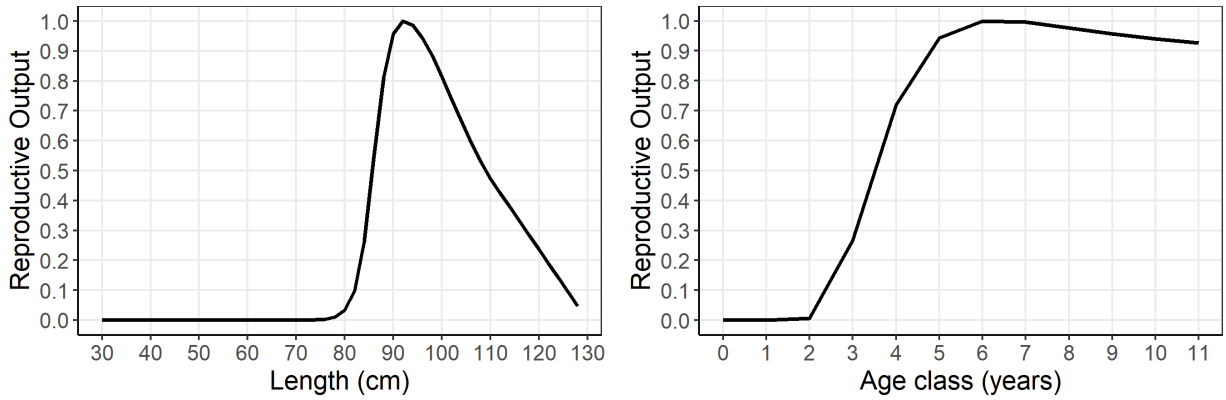


Figure 11: Reproductive potential ogives for South Pacific albacore by length (left) and converted to age (years; right).

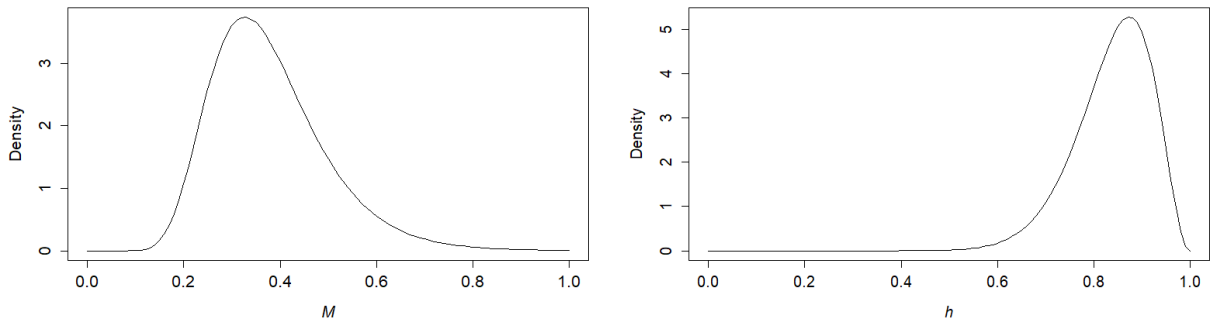


Figure 12: Assumed prior distributions of average M (a) and steepness rates (b) considered in the model ensemble uncertainty characterization.

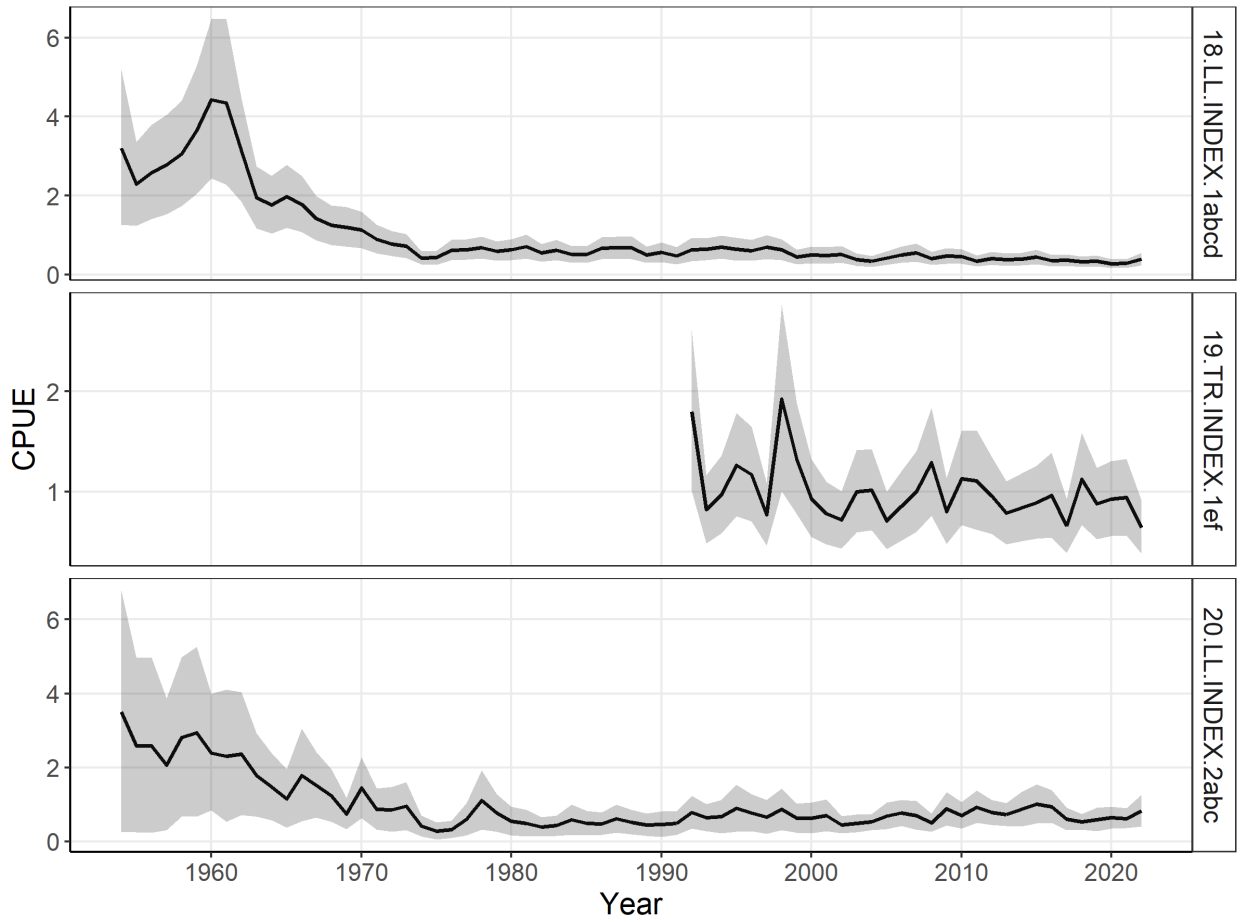


Figure 13: standardised relative abundance indices by region plotted with 95% confidence intervals (shading), from 1954–2022.

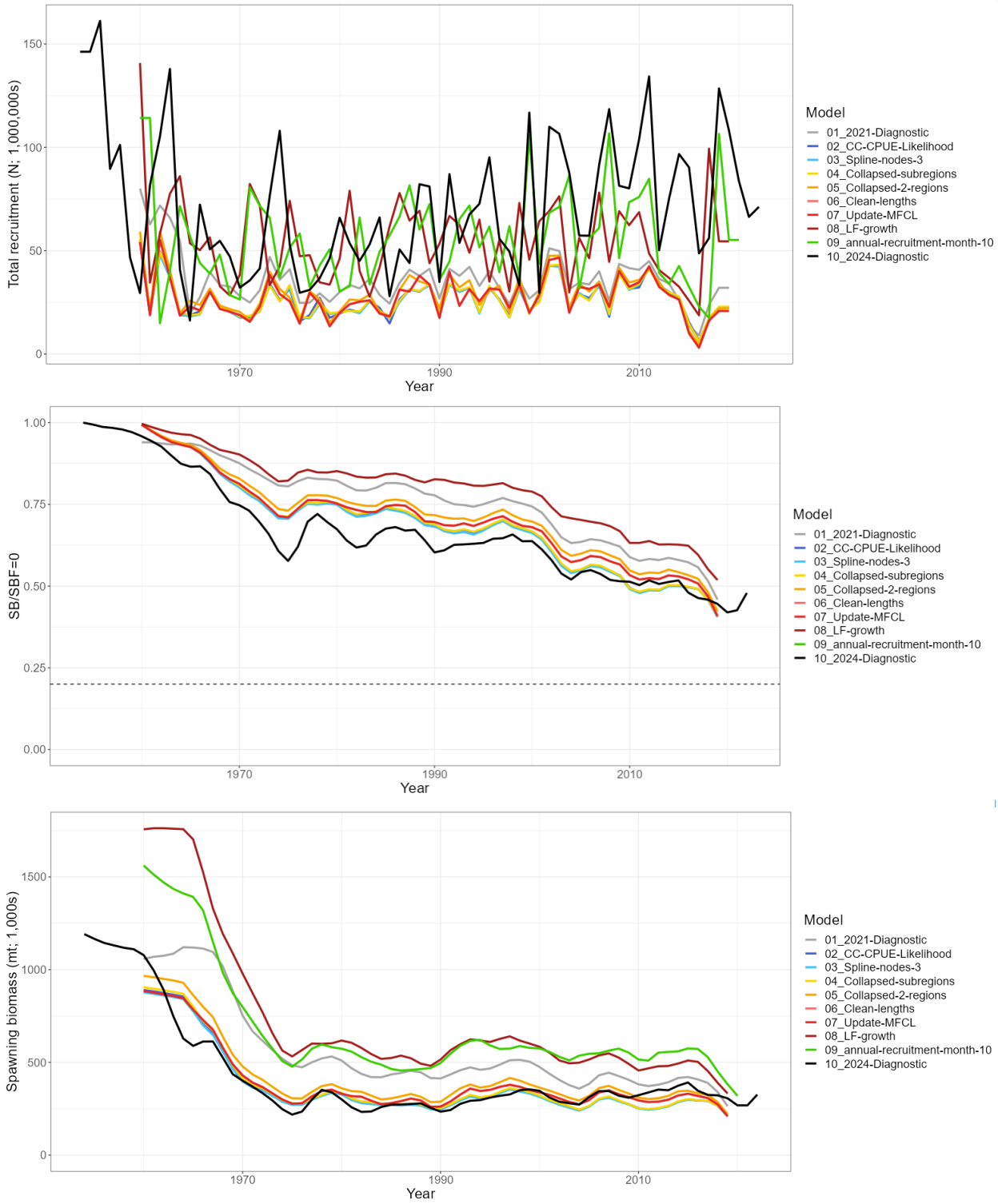


Figure 14: Model development change in recruitment, SB, and $SB_t/SB_{F=0(t)}$ from the 2021 reference case model to the 2024 diagnostic case model.

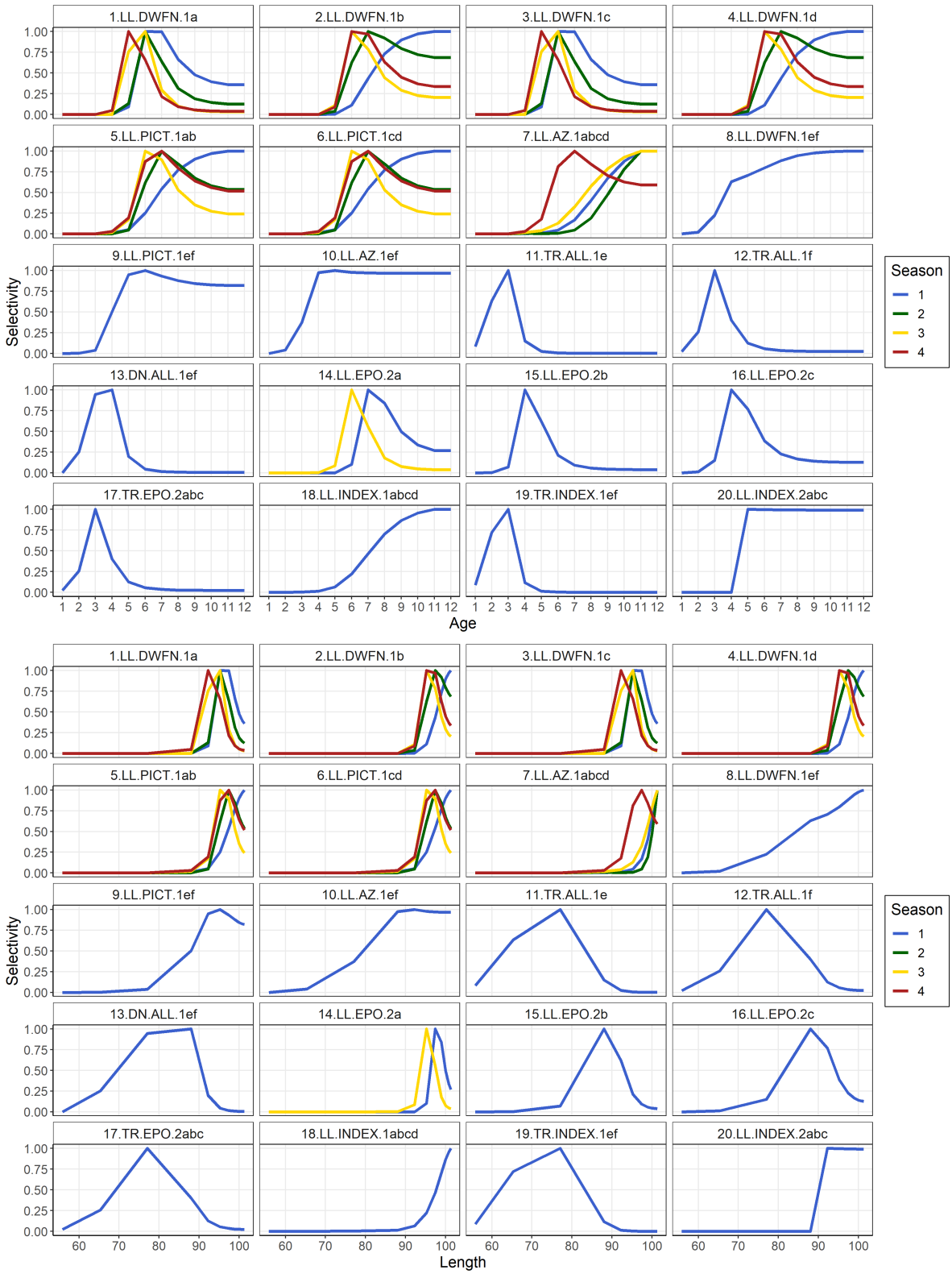


Figure 15: Estimated age-specific (top) and length-specific (bottom) selectivity coefficients by fishery for the diagnostic case model.

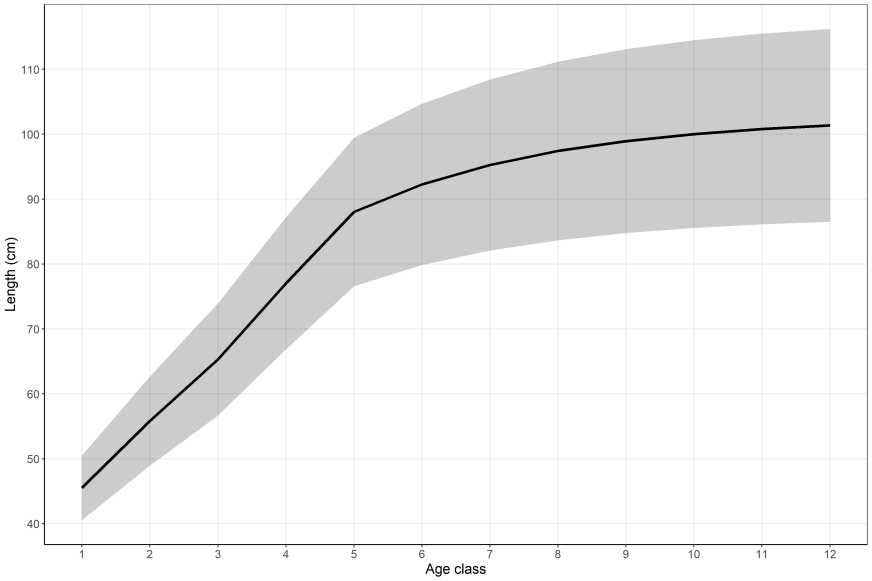


Figure 16: Estimated length-at-age (Fork Length) von Bertalanffy (VB) relationship with 95% confidence intervals from the diagnostic case model. Note ages 2–4 were modelled with estimated offsets from the VB curve.

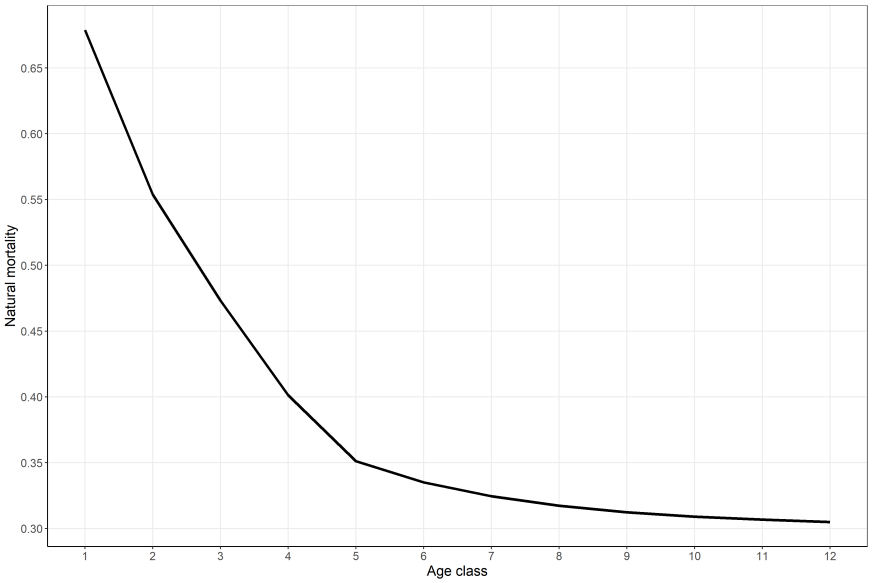


Figure 17: Estimated M-at-age from the von Bertalanffy (VB) growth parameters from the diagnostic case model. Average $M_{a=4:12} = 0.329$ and $M_{a=4:12}/M_{a=12} = 1.079$.

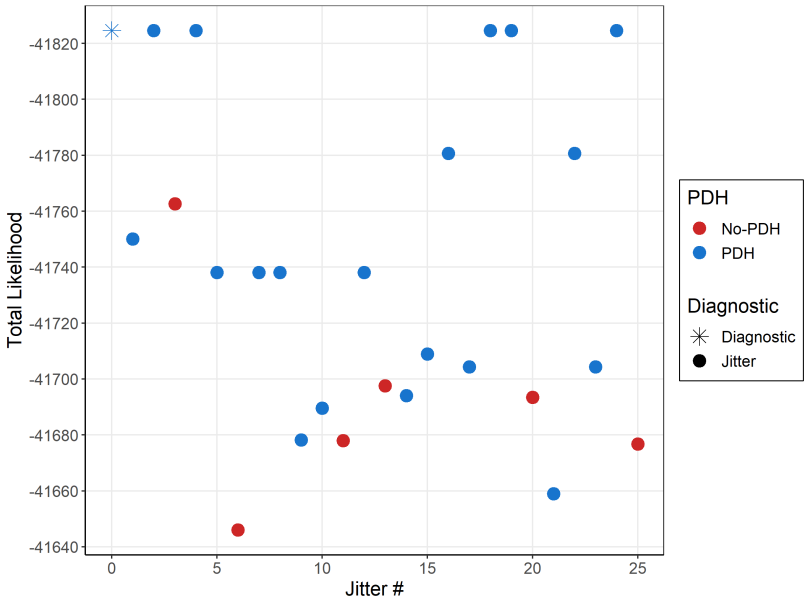


Figure 18: Negative log likelihood for 25 jittered models (coloured points; colours represent whether a PDH was achieved) and diagnostic case model (represented by an asterisk). Y-axis has been reversed. Lower values (more negative) of negative log likelihood indicate an improvement in fit.

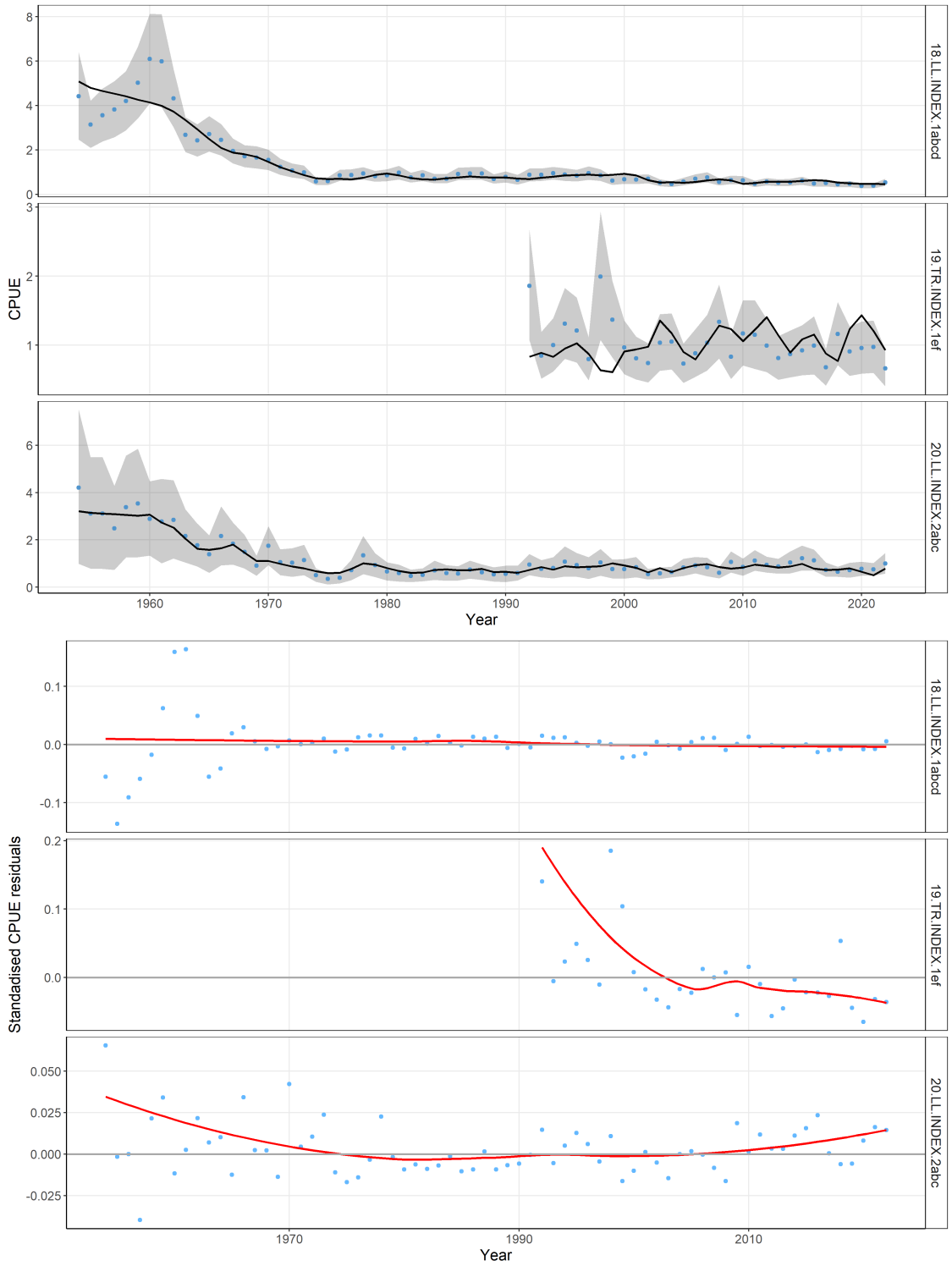


Figure 19: Model fits (top, black line) to observed standardised CPUE (blue dots) with 95% confidence intervals (grey shading) for the three index fisheries, and standardised residuals (bottom) of model fits to observed CPUE indices. The red line represents a lowess smoother fit to the residuals.

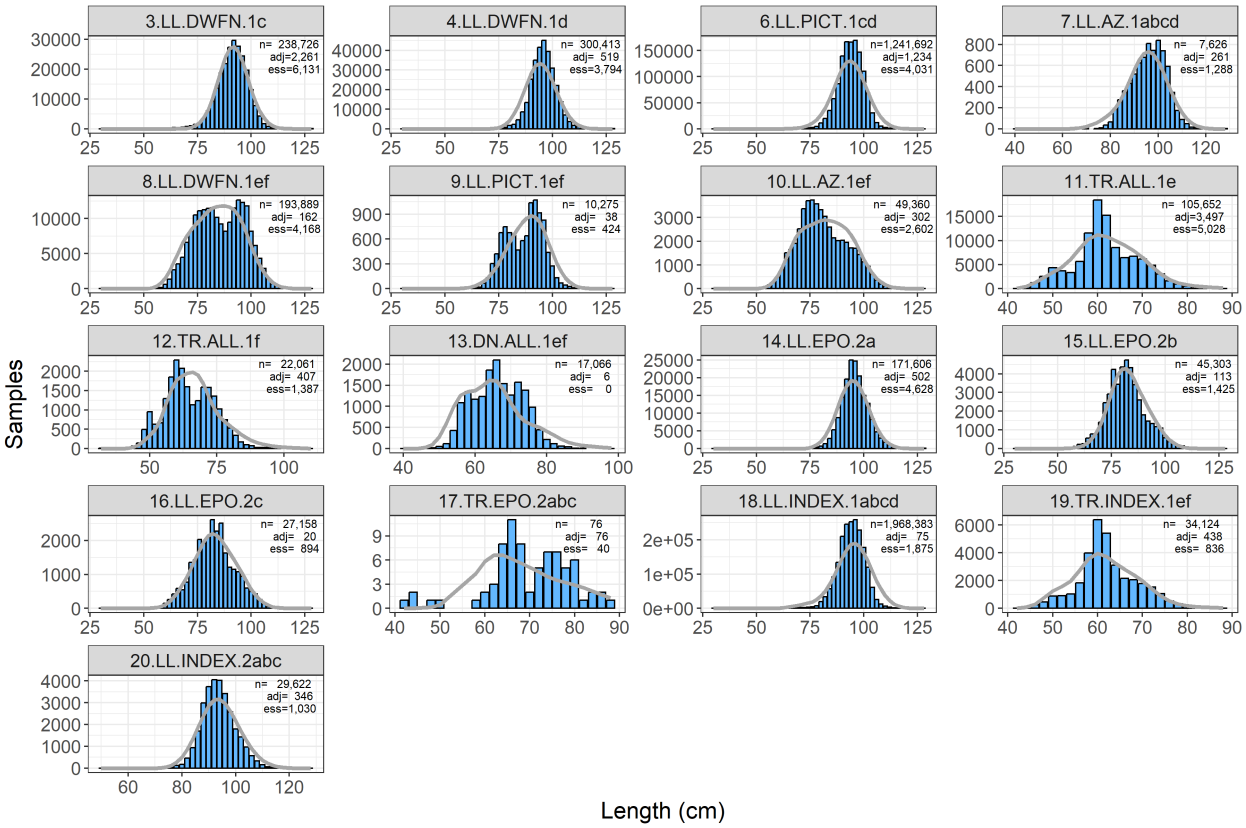


Figure 20: Aggregated (over time) observed (blue histograms) and predicted (grey lines) catch-at-length for longline, troll, and driftnet fisheries with sample sizes for the diagnostic case model. Includes the estimated effective sample size (ess) as derived from the robust normal likelihood, the adjusted input sample size (adj) as scaled by the Francis weighting method, and the observed sample size (n).

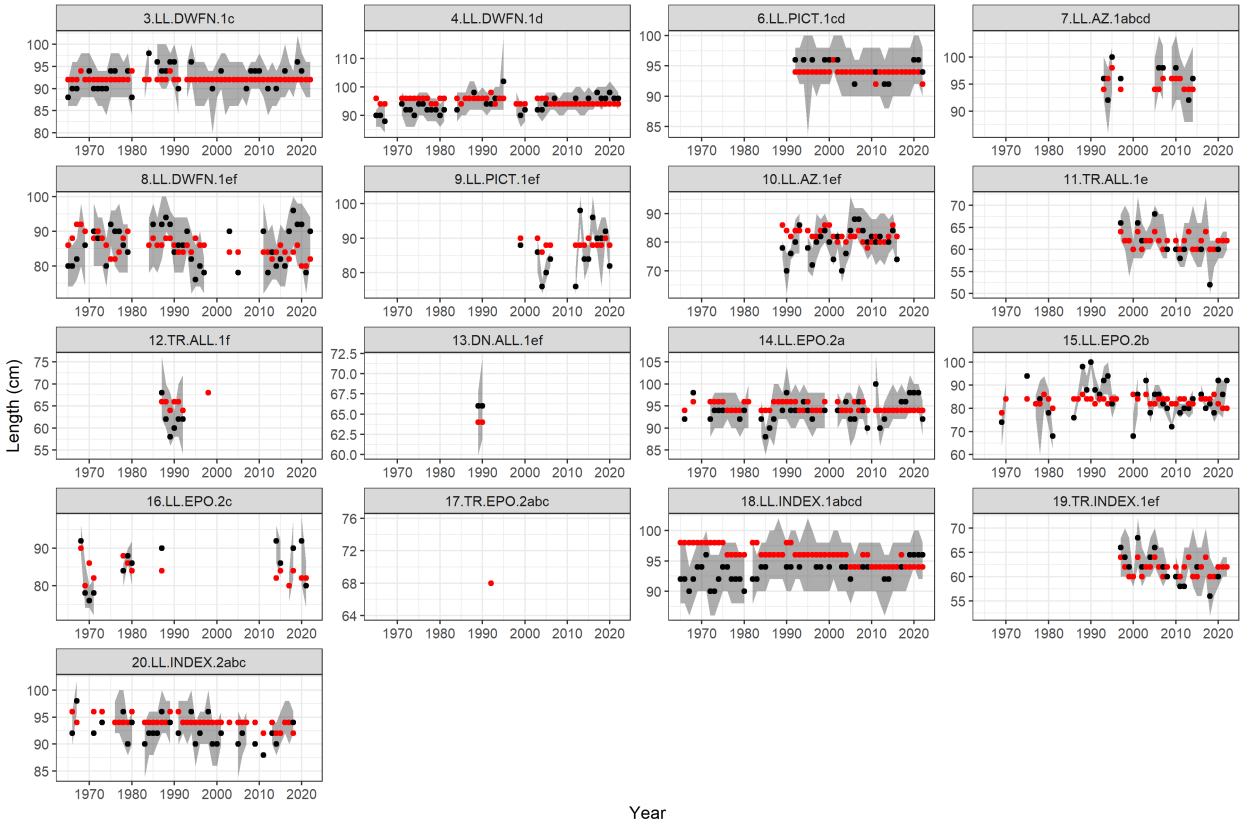


Figure 21: Comparison of the observed (black points) and predicted (red points) median length by fishery in the diagnostic case model. The intervals (grey shading) represent the observed values encompassed by the 25% and 75% quantiles. Sampling data are aggregated by year.

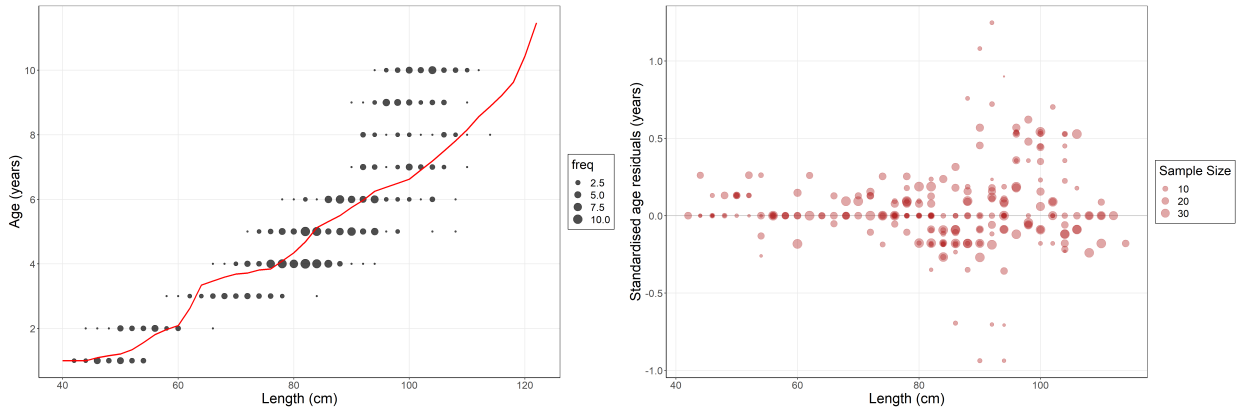


Figure 22: Comparison of the observed (black points sized based on number of observations) and predicted (red line) age-at-length data (left) with corresponding standardised residuals (right) for the diagnostic case model.

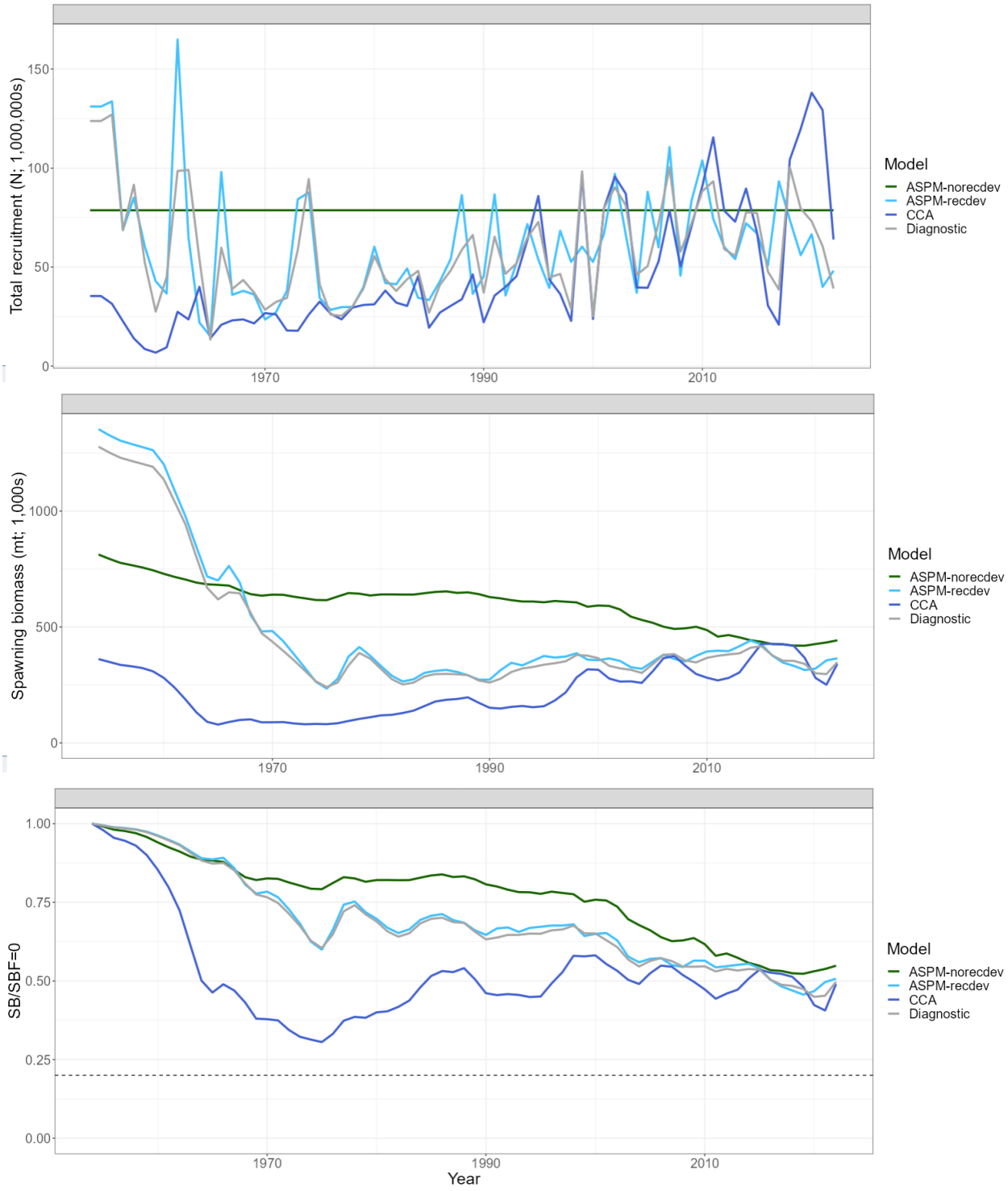


Figure 23: Recruitment (top), SB (middle), and $SB_t/SB_{F=0(t)}$ (bottom) for the age-structured production model (ASPM) with and without recruitment deviations (-recdev, -norecdev), catch-curve analysis (CCA), and the diagnostic case model (Diagnostic).

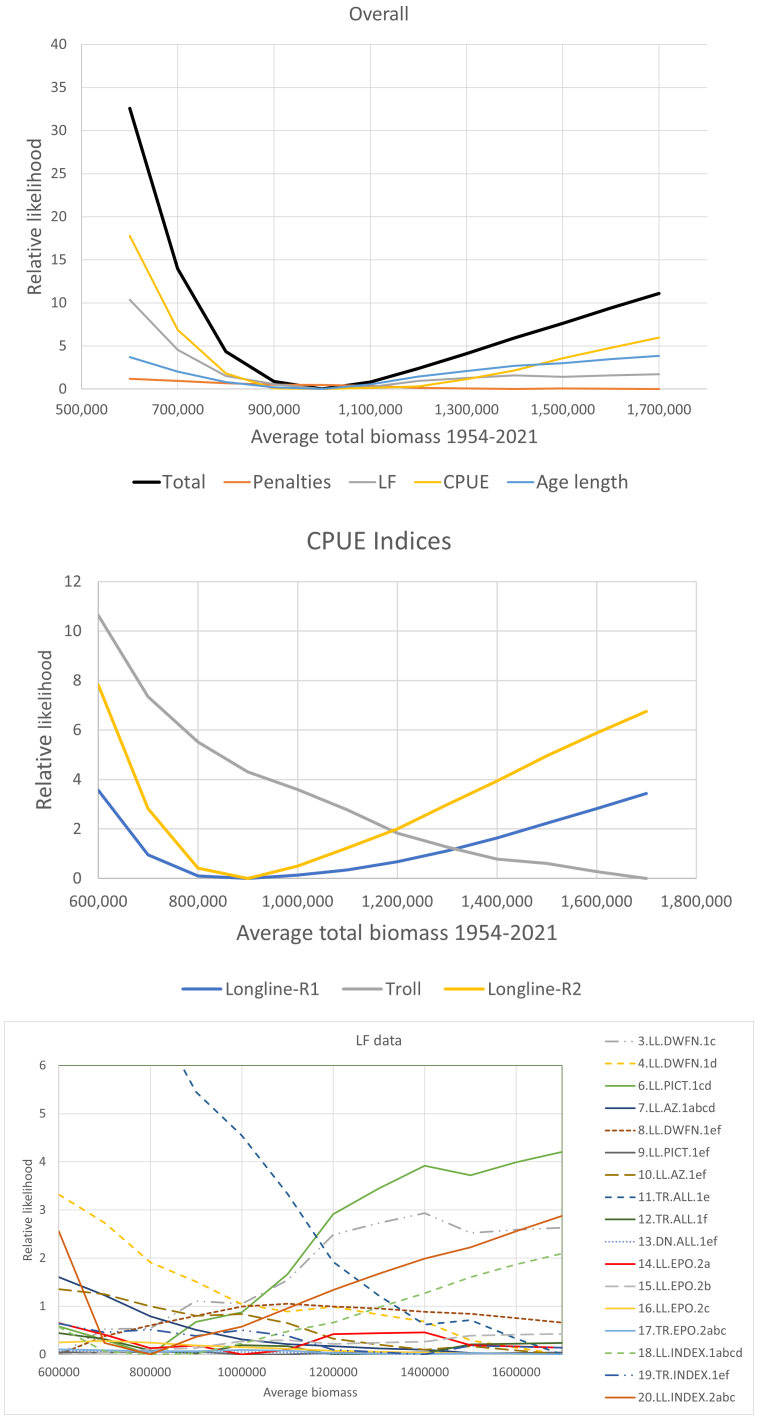


Figure 24: Relative likelihood (top) with total, penalties, length-frequency (LF), CPUE, and age-length data. Relative likelihood of CPUE indices (middle) for region 1 (R1) and region 2 (R2) longline indices and troll index. Relative likelihood of length-frequency (bottom) by fisheries. All relative likelihoods were performed on average total biomass (metric tons; 1954–2021).

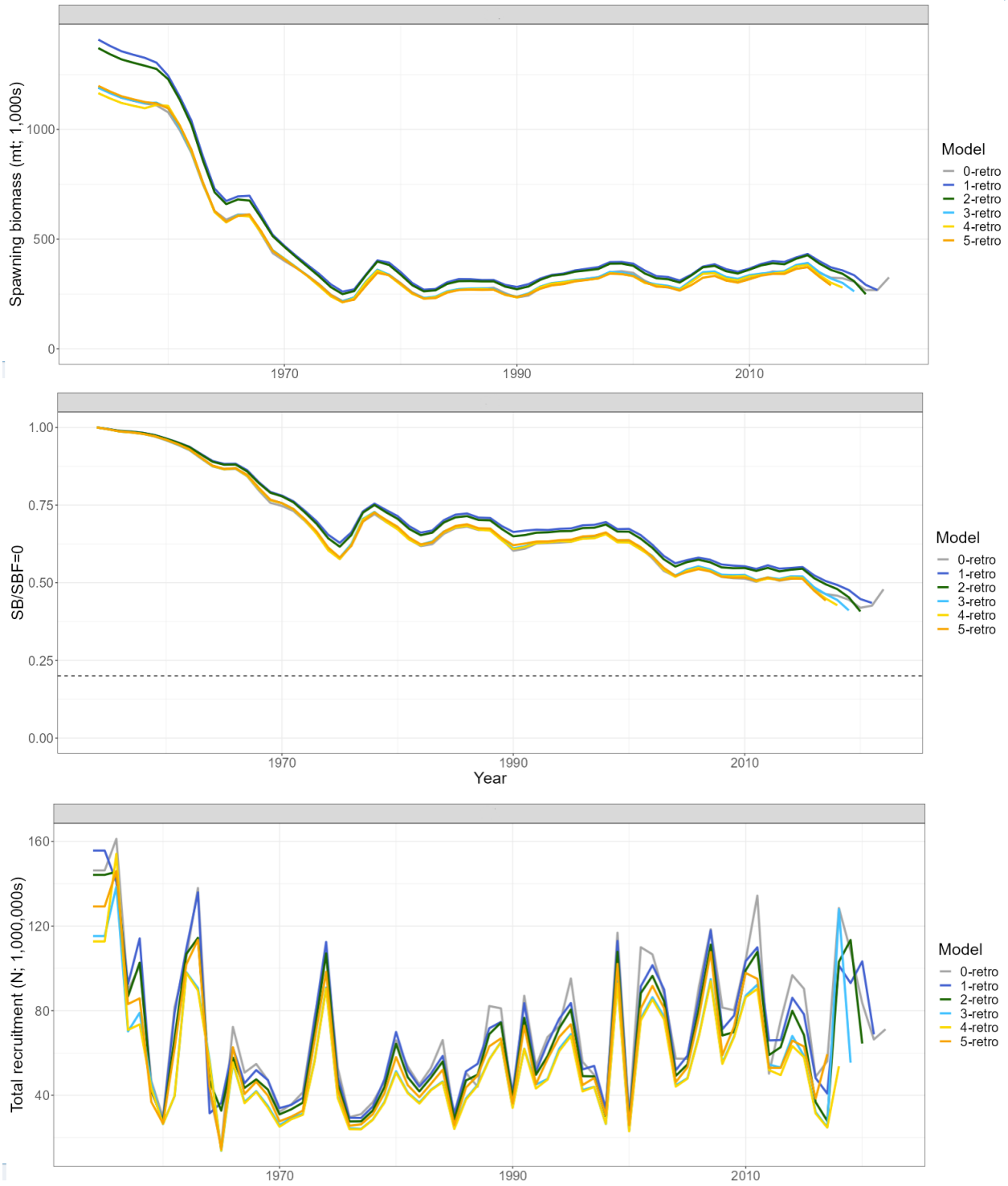


Figure 25: Retrospective analysis results for SB, $SB_t/SB_{F=0(t)}$, and recruitment for diagnostic case model.

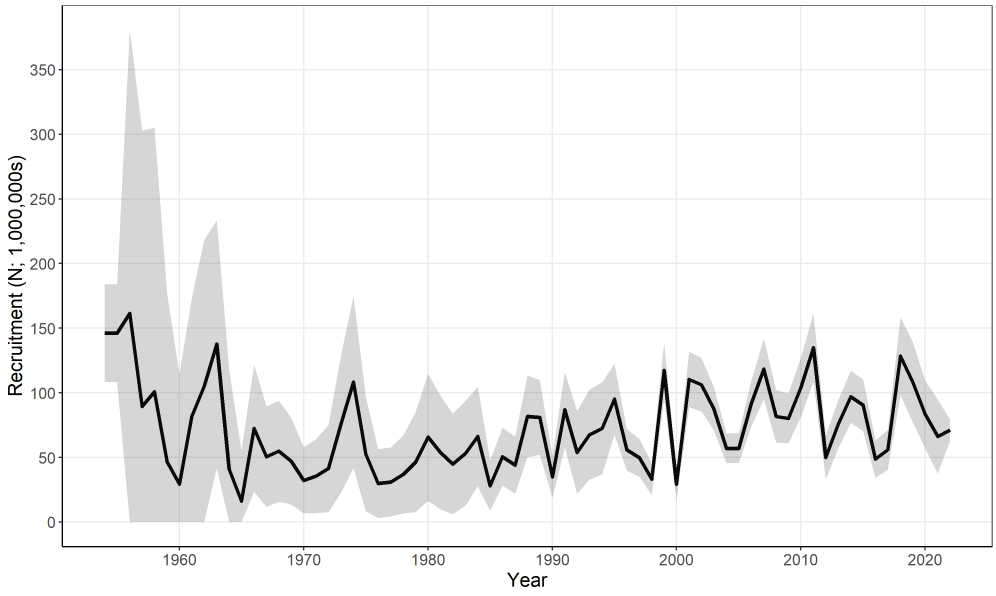


Figure 26: Estimated annual recruitment with 95% confidence intervals across model regions for the diagnostic case model.

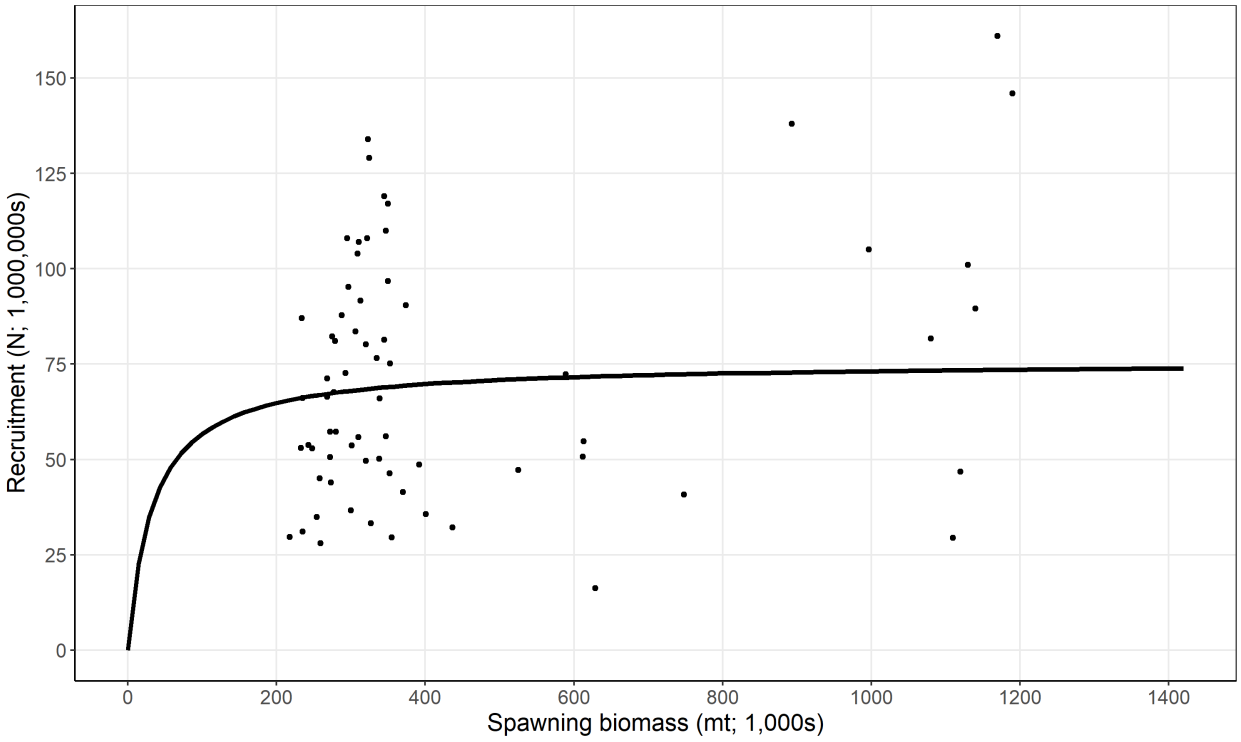


Figure 27: Estimated relationship between recruitment and SB based on annual values for the diagnostic case model.

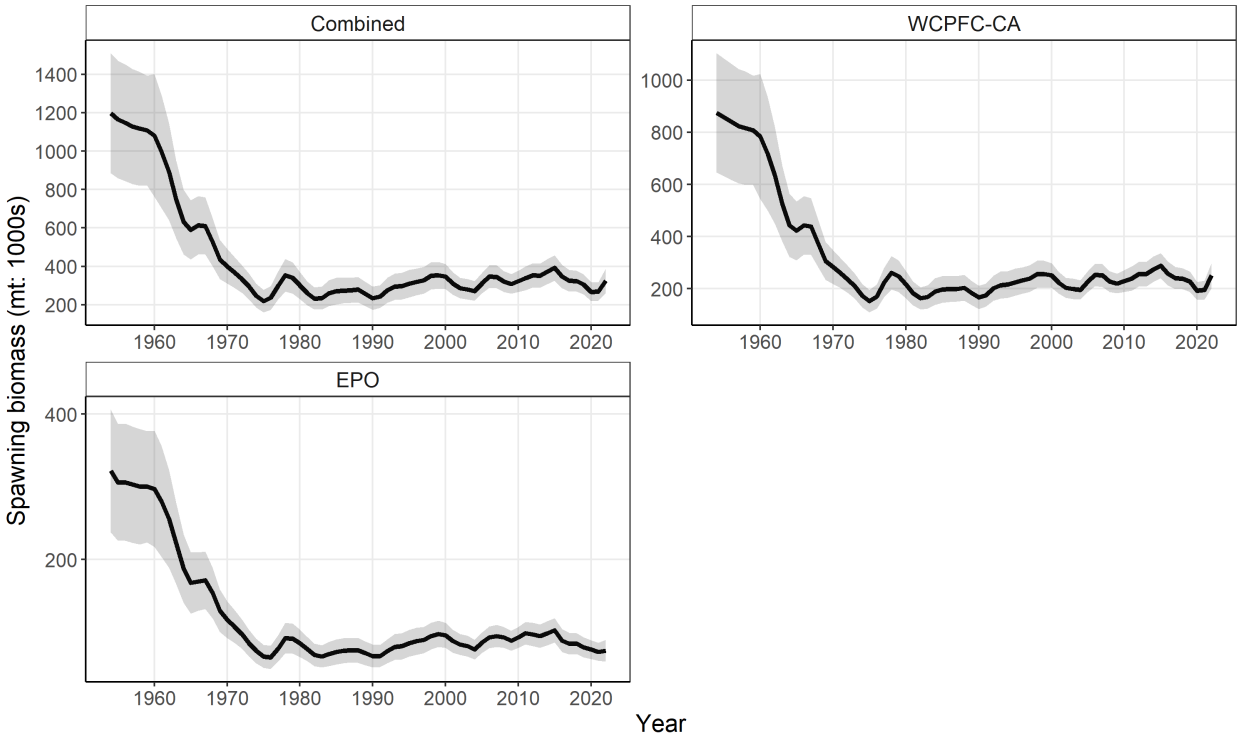


Figure 28: Estimated temporal SB with 95% confidence intervals by model region and the South Pacific as a whole, for the diagnostic case model.

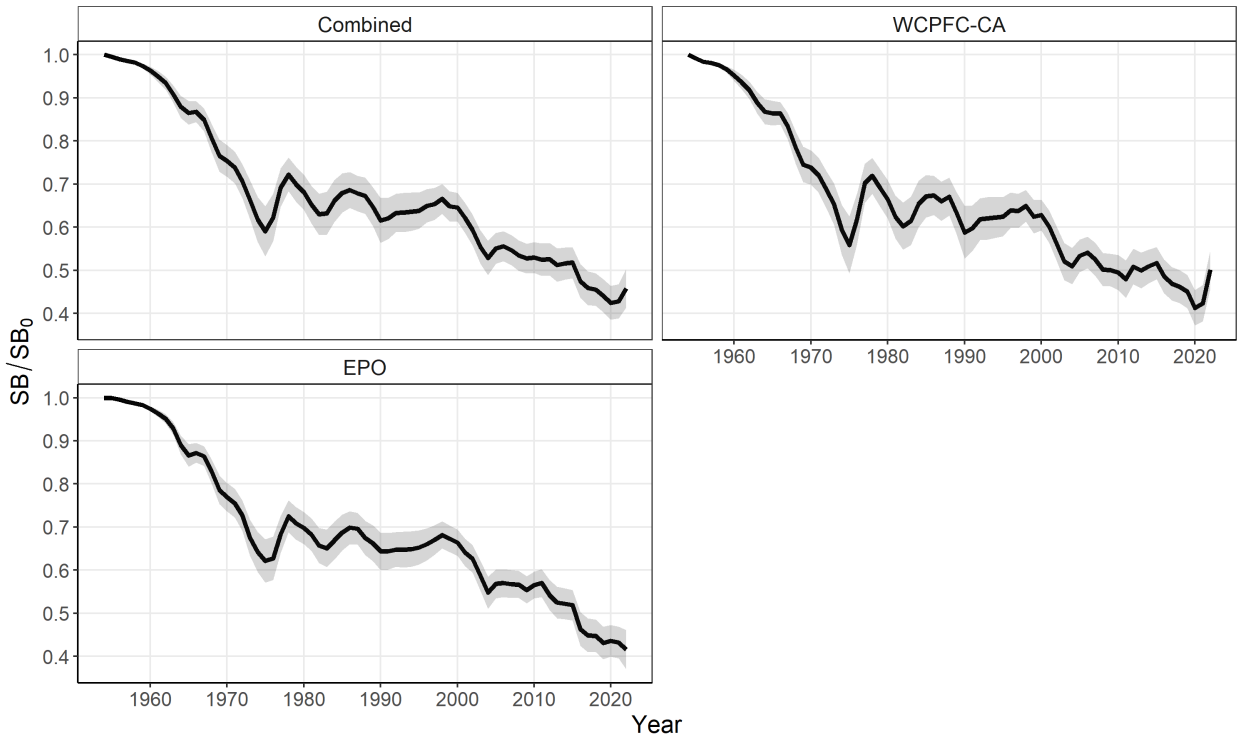


Figure 29: Estimated temporal $SB_t/SB_{F=0(t)}$ with 95% confidence intervals by model region and the South Pacific as a whole, for the diagnostic case model.

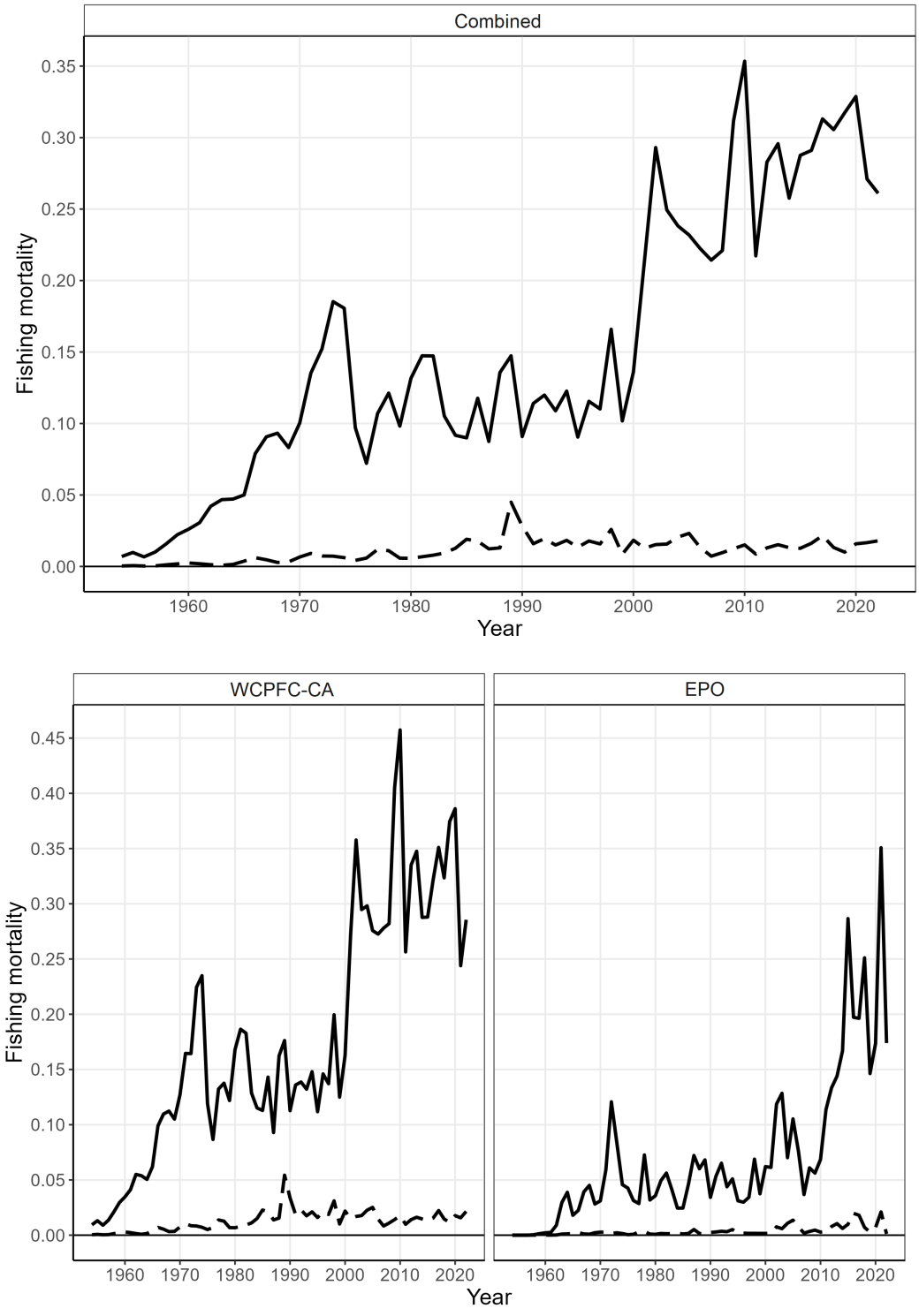


Figure 30: Estimated annual juvenile (dashed line) and adult (solid line) fishing mortality for the diagnostic case model.

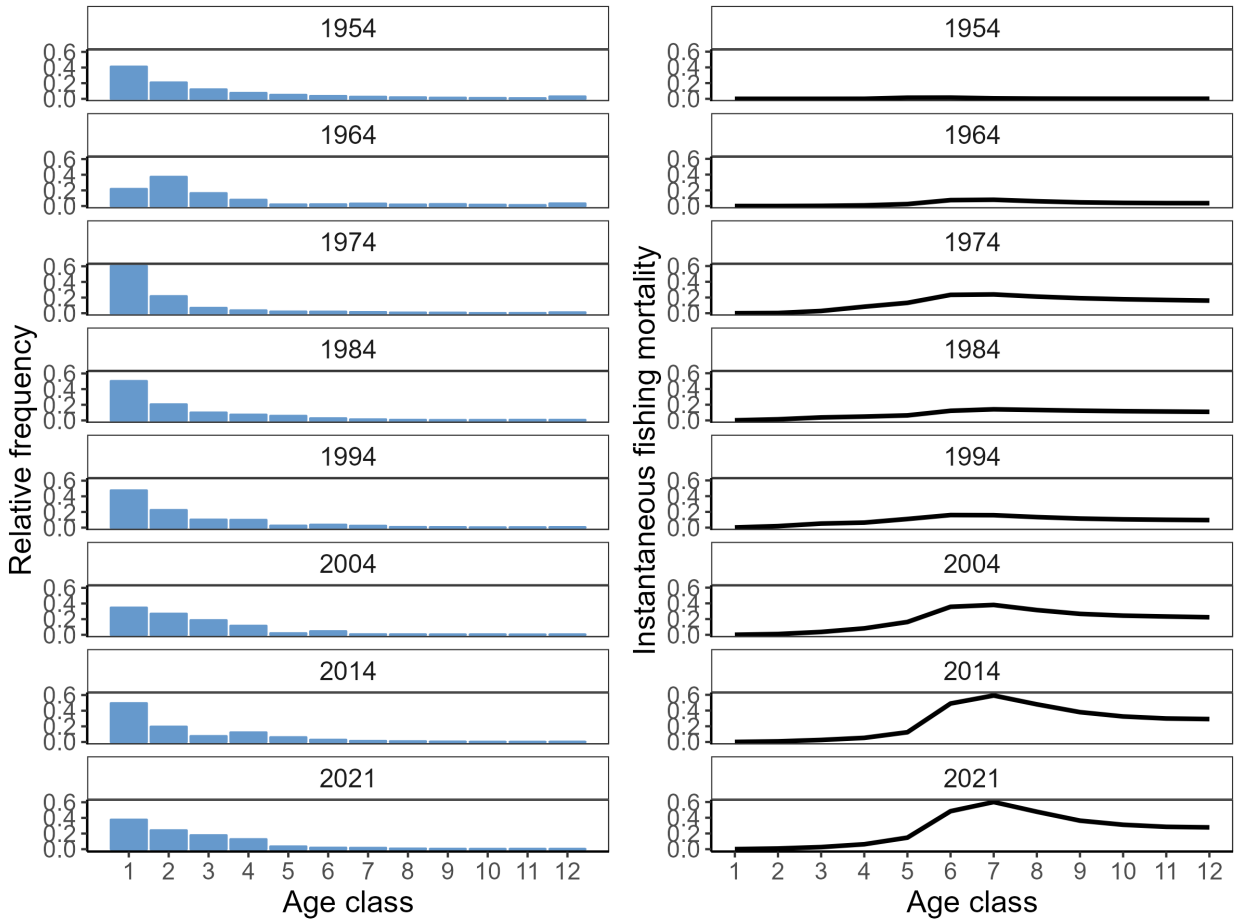


Figure 31: Estimated relative frequency by age class and decade (numbers-at-age; left) and estimated instantaneous fishing mortality by age class and decade (right) for the diagnostic case model.

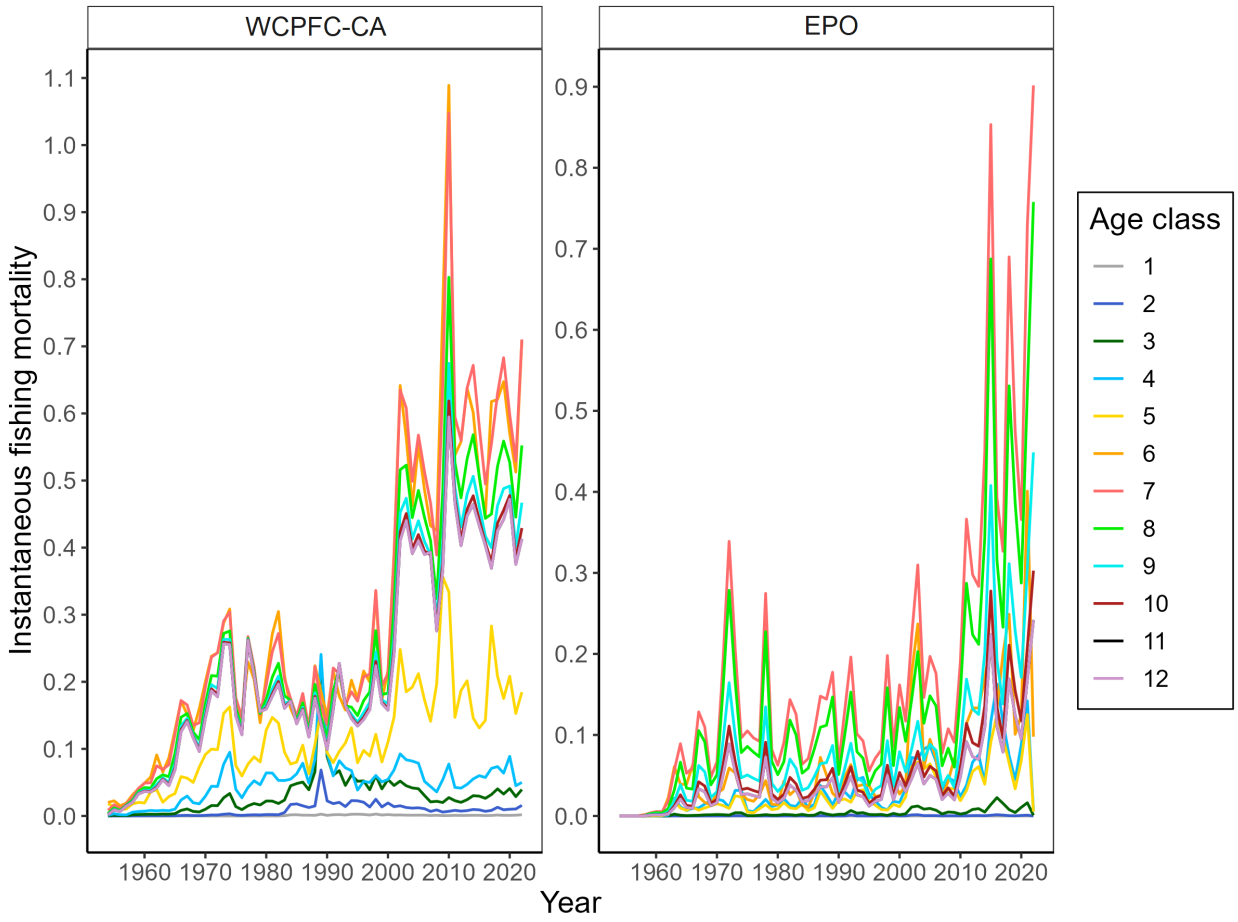


Figure 32: Estimated annual instantaneous fishing mortality by age class and area for the diagnostic case model.

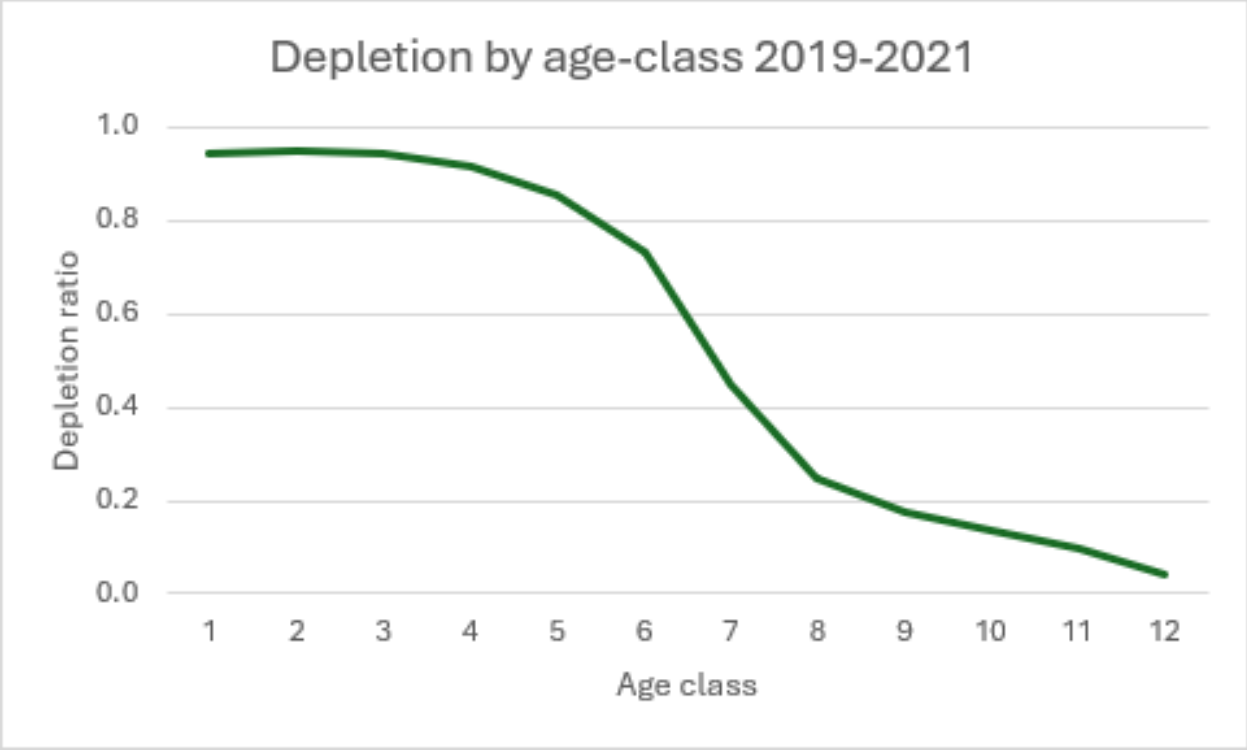


Figure 33: Estimated $SB_t/SB_{F=0(t)}$ by age class for the diagnostic case model.

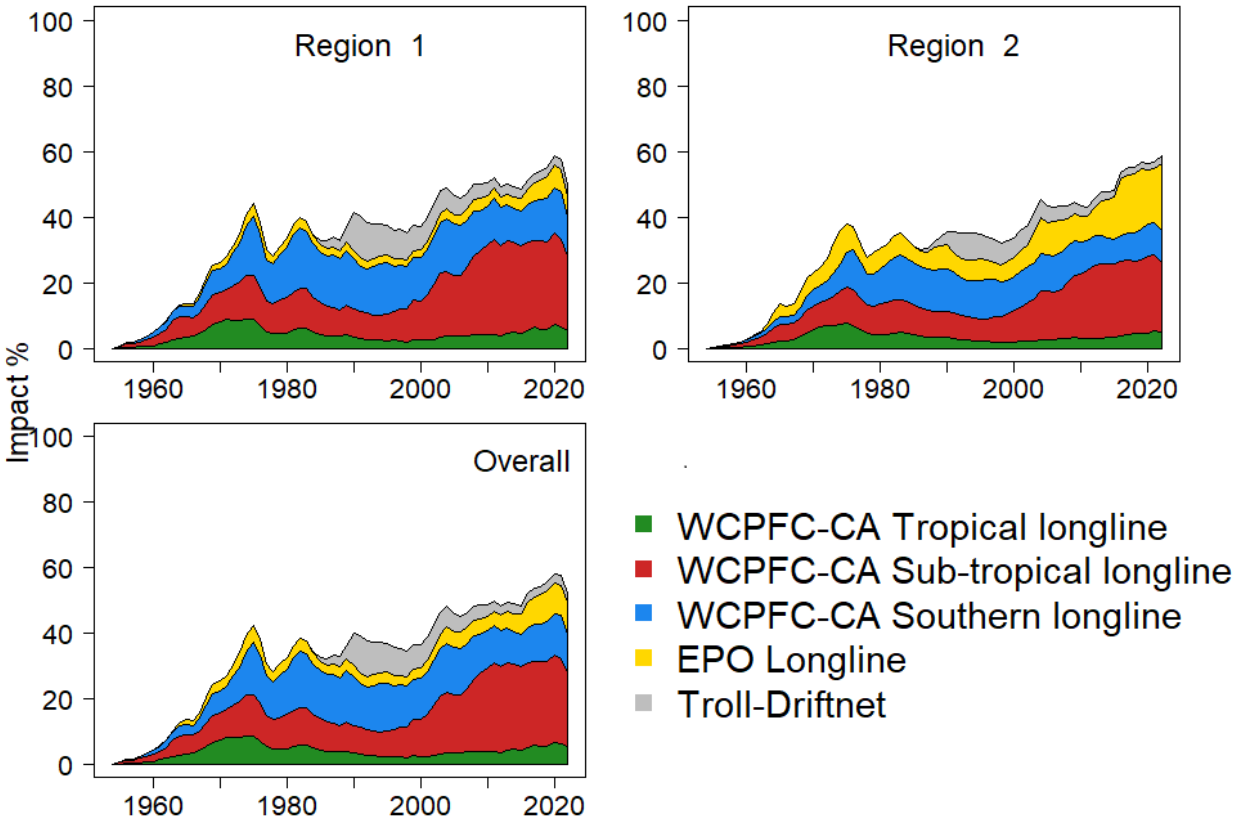


Figure 34: Estimates of fishery impact, or reduction in SB due to fishing (Fishery impact = 1 - $SB_t/SB_{F=0(t)}$) by region (region 1 – WCPFC-CA and region 2 – EPO) and combined regions attributed to various fishery groups for the diagnostic case model.

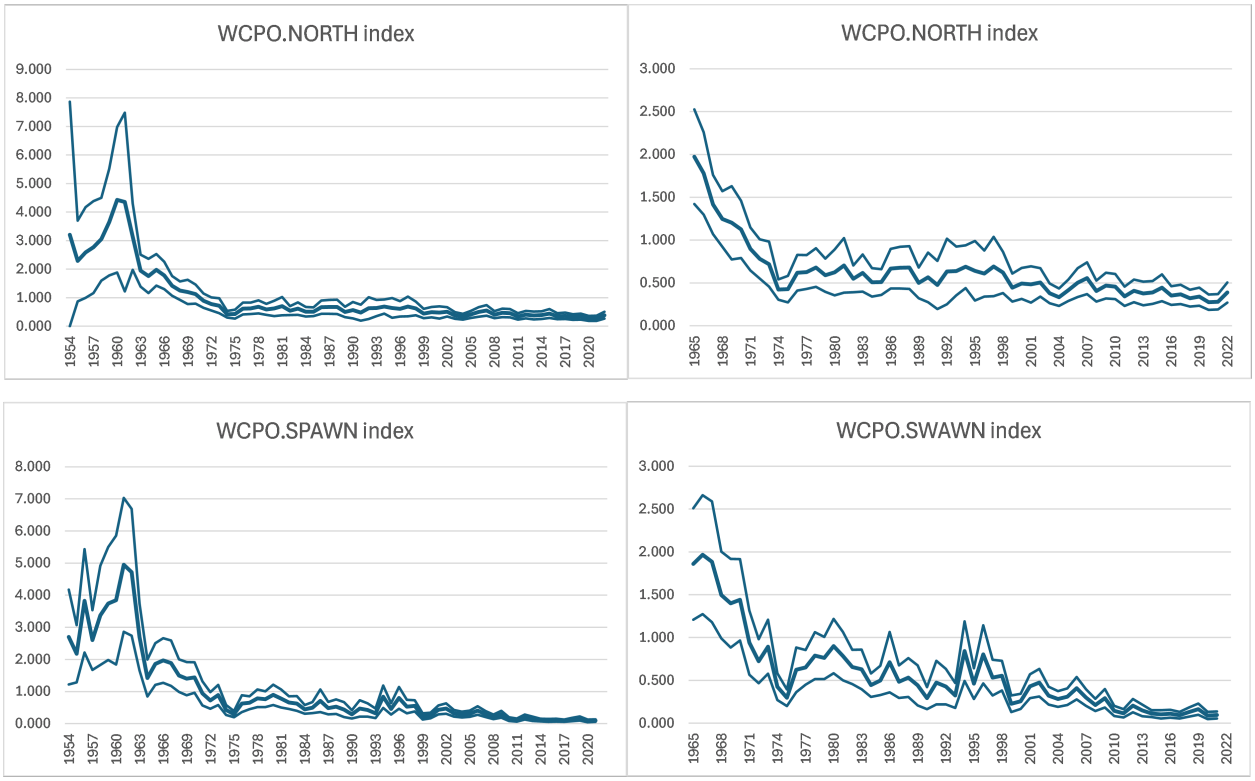


Figure 35: The WCPFC-CA.NORTH and WCPFC-CA.SPAWN indices of relative abundance, and their 95% confidence intervals, tested in the sensitivity analysis. The panels on the left show the full time series, while the panels on the right omit the initial ten years to better highlight recent trends.

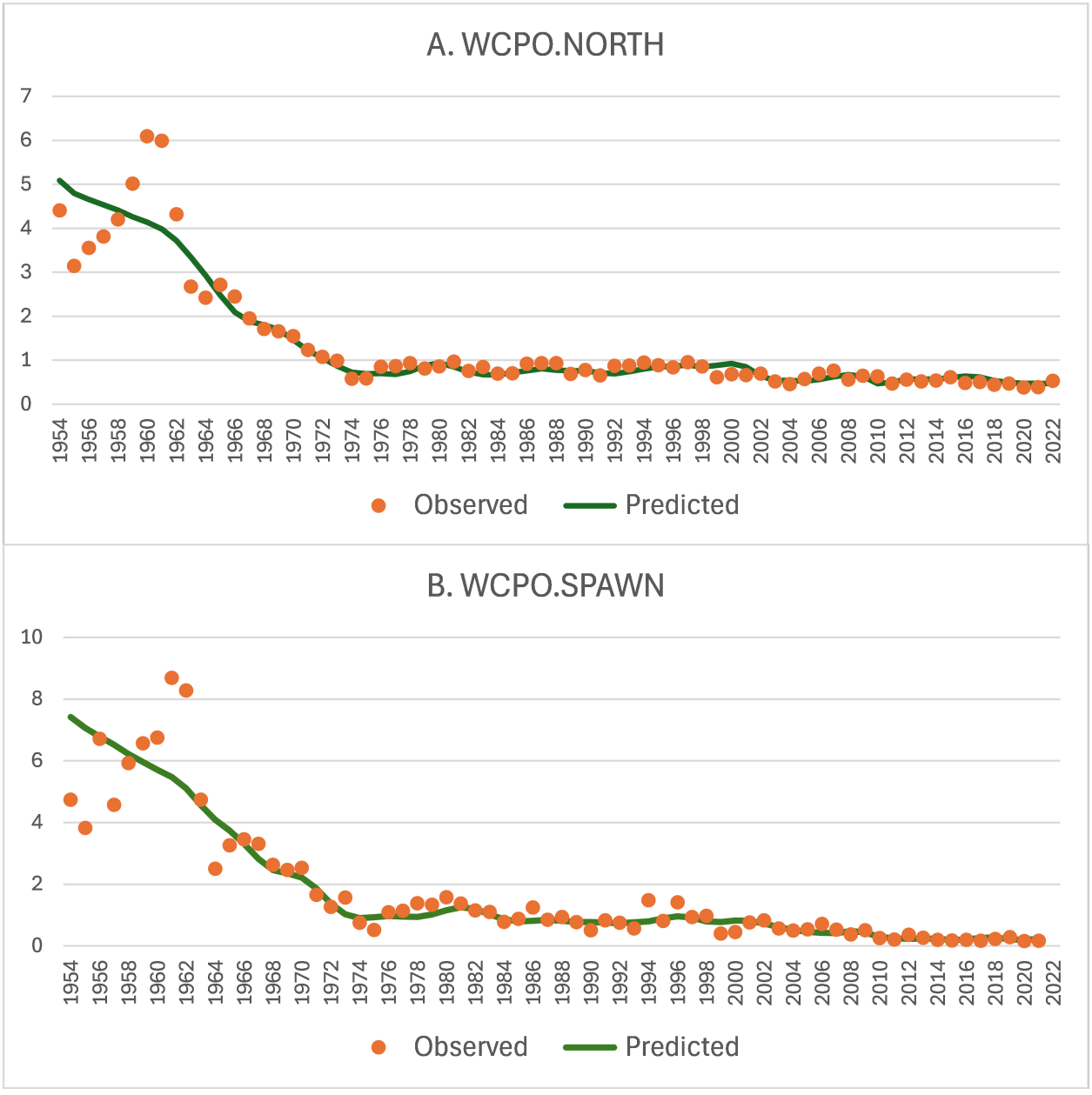


Figure 36: Fits to the WCPFC-CA longline indices.

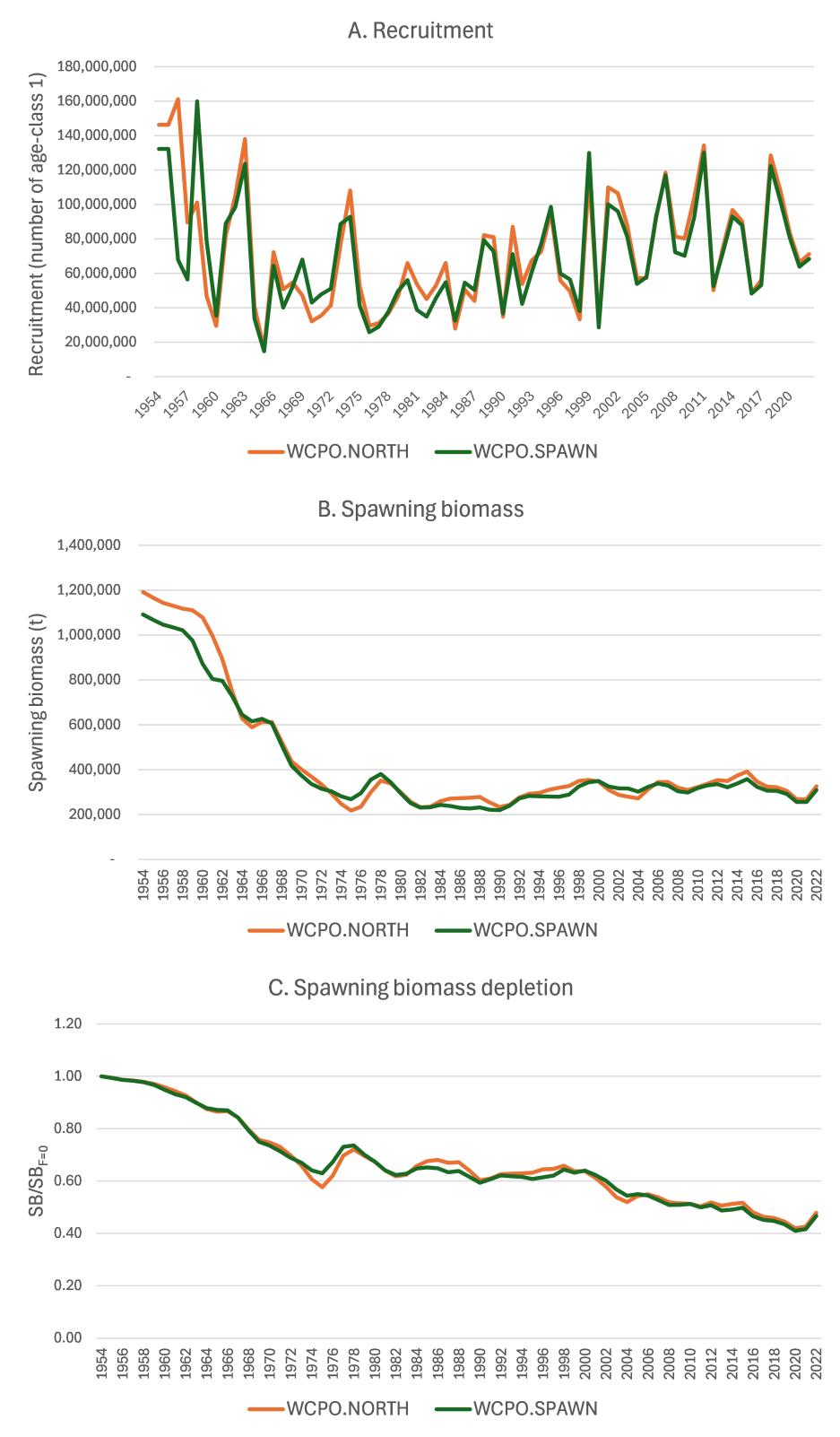


Figure 37: Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ from models using the WCPFC-CA.NORTH and WCPFC-CA.SPAWN indices of abundance.

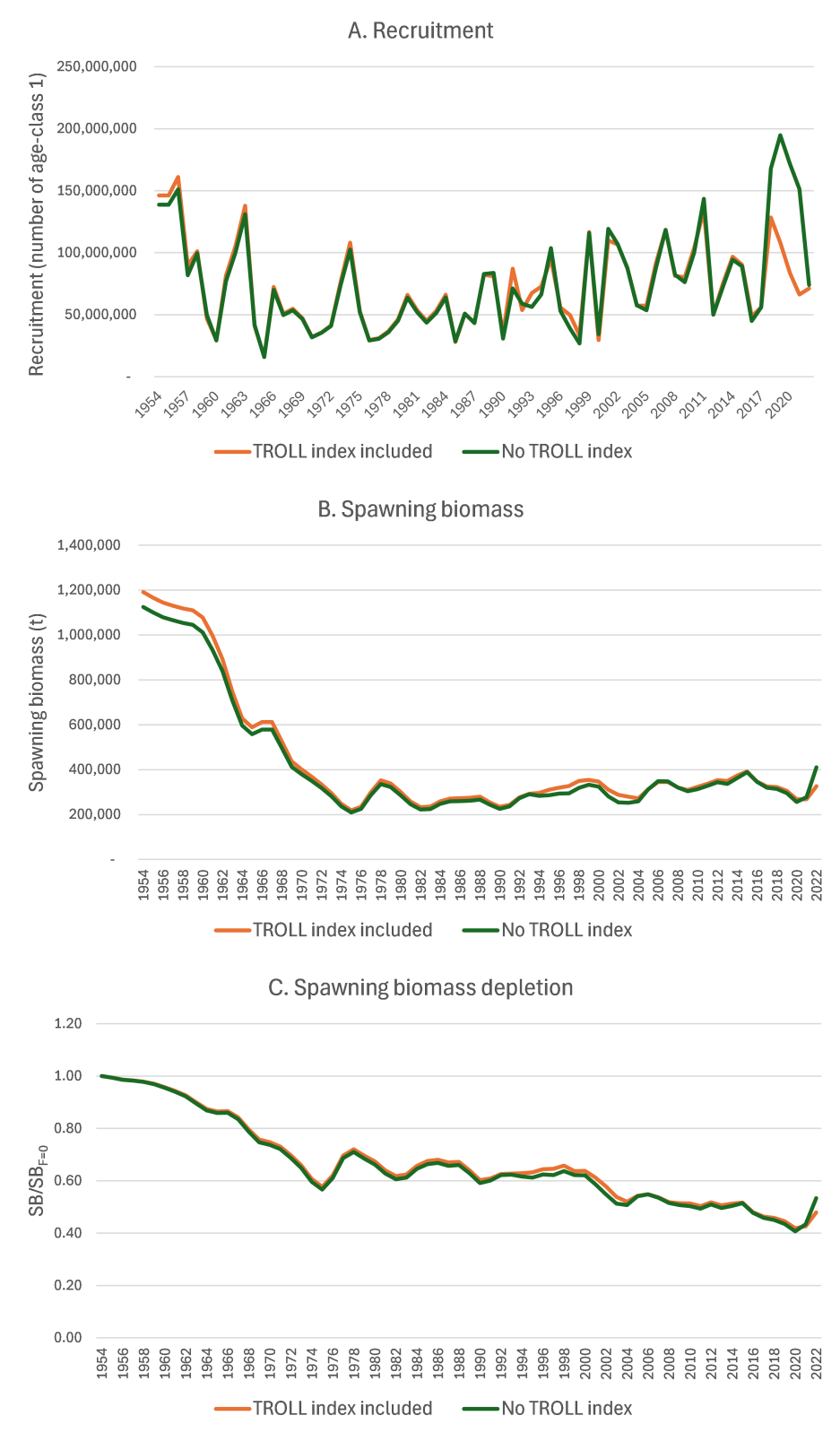


Figure 38: Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ from models using the WCPFC-CA.NORTH index of abundance with and without the inclusion of the NZ troll index of abundance.

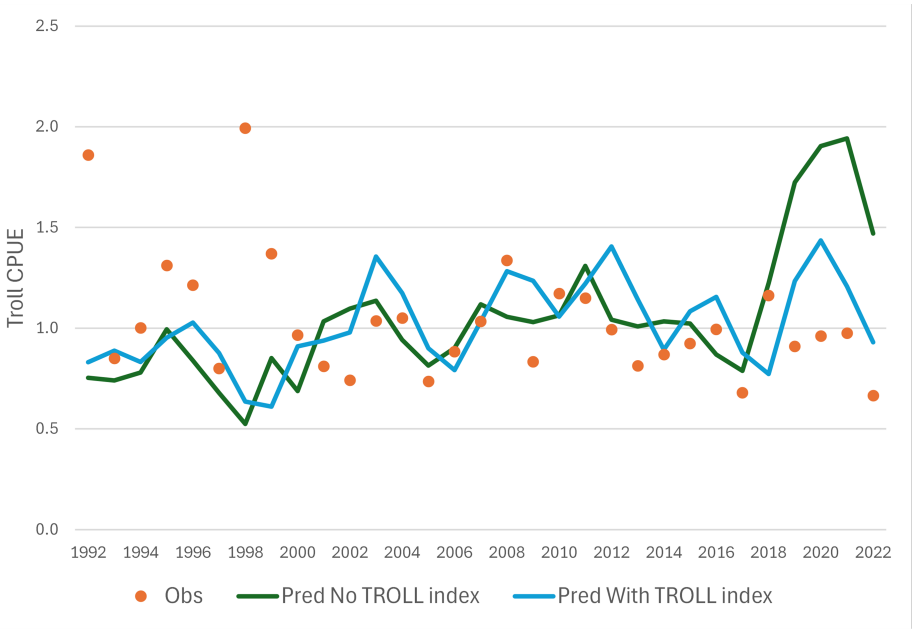


Figure 39: NZ troll fishery CPUE observations (red circles) and model predictions of the index. The blue line is the model predictions for the diagnostic model in which the troll index is fitted (average CV of 0.2). The green line is from a model in which the troll index has been down-weighted to a trivially small level (CV of 10,000).

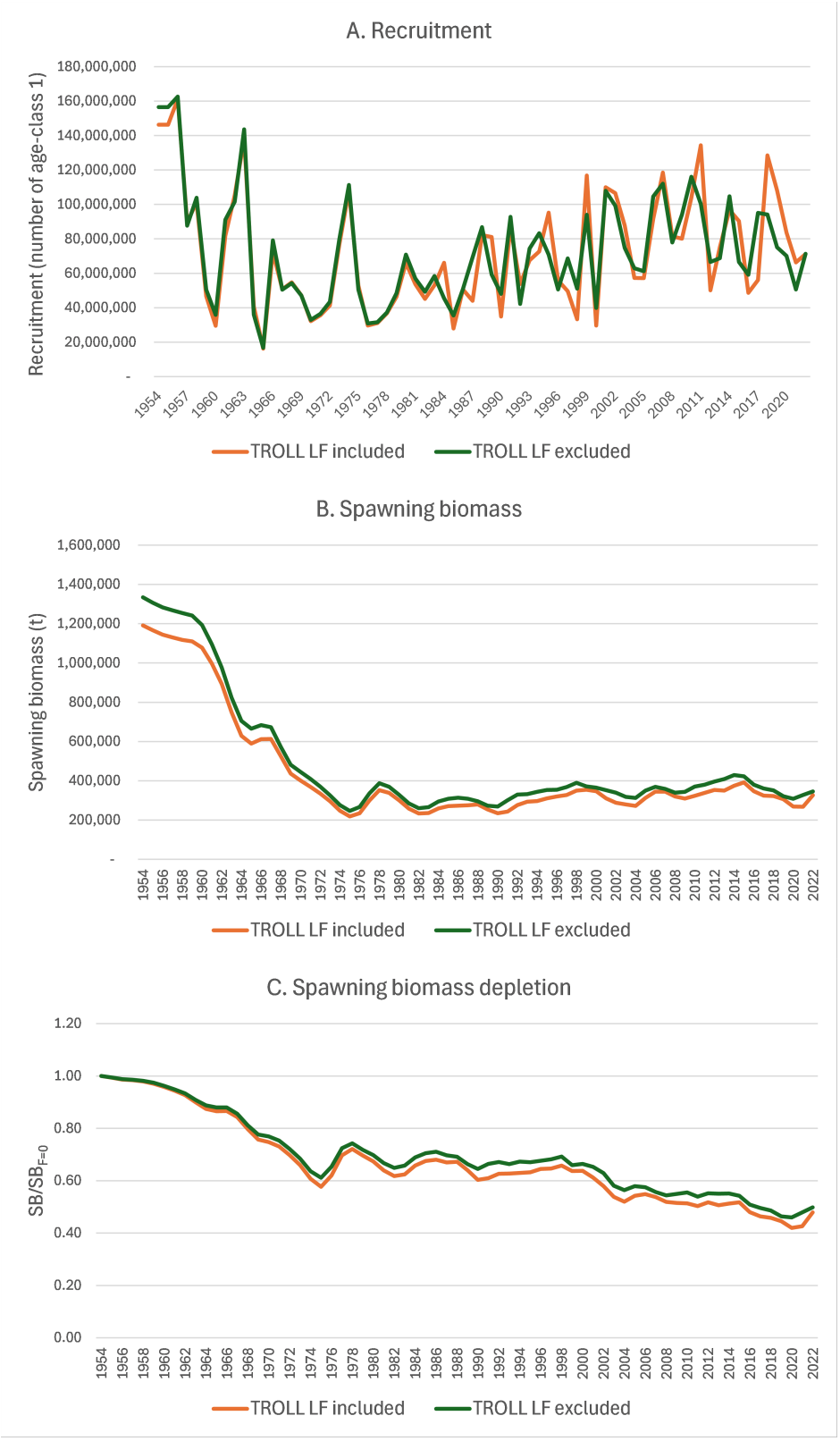


Figure 40: Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ from models using the WCPFC-CA.NORTH index of abundance with and without troll length frequency data.

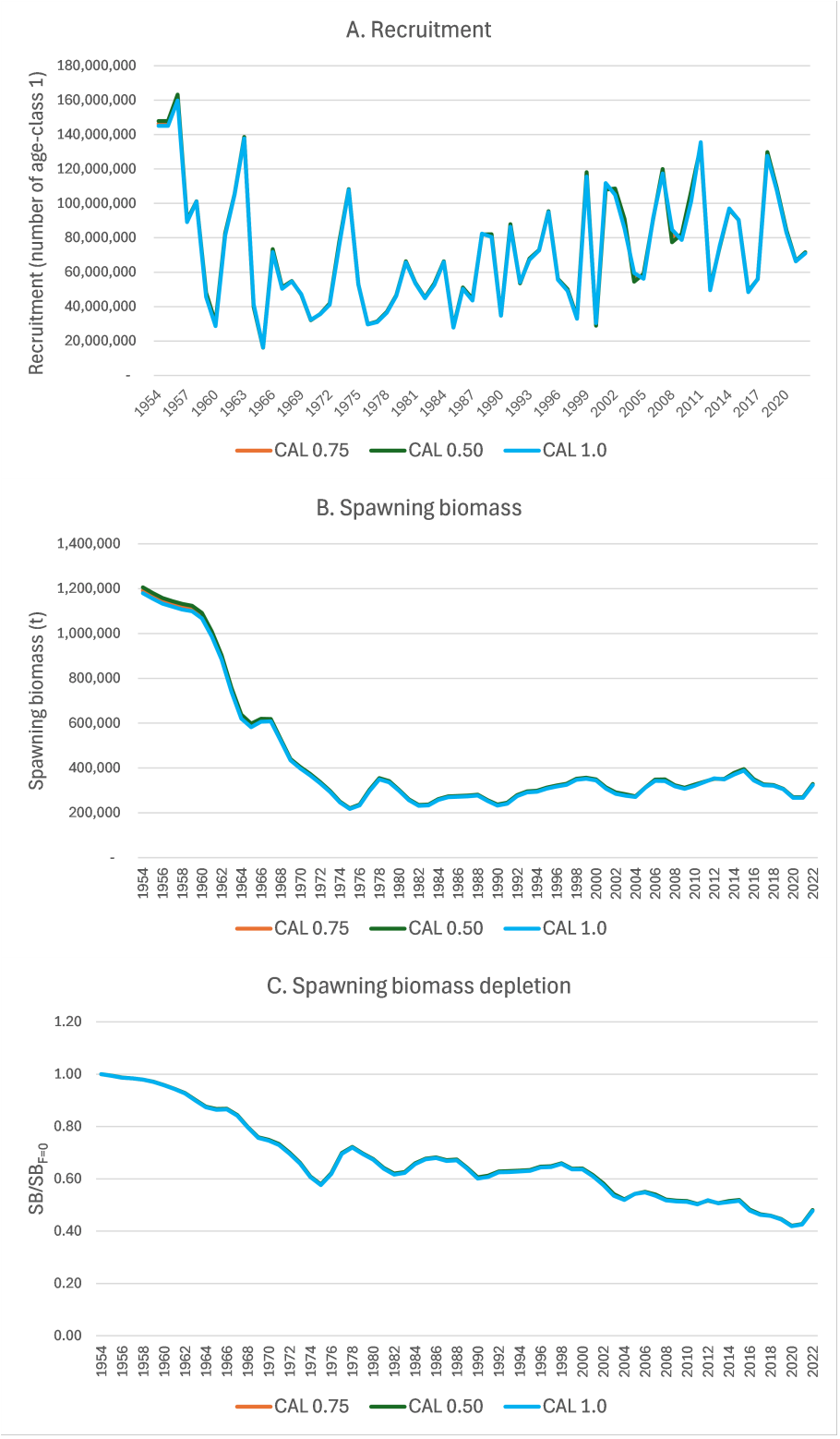


Figure 41: Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ from models with conditional age-at-length (CAAL) data set to an effective sample size (ESS) of 0.75 (diagnostic case), 0.5 and 1.0 times the observed sample size.

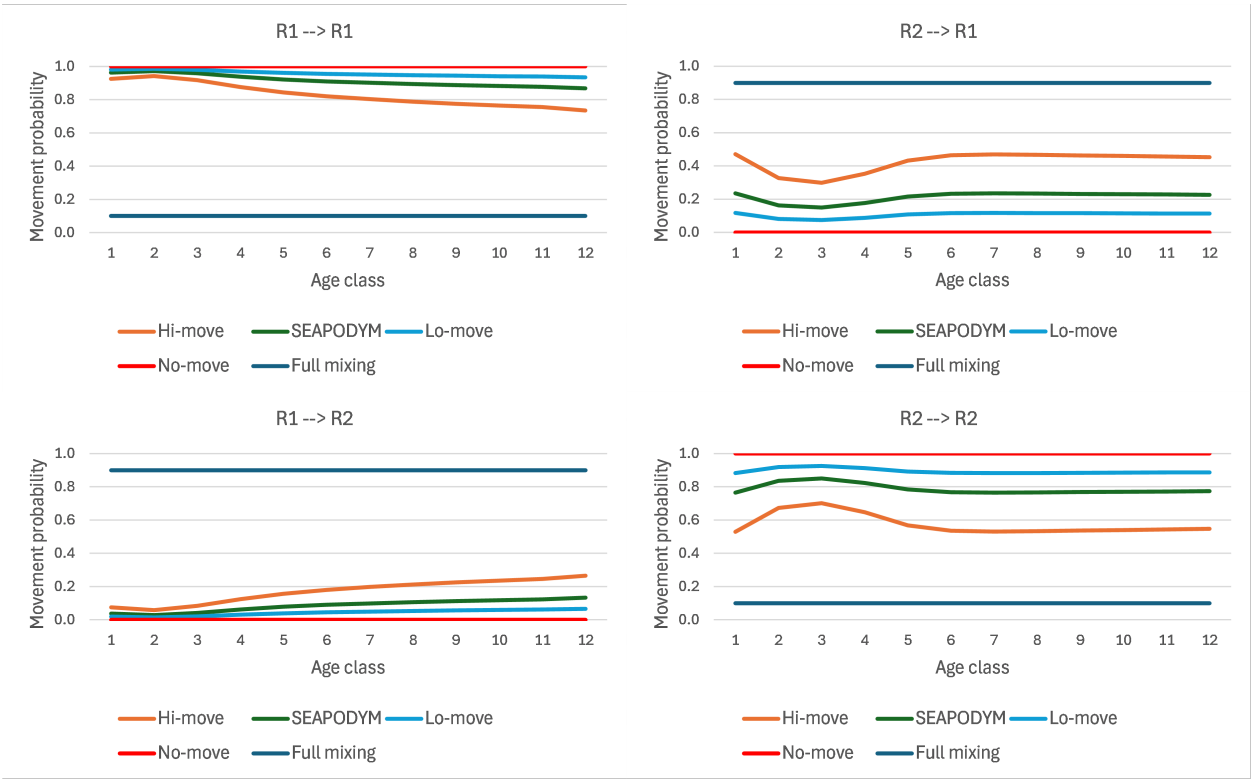


Figure 42: Movement probabilities for the first quarter applied in the sensitivity analysis of movement. Hi-move is 2 x SEAPODYM movement, Lo-move is 0.5 x SEAPODYM movement, Full-move is movement probabilities of 0.9 per period in both directions.

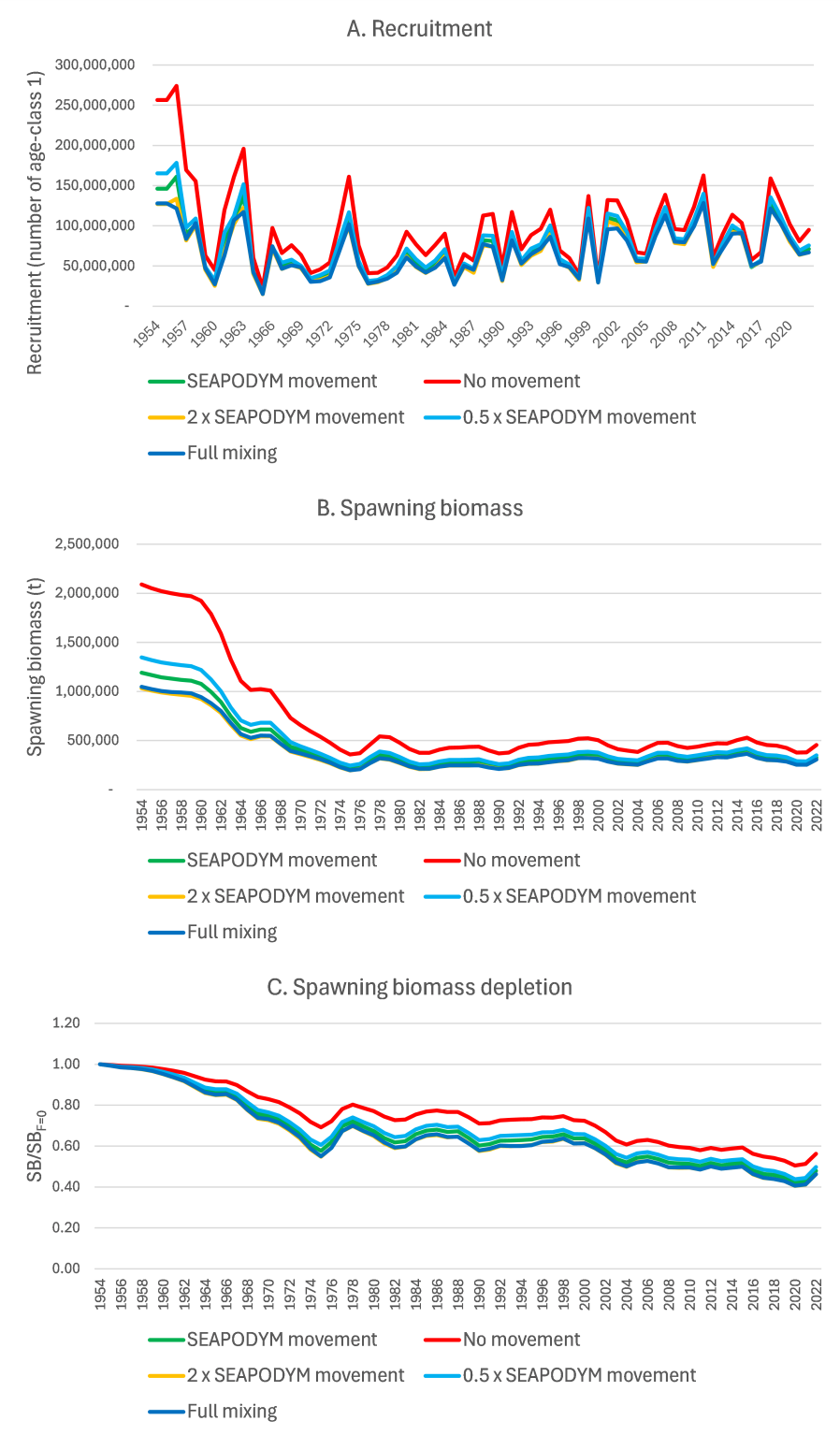


Figure 43: Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ from models with different movement scenarios. Full mixing has movement probabilities in both directions of 0.9 per period.

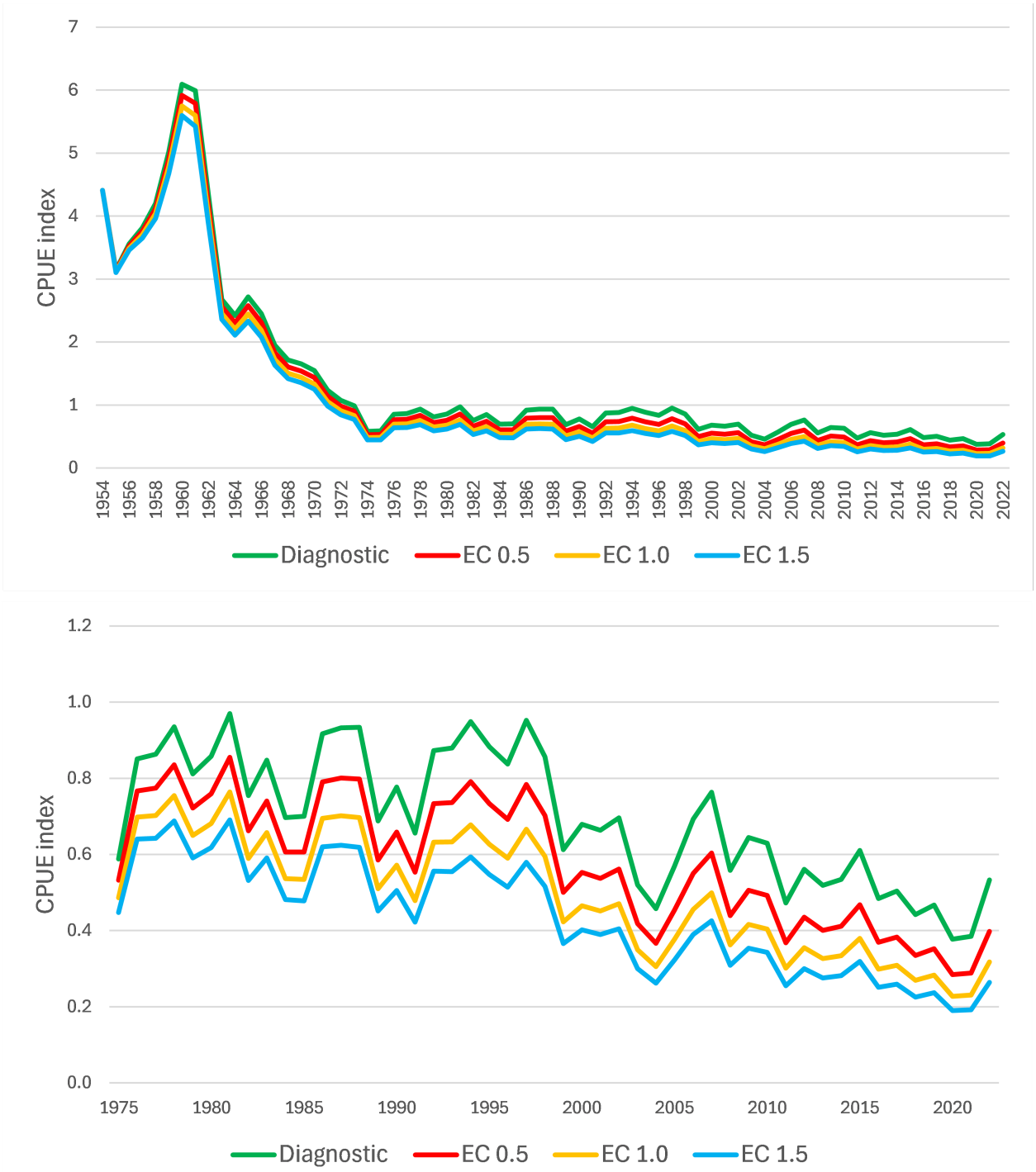


Figure 44: CPUE indices for the WCPFC-CA.NORTH index under a range of effort creep scenarios. The upper panel shows the full time series while the lower panel shows the series from 1975 to highlight more recent trends.

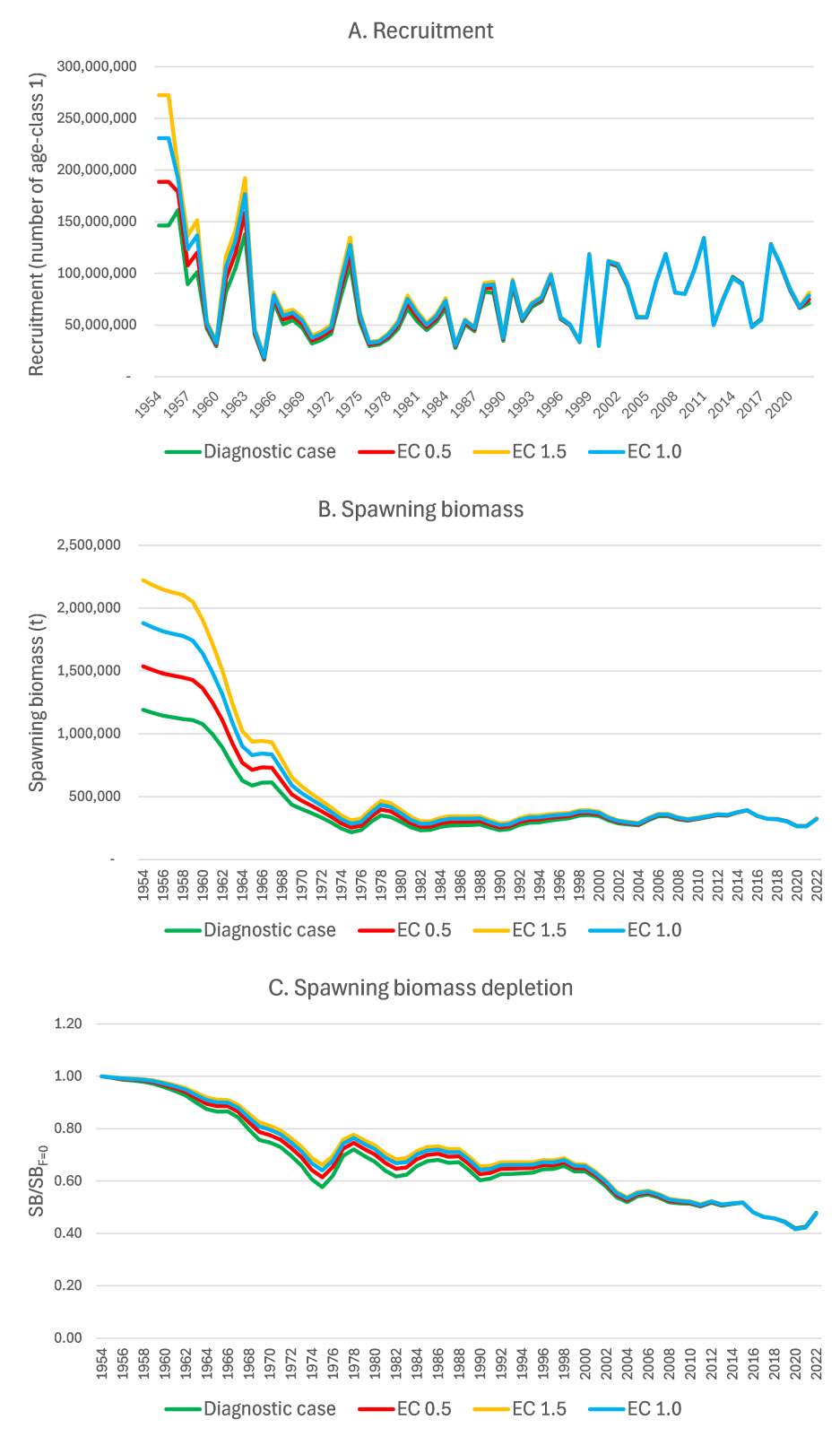


Figure 45: Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ from models with different effort creep scenarios.

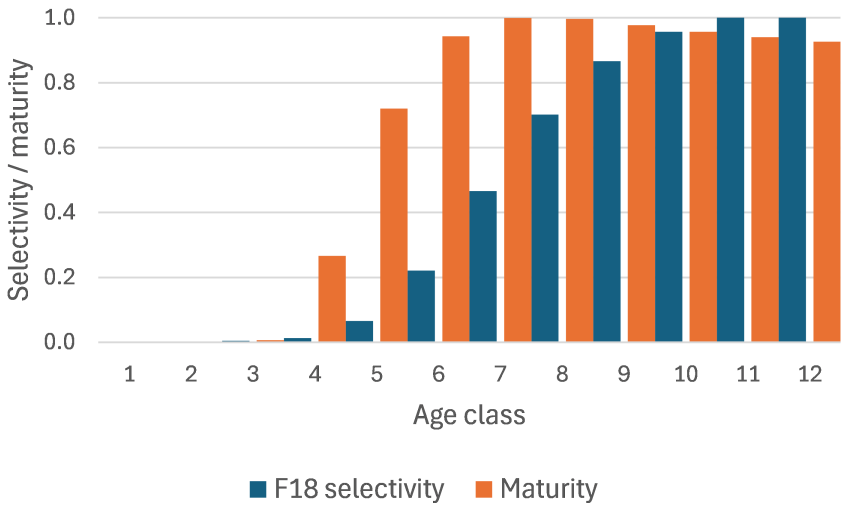


Figure 46: Estimated age-specific selectivity for the longline WCPFC-CA.NORTH index (F18) and maturity.

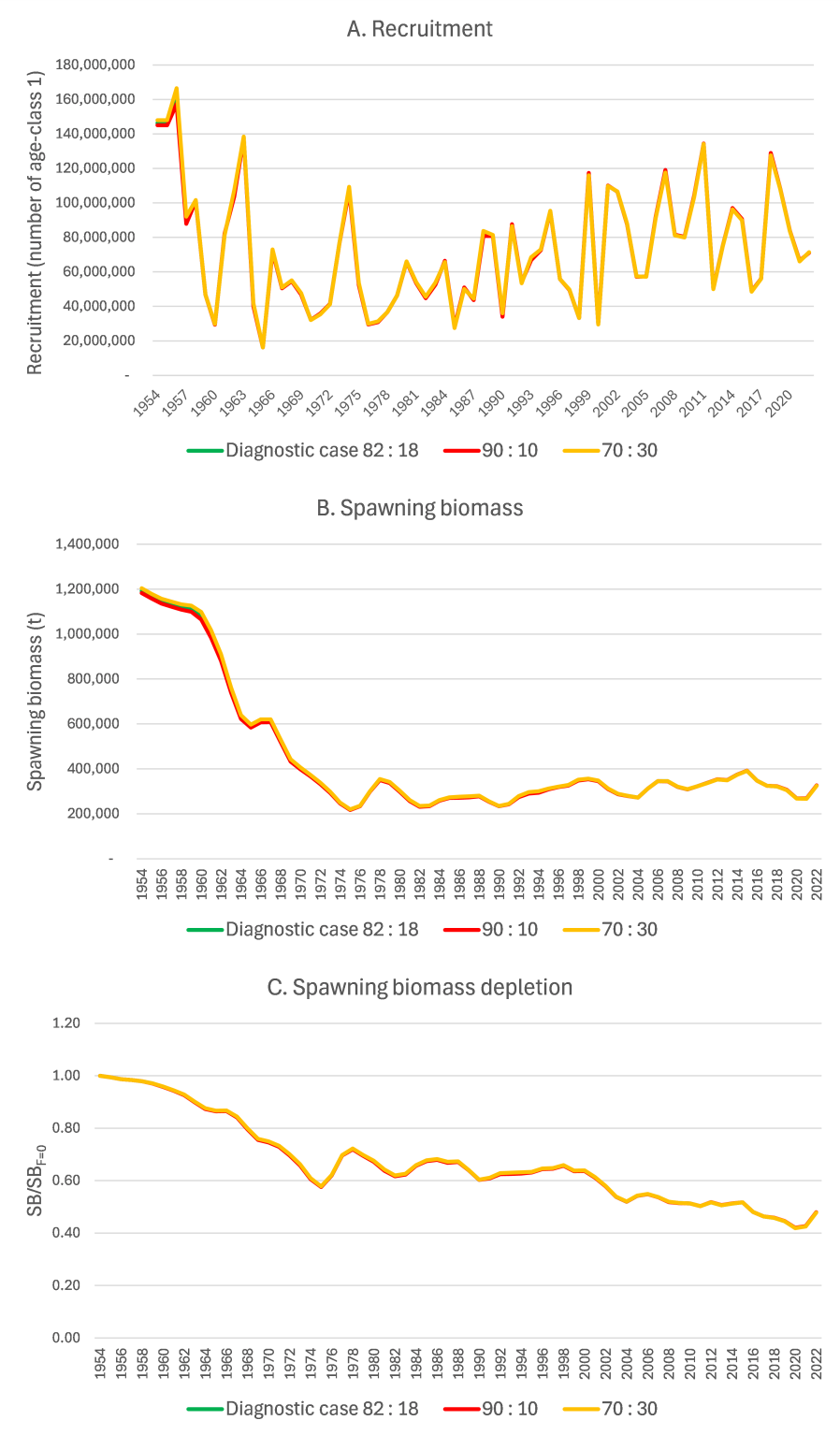


Figure 47: Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ from models with different assumptions regarding the distribution of recruitment between regions 1 (WCPFC-CA) and 2 (EPO).

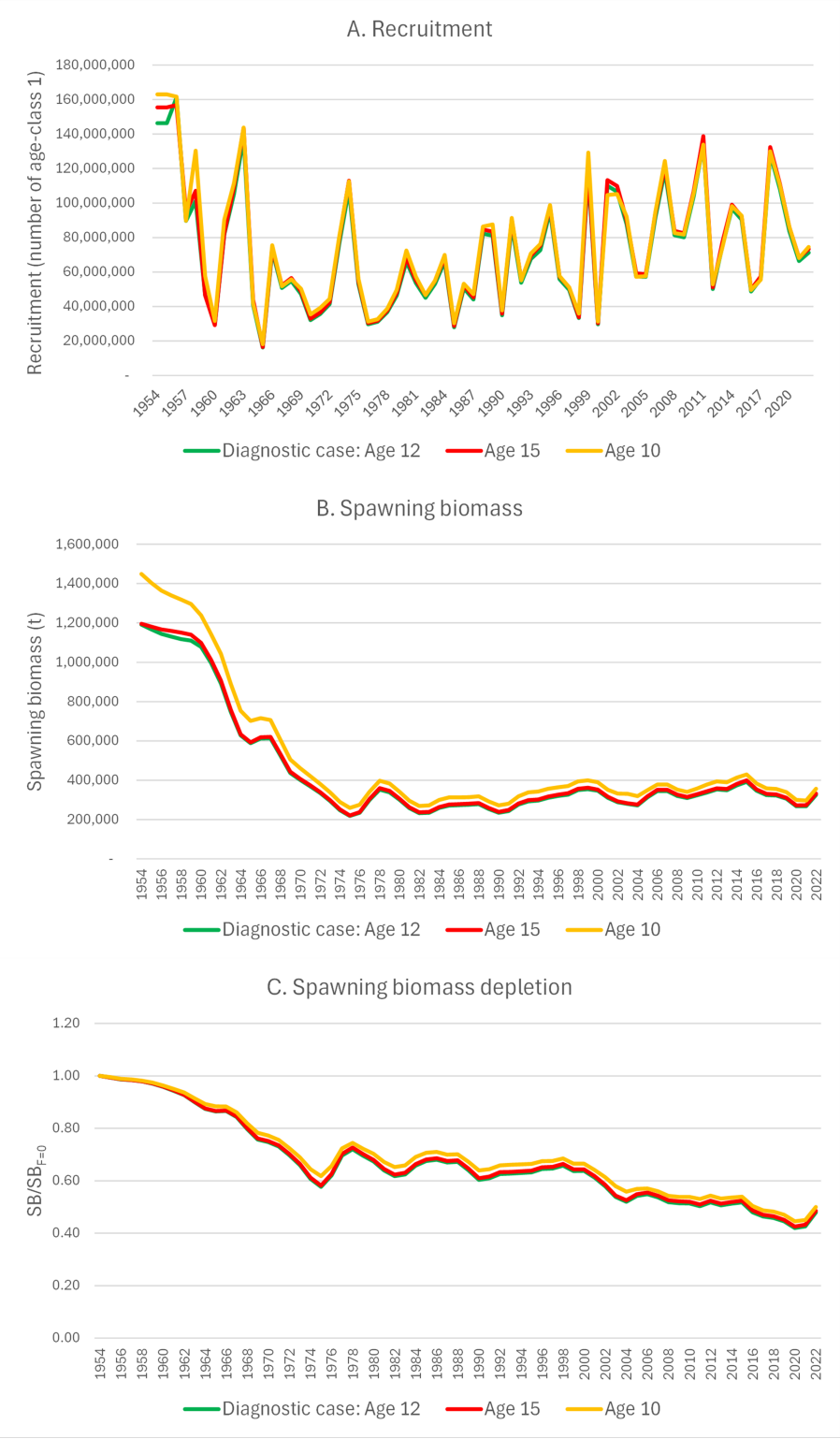


Figure 48: Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ from models with different assumptions regarding the number of annual age classes in the model.

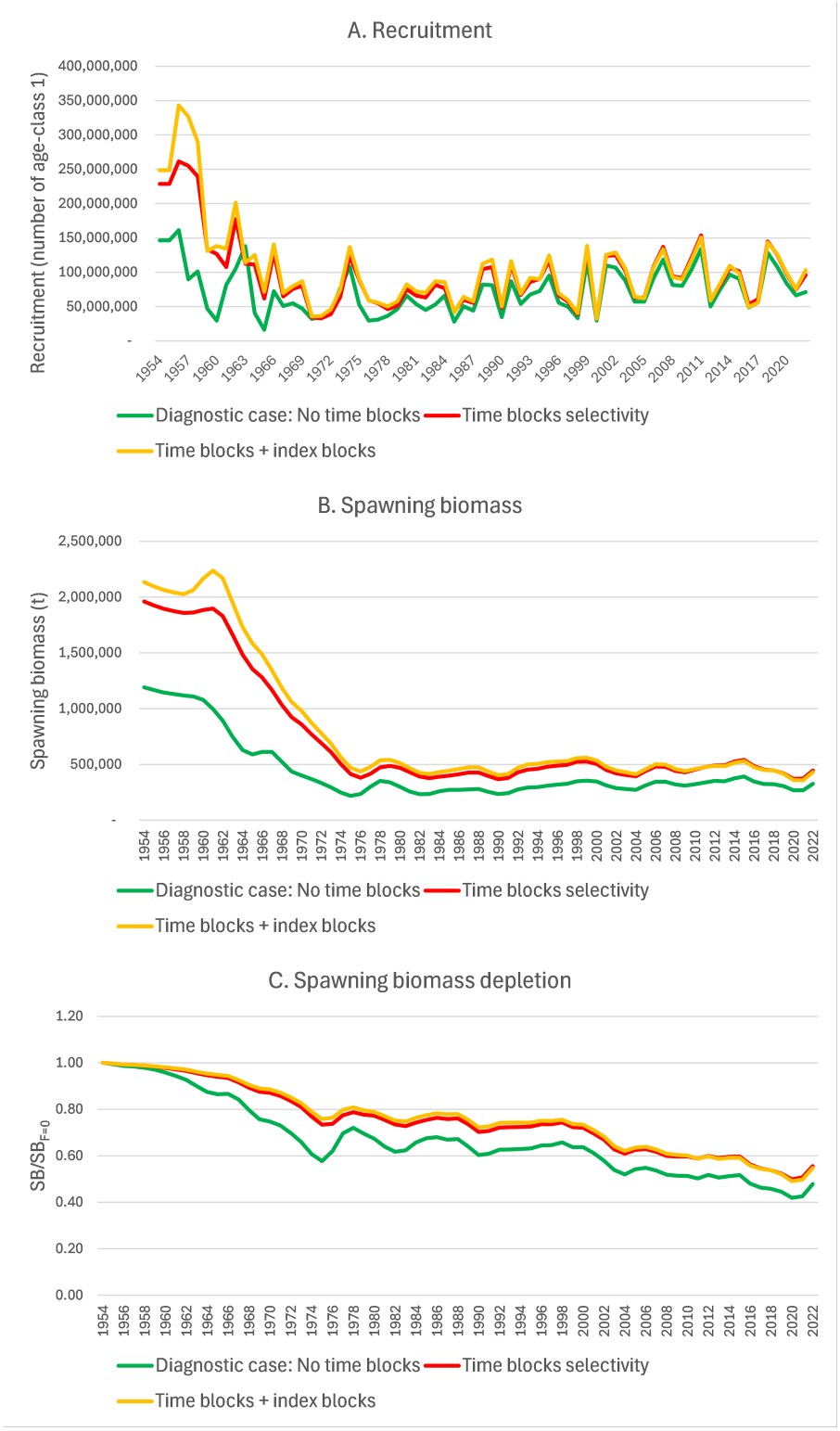
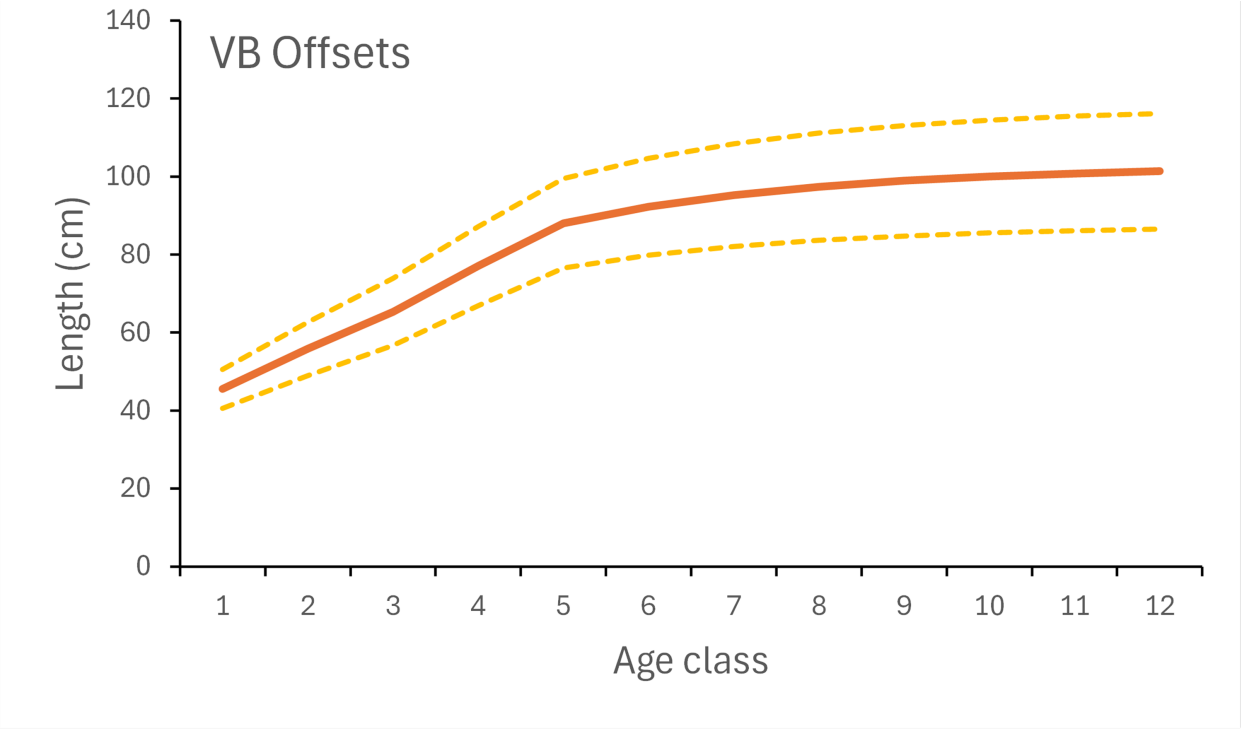


Figure 49: Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ from models with different assumptions regarding time blocks of selectivity and index fishery time blocks.



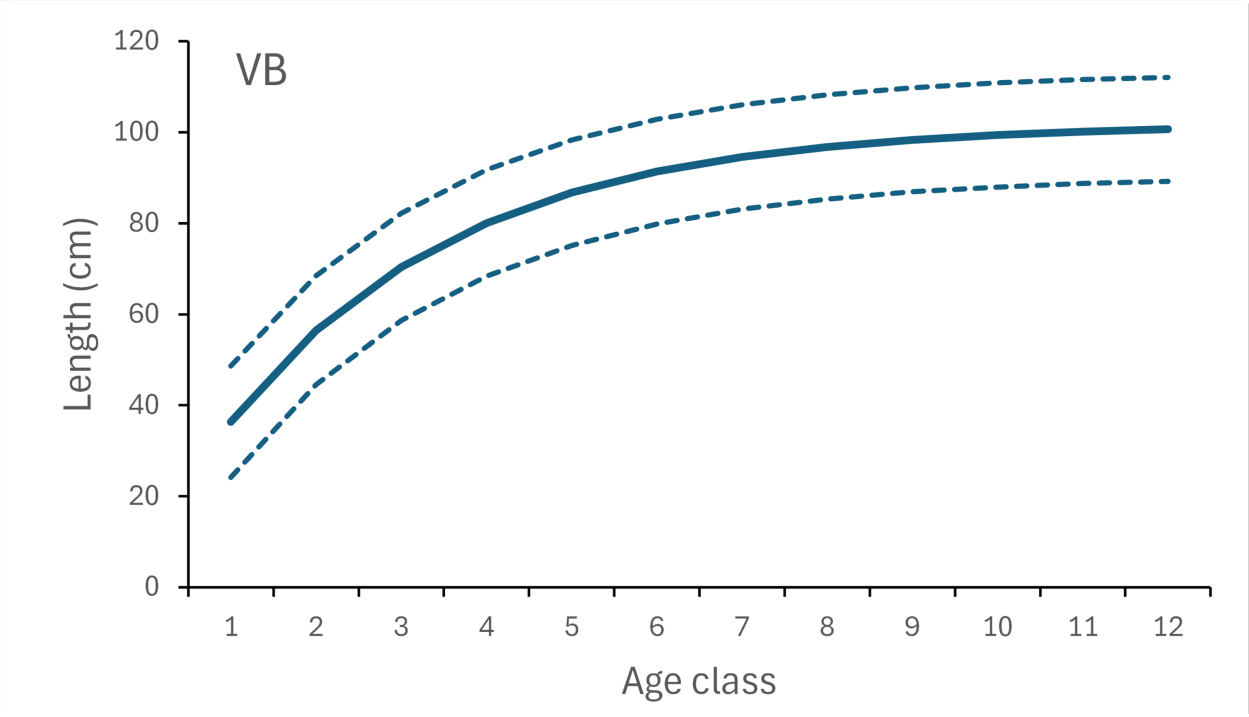


Figure 50: Estimated growth models. The upper panel is the von Bertalanffy offsets model and the lower panel is the standard von Bertalanffy model.

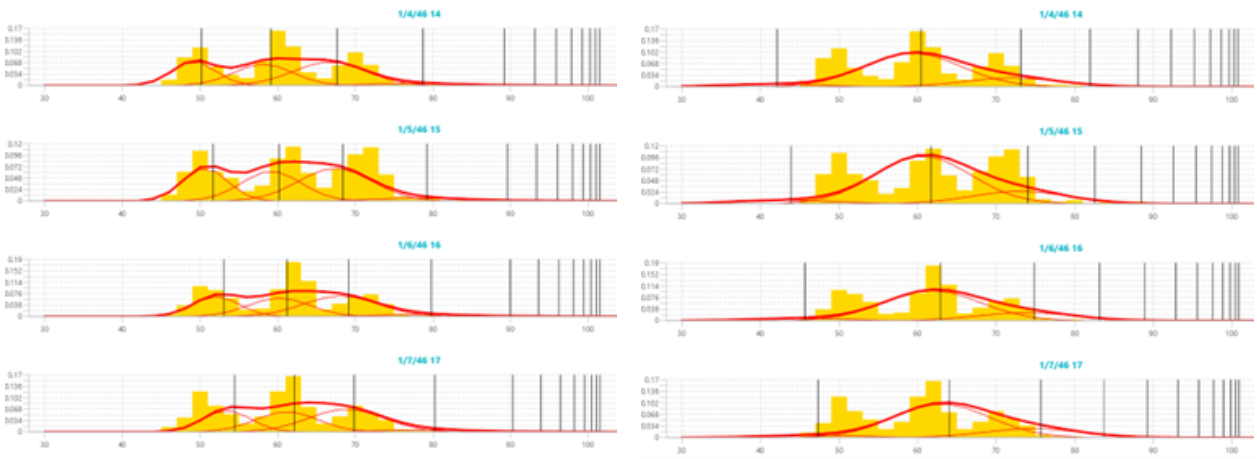


Figure 51: Observed (yellow histograms) and predicted (red lines) length frequency data for a segment of the NZ troll fishery showing clear modal structure. The left panel shows predictions for the VB offsets model and the right panel for the standard VB model.

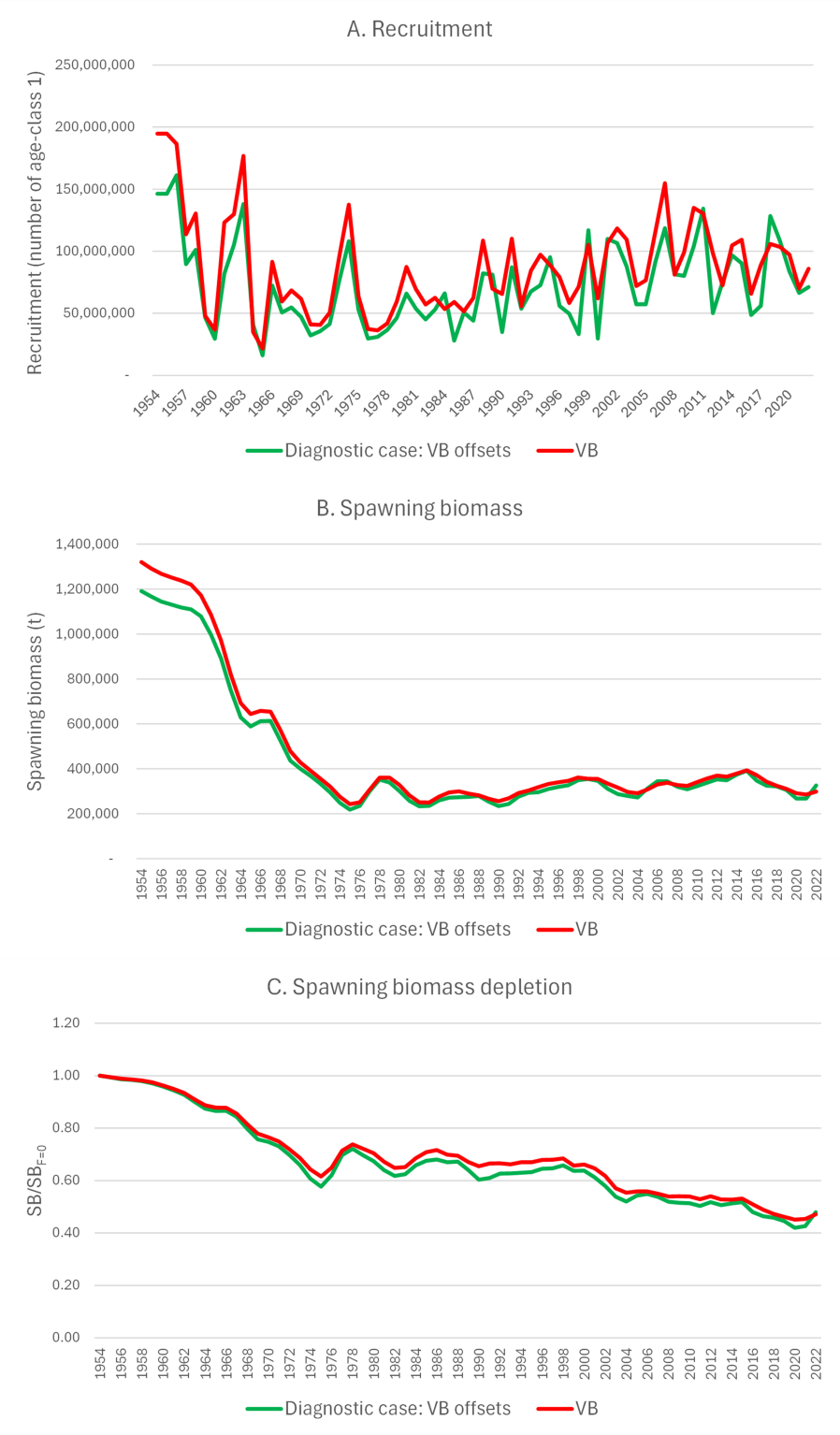


Figure 52: Estimates of recruitment, SB, and $SB_t/SB_{F=0(t)}$ from models with growth model formulations – von Bertalanffy offsets model and standard von Bertalanffy model.

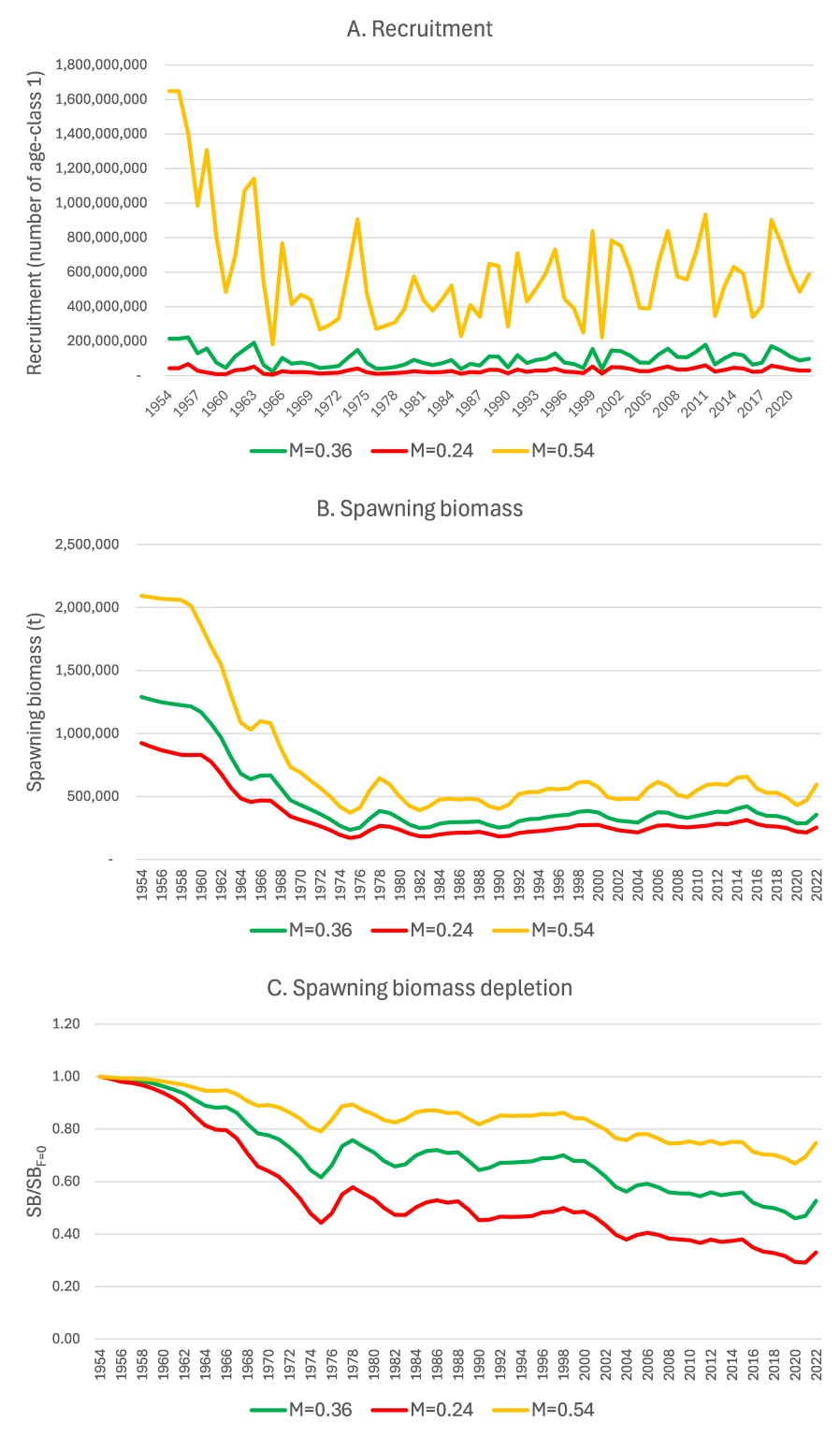


Figure 53: Estimates of recruitment, SB and $SB_t/SB_{F=0(t)}$ from models with different settings for average natural mortality ($\overline{M}$).

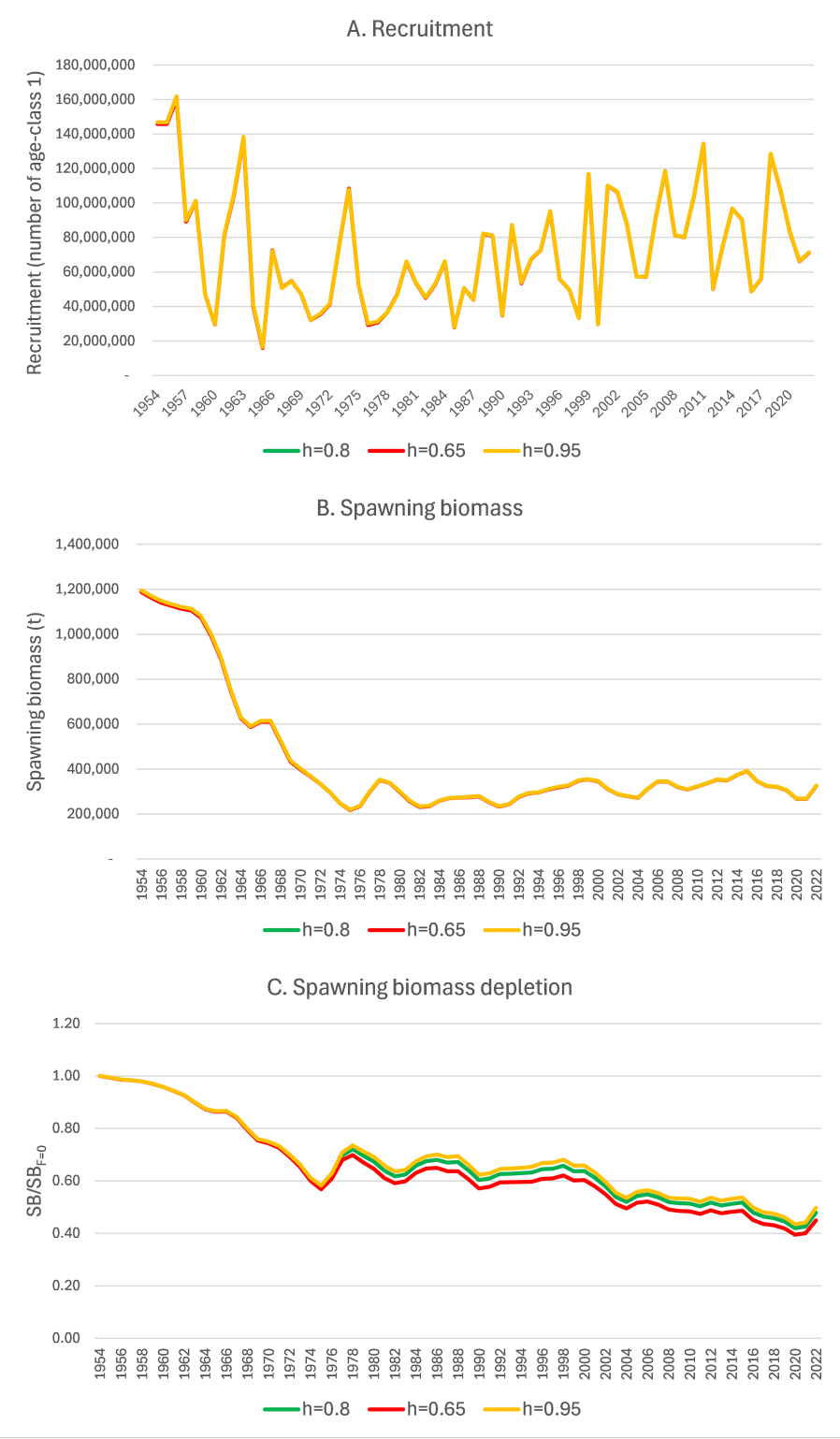


Figure 54: Estimates of recruitment, SB and $SB_t/SB_{F=0(t)}$ from models with different settings for steepness (h).

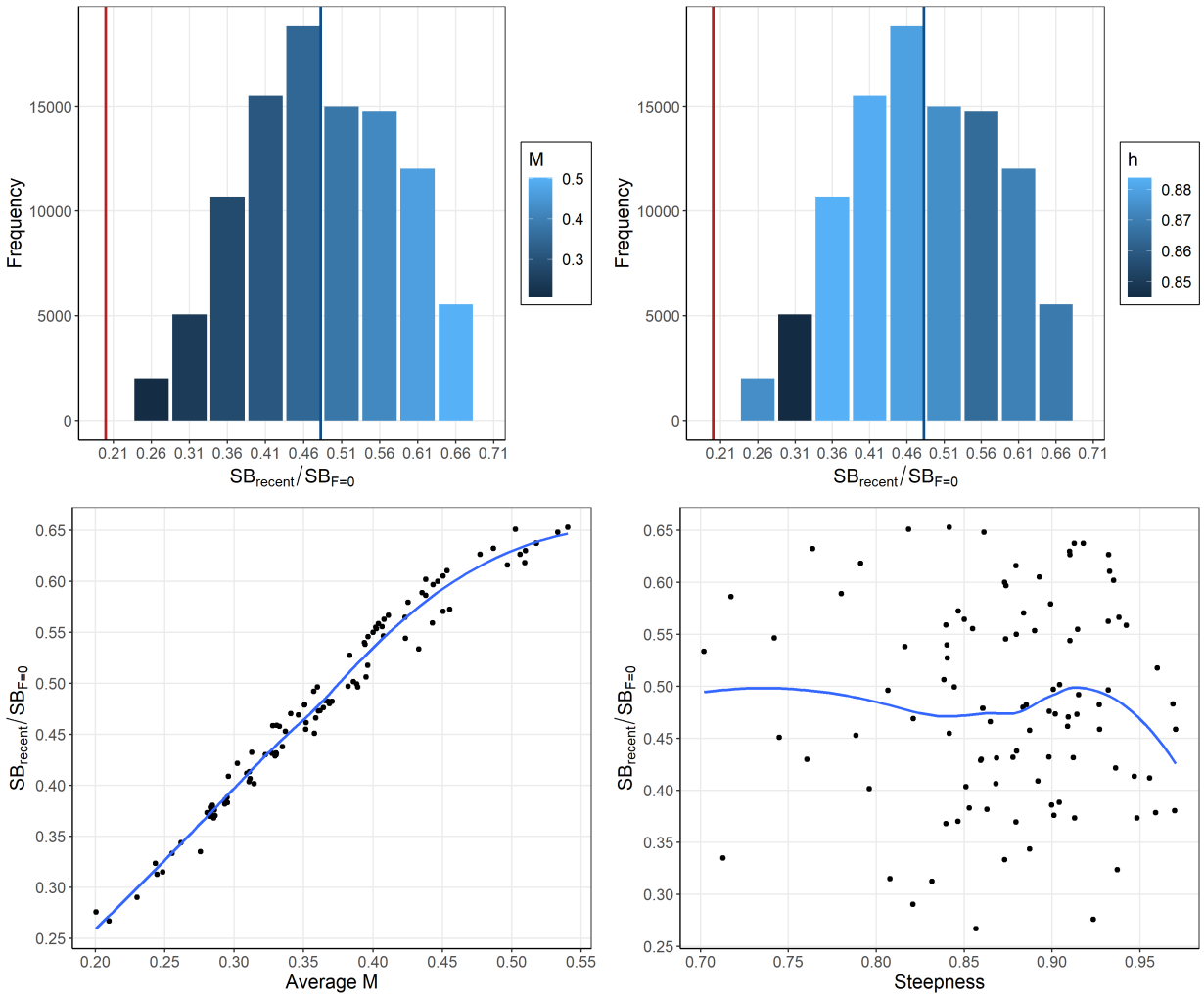


Figure 55: Histograms of Monte-Carlo estimated model uncertainty for $SB_{\text{recent}}/SB_{F=0}$ by $\overline{M}$ (top-left) and h (top-right) with mean line (blue) and $SB_{recent}/SB_{F=0} = 0.2$ line (red). Also includes estimated $SB_{\text{recent}}/SB_{F=0}$ by $\overline{M}$ (bottom-left) and h (bottom-right) for each model in the ensemble with loess smoother.

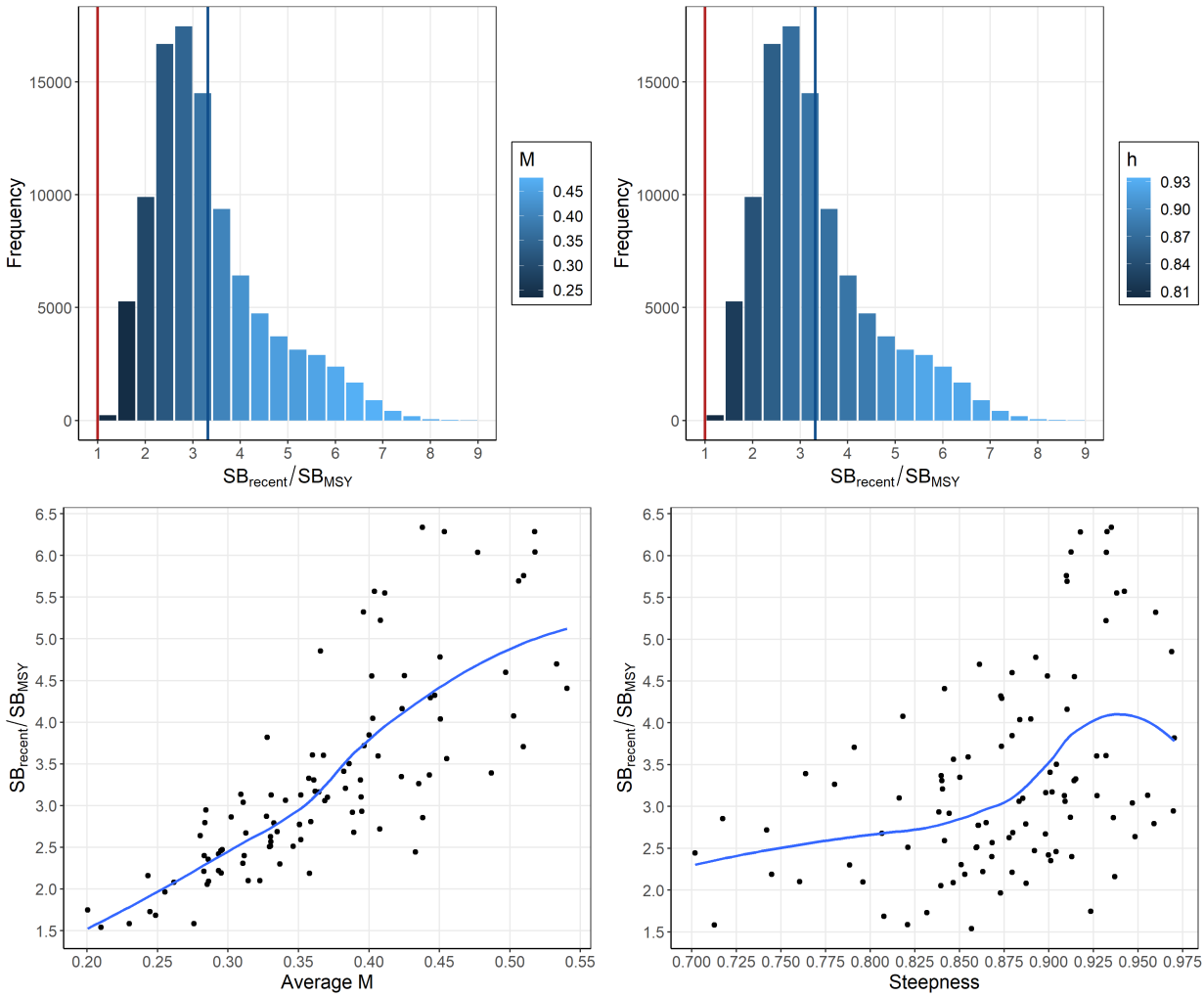


Figure 56: Histograms of Monte-Carlo estimated model uncertainty for $SB_{\text{recent}}/SB_{\text{MSY}}$ by $\overline{M}$ (top-left) and h (top-right) with mean line (blue) and $SB/SB_{MSY} = 1$ line (red). Also includes estimated $SB_{\text{recent}}/SB_{\text{MSY}}$ by $\overline{M}$ (bottom-left) and h (bottom-right) for each model in the ensemble with a loess smoother.

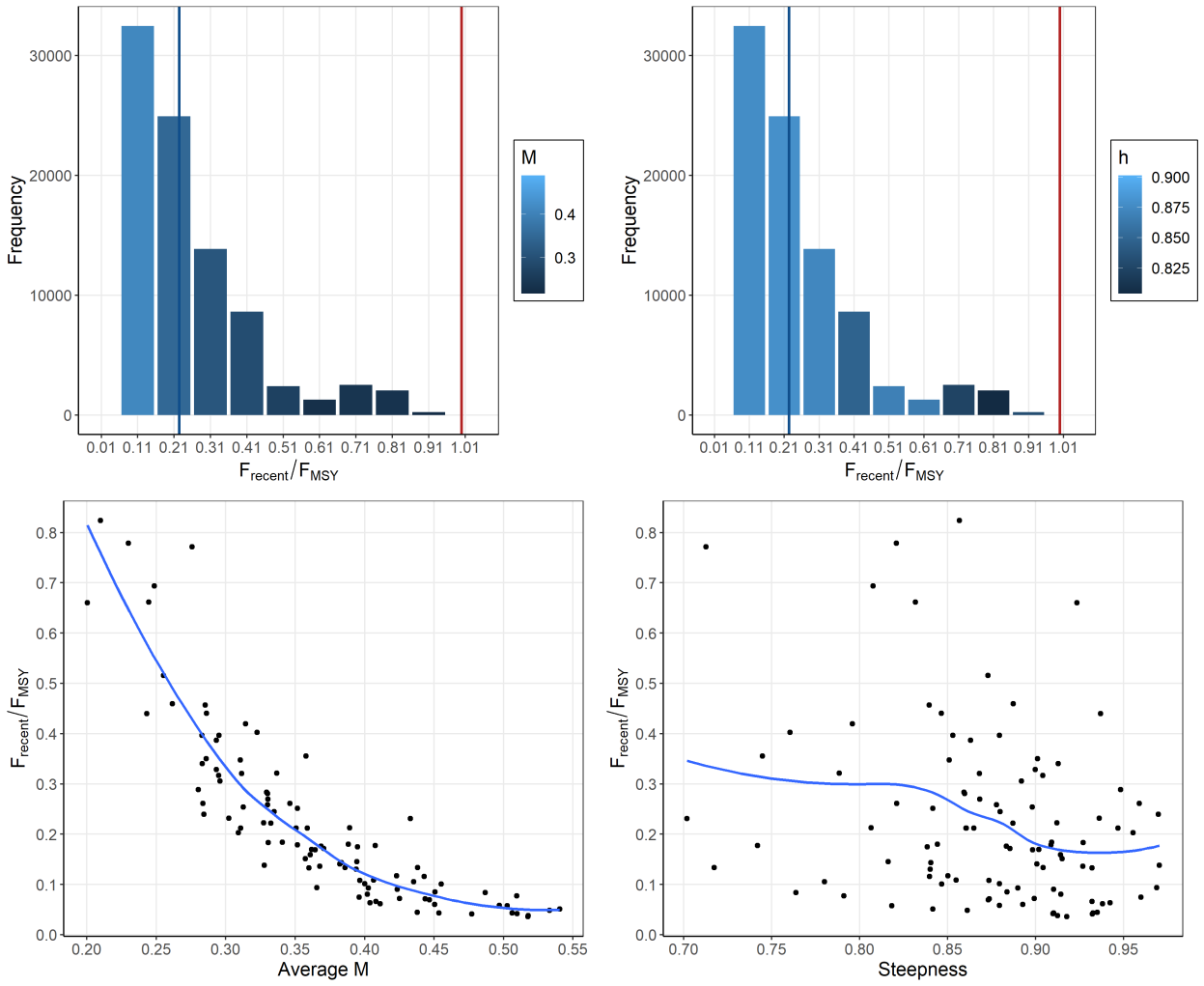


Figure 57: Histograms of Monte-Carlo estimated model uncertainty for $F_{\text{recent}}/F_{\text{MSY}}$ by $\overline{M}$ (top-left) and h (top-right) with mean line (blue) and $F/F_{MSY} = 1$ line (red). Also includes estimated $F_{\text{recent}}/F_{\text{MSY}}$ by $\overline{M}$ (bottom-left) and h (bottom-right) for each model in the ensemble with a loess smoother.

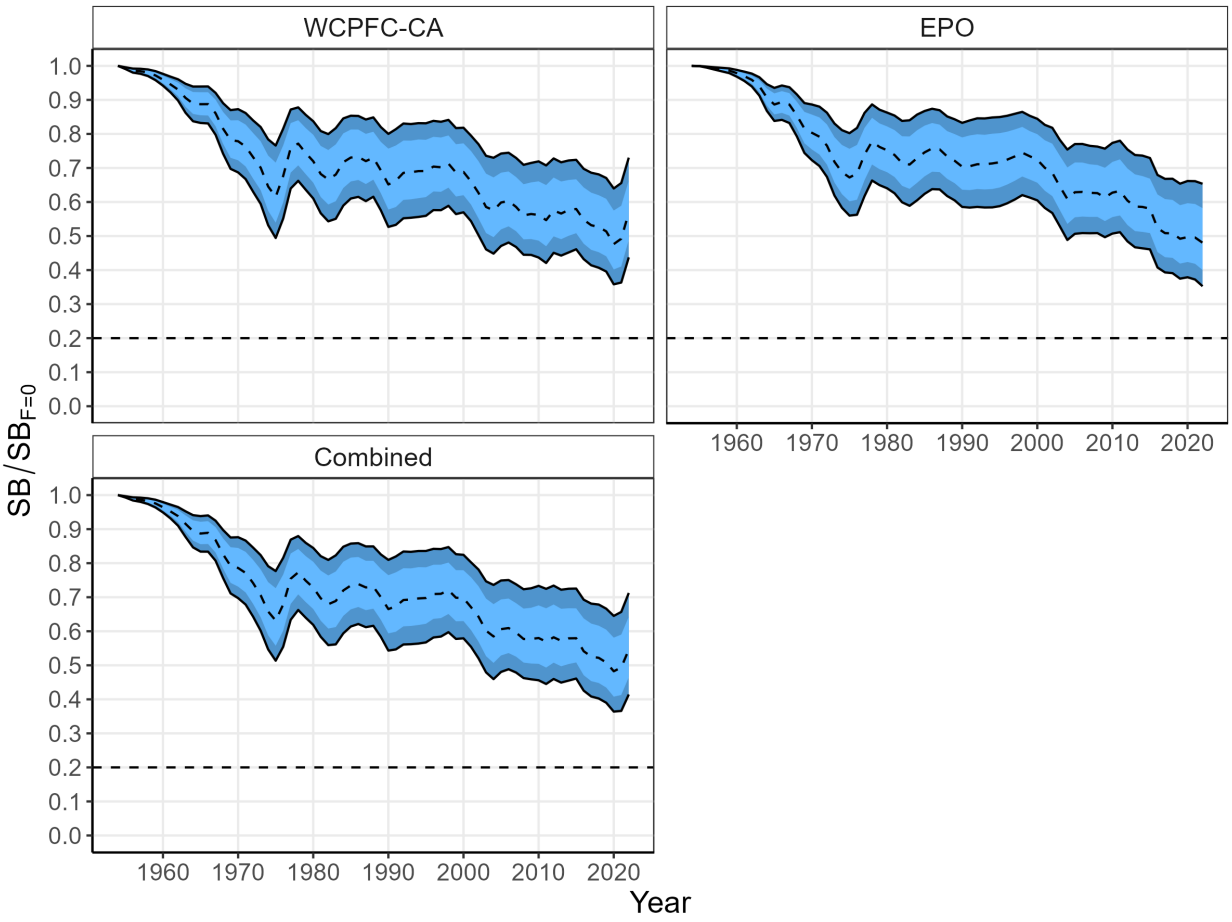


Figure 58: Annual estimated 90% (dark blue) and 75% (light blue) quantiles of $SB_t/SB_{F=0(t)}$ by region from the model ensemble. The dashed line within the interval indicates the median.

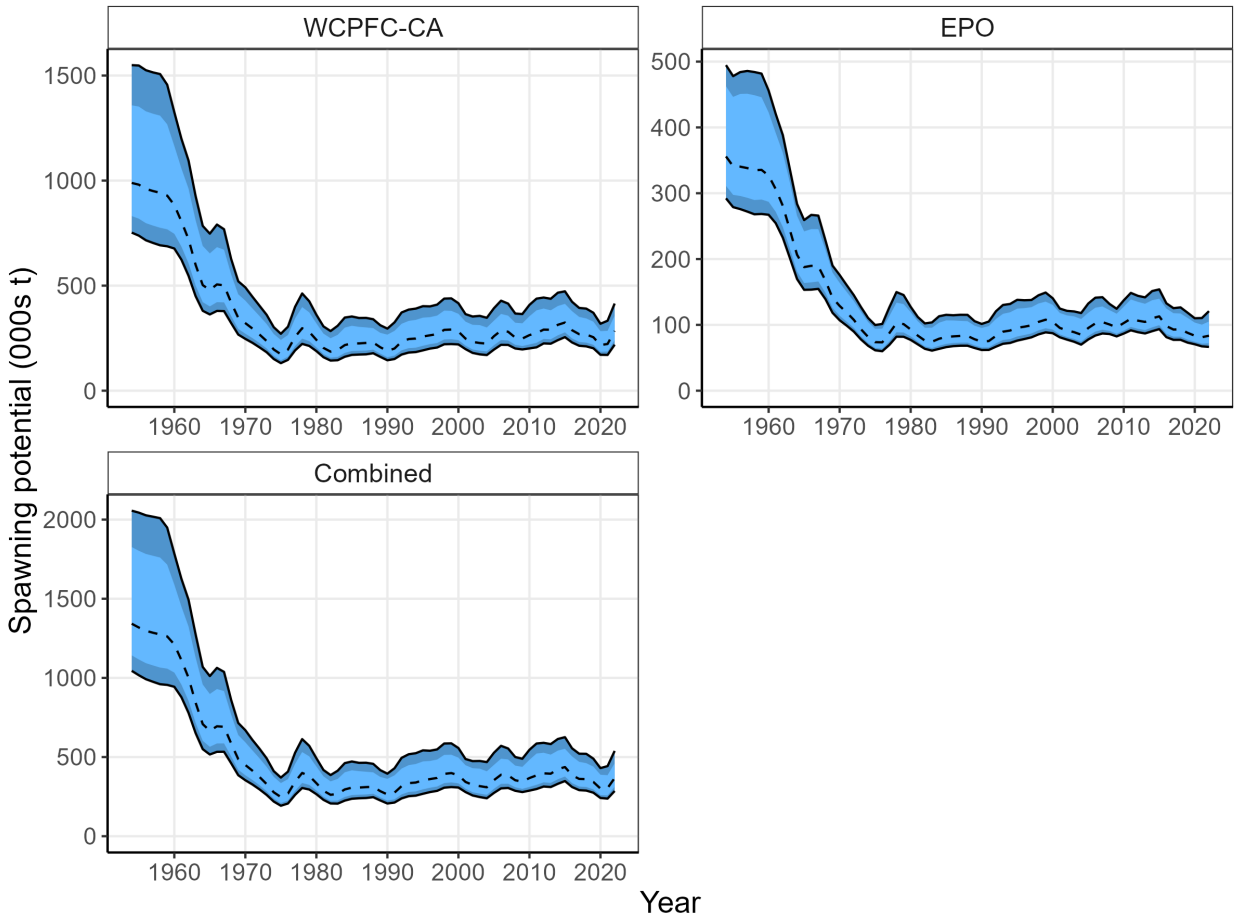


Figure 59: Annual estimated 90% (dark blue) and 75% (light blue) quantiles of SB by region from the model ensemble. The dashed line indicates the median.

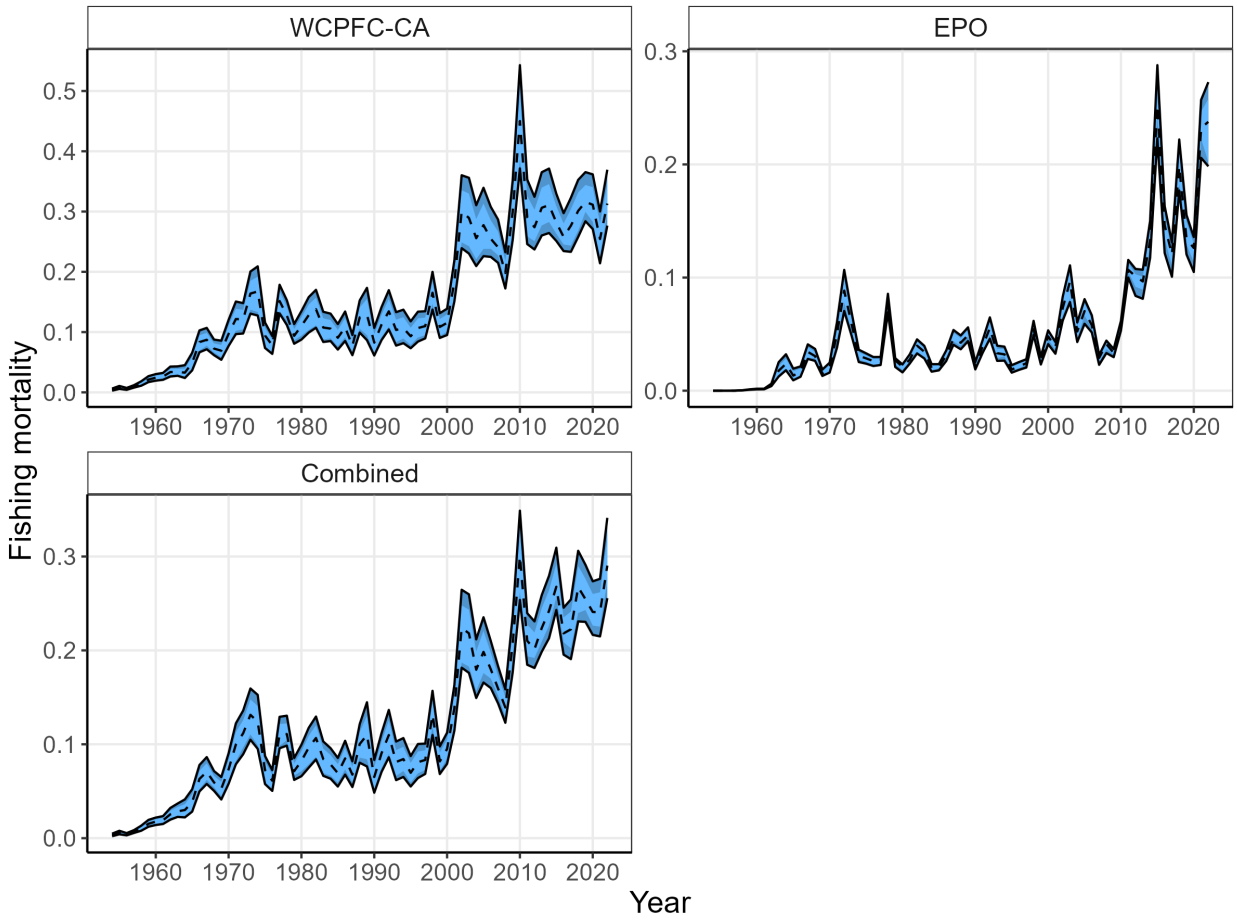


Figure 60: Annual estimated 90% (dark blue) and 75% (light blue) quantiles of aggregated fishing mortality by region from the model ensemble. The dashed line indicates the median.

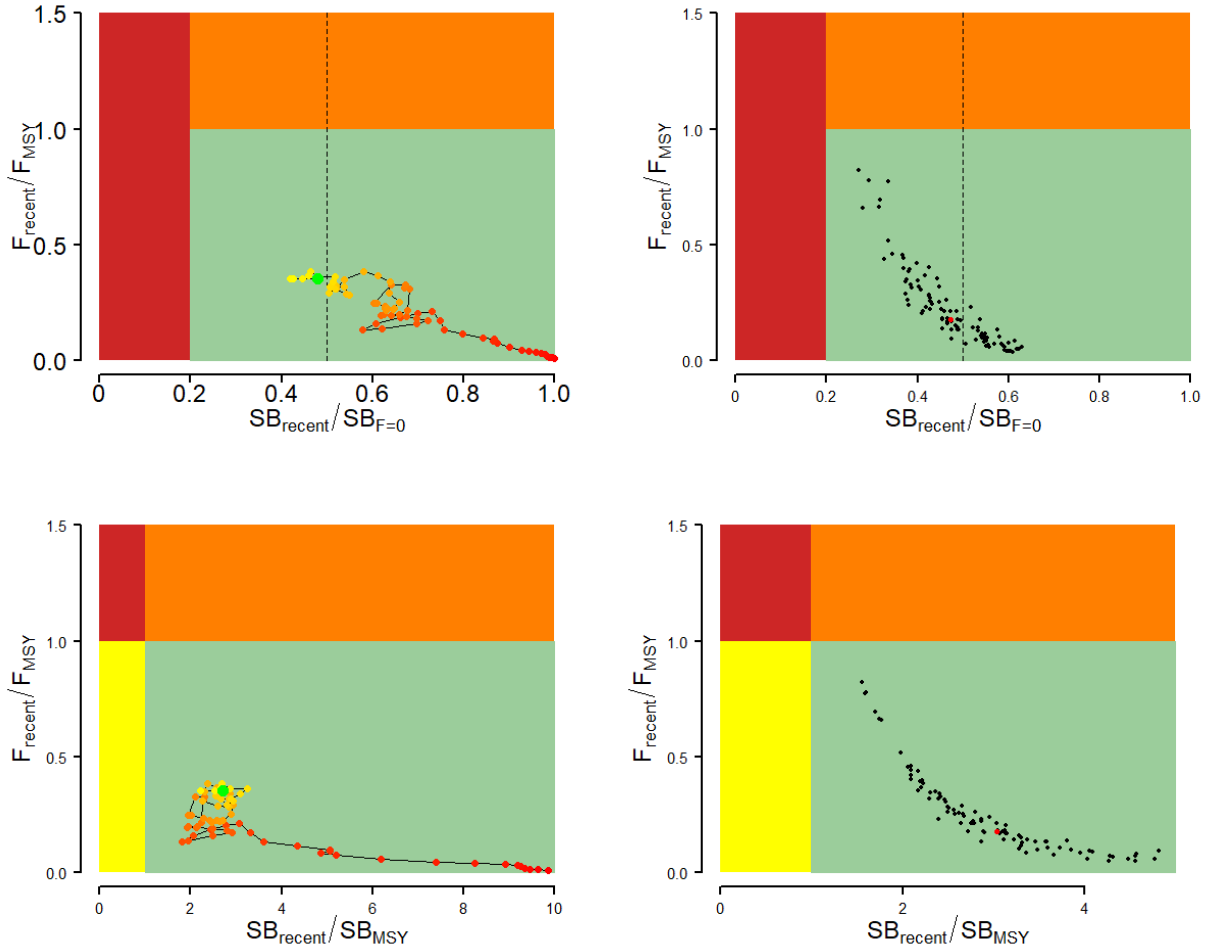


Figure 61: Majuro plots (top) and Kobe plots (bottom) summarising the results for the dynamic MSY analysis (left) and each of the models in the model ensemble for the recent period (2019–2022; right). Majuro plots include dashed line at iTRP estimate (0.5), calculated from the current assessment (Pilling et al., 2024). Colors for dynamic MSY go from red to green over time. The red point in model ensemble represents the median.

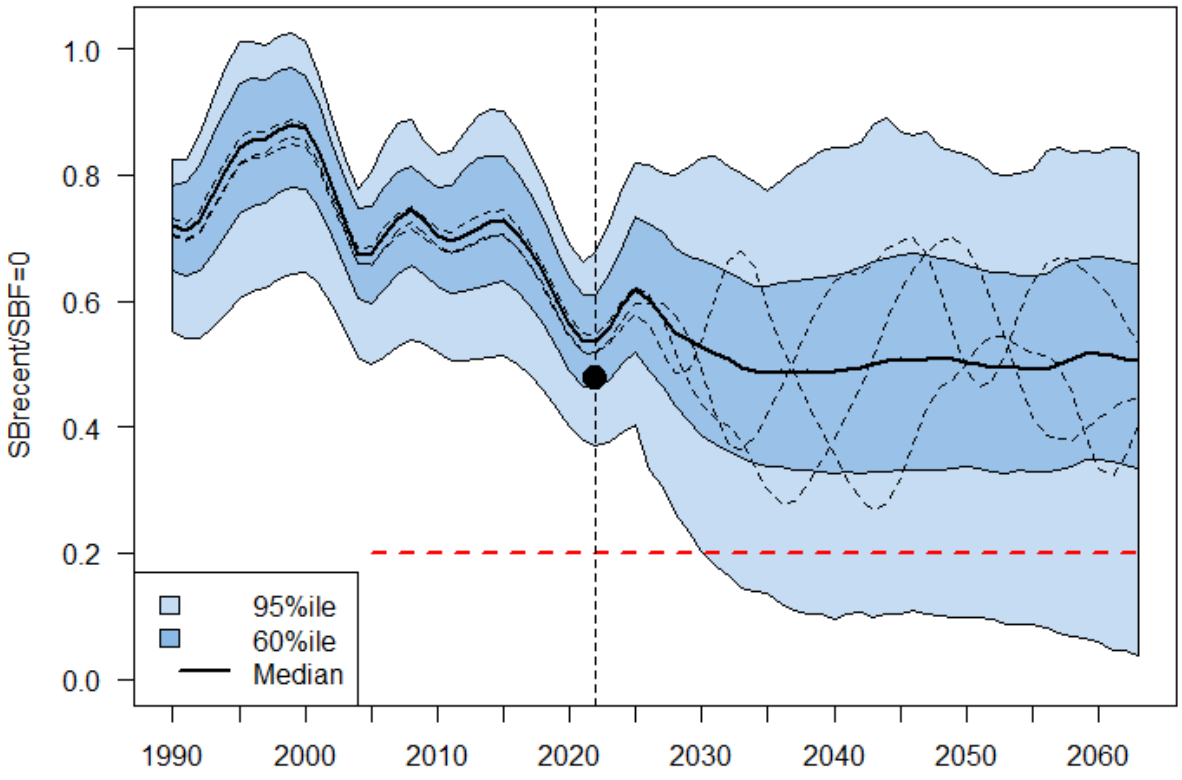


Figure 62: South Pacific albacore SB depletion for the WCPFC Convention Area from the uncertainty grid of assessment model runs for the period 1990 to 2022 (the vertical line at 2022 represents the last year of the assessment), and stochastic projection results for the period 2023 to 2062 assuming actual catch and effort levels in 2022, and that 2022 fishing levels continued. Prior to 2022 the data represent the 60th and 95th percentiles of the uncertainty grid from the assessment models and the median. During the projection period (2023-2062) levels of recruitment variability estimated over the period used to estimate the stock-recruitment relationship (1972-2020) are assumed to continue in the future. The dashed lines indicate three example trajectories (chosen randomly out of 5000) from the model grid. The red dashed line represents the WCPFC agreed limit reference point (0.20). Point represents $SB_{\text{recent}}/SB_{F=0}$, as defined within the stock assessment process.

16 Appendix 1

The text below (known as the "doitall" file) provides the flag settings and phase operations for running the diagnostic case model within a bash shell.

```
# ——————————
# Create initial 00.par file
# ——————————
# MULTIFAN-CLo64 alb.frq alb.ini 00.par # # does not work within a script
# ——————————
#
# ——————————
# PHASE 1 - initial fit with control phases
# ——————————
#
MULTIFAN-CLo64 alb.frq 00.par 01.par # - PHASE1
#
# - control phase type
1 32 7 # sets control, but don't estimate growth
1 387 1
1 246 0 # Produces independent variables report
#
# - initial equilibrium population conditions
2 177 1 # use old totpop scaling method
2 32 1 # and estimate the totpop parameter
2 94 1 # initial age structure based on Z
2 128 10 # average Z for 1st 20 periods
#
#————————————————————————————————————————
# Catch conditioned flags
# general activation
1 373 1 # activated CC with Baranov equation
1 393 0 # activate estimation of: kludged equilib coffs, and implicit fm level regression pars
2 92 2 # specifies the catch-conditioned option with Baranov equation
# - catch equation bounds
2 116 80 # value for Zmax fish in the catch equations
2 189 80 # the fraction of Zmax fish above which the penalty is calculated
1 382 300 # weight for Zmax fish penalty - set to 300 to avoid triggering Zmax flag=1
# De-activate any catch errors flags
-999 1 0
```

-999 4 0
-999 10 0
-999 15 0
-999 13 0
– survey fisheries defined
Index wt Time varying CV
-18 92 20 -18 66 1
-19 92 20 -19 66 1
-20 92 20 -20 66 1
– Grouping flags for survey CPUE
#
- size data
1 141 3 # sets likelihood function for LF data to normal
1 311 1 # tail compression - necessage to limit min sample size
1 312 50 # these settings omit LF samples ¡50 fish
- effective size data sample size based on Francis weighting
-1 49 80
-2 49 183
-3 49 52
-4 49 192
-5 49 287
-6 49 123
-7 49 28
-8 49 428
-9 49 235
-10 49 142
-11 49 28
-12 49 53
-13 49 271
-14 49 128
-15 49 242
-16 49 477
-17 49 1
-18 49 179
-19 49 53
-20 49 64
#
#
- maturity

2 188 2 # Sets option to use weighted spline to dynamically convert
maturity specified in length to maturity at age
- growth
1 173 4 # old comment: # 1st n lengths are independent pars
1 12 0 # turn off estimation growth mean length of first age class
1 13 0 # turn off estimation growth mean length of last age class
1 14 0 # turn off estimation growth K
1 15 0 # turn off estimation growth generic standard deviation length-at-age
1 16 0 # turn off estimation growth length-dependent standard deviation
1 184 0 # turn off estimation growth independent mean length parameters
1 227 0 # turn off estimation growth Richards parameter
#
- recruitment
2 57 1 # sets no. of recruitments per year to 1
2 93 1 # sets no. of recruitments per year to 1
1 400 1 # assume constant mean recruitment for last 2 years
1 398 1 # sets terminal recruitment to arithmetic mean of estimated recrui
1 149 100 # initial recruitment deviate penalties (from average)
#
natural mortality
2 109 3 # set Lorenzen M - turn off as using specified age pars(2)
2 121 0 # do not estimate parameters (set for length)
- movement
2 114 1 # take movement from matrices
#
- selectivity
-999 26 2 # sets length-dependent selectivity option
-999 57 3 # uses cubic spline selectivity
-999 61 3 # with 3 nodes for cubic spline
-8 61 4 # moreflexibility for southern LL
-9 61 4
-10 61 4
#
Selectivity grouping - EPO troll fisheries grouped with US troll, northern DWFN and PICT fisheries grouped across 10S boundary
-1 24 1
-2 24 2
-3 24 1
-4 24 2

-5 24 3
-6 24 3
-7 24 4
-8 24 5
-9 24 6
-10 24 7
-11 24 8
-12 24 9
-13 24 10
-14 24 11
-15 24 12
-16 24 13
-17 24 9
-18 24 14
-19 24 15
-20 24 16
#
#
-999 71 0 # no time-block selectivities
#
sets non-decreasing selectivity for longline fisheries and zero selectivity for age classes 19 and 20 in troll and driftnet fisheries
#
-1 16 0
-2 16 0
-3 16 0
-4 16 0
-5 16 0
-6 16 0
-7 16 0
-8 16 0
-9 16 0
-10 16 0
-11 16 0
-12 16 0
-13 16 0
-14 16 0
-15 16 0
-16 16 0

-17 16 0
-18 16 1
-19 16 0
-20 16 1
#
apply constraint to selectivity = 0 for young age classes for particular fisheries
-1 75 2
-2 75 2
-3 75 2
-4 75 2
-5 75 3
-6 75 3
-7 75 3
-8 75 1
-9 75 1
-10 75 1
-11 75 0
-12 75 0
-13 75 1
-14 75 3
-15 75 1
-16 75 1
-17 75 0
-18 75 2
-19 75 0
-20 75 4
#
PHASE1
#
———
PHASE 2
———
MULTIFAN-CLo64 alb.frq 01.par 02.par # - PHASE2
#
2 113 0 # estimate initpop/totpop scaling parameter - turned off
selectivity
-999 3 0 # all selectivities equal for age classes 47 and 48 (default)
#
catch equation fishing mortality bound

```
# -999 14 10 # Penalties to stop F blowing out
#
# output
1 190 1 # write plot.rep
#
# number of evaluations and convergence threshold
1 1 500 # set max. number of function evaluations per phase to 1000
1 50 -1
#
PHASE2
#
# ———
# PHASE 3
# ———
MULTIFAN-CLo64 alb.frq 02.par 03.par # - PHASE3
#
# fishing impact analysis
-999 55 1 # activate fishery impact analyis - run zero F option
2 171 1 # unfished calculations use BH-SRR multiplier on recruitments
#
# BH-SRR and Yield calculation
#
2 182 0 # Fit BH-SRR to annual recruitments
2 146 1 # estimate SRR parameters
2 145 1 # activates SRR estimation with penalty 1
1 149 0 # recr dev pen set to 0
2 162 0 # don‘t estimate steepness
2 163 0 # use fixed steepness in BH-SRR
2 147 1 # lag between spawning and recruitment
2 148 5 # no. years for averaging F (same as yft)
2 155 1 # but omits the last year
2 161 1 # log-normal bias correction in BH-SRR predictions used for yields
2 199 57 # start period for recruitments used in BH-SRR estimation - 1965 onwards
2 200 1 # end period for recruitments used in BH-SRR estimation - exclude last year
#
# catch equation
-999 14 0 # limit on F per fishing incident - turned off
#
# number of evaluations and threshold max.gradient
```

1 1 2000
1 50 -5
Period of average recruitment used for impact analysis or projections
2 190 0 # Turn off - use entire model period for calculating the average
2 191 0
1 189 1 1 190 1 1 188 1 1 187 1 1 186 1 # full output
2 116 300
PHASE3
———
PHASE 4
———
MULTIFAN-CLo64 alb.frq 03.par 04.par # - PHASE4
Estimate K
1 14 1
1 173 4
1 184 1
1 240 1
1 1 5000
#
#
PHASE4
———
PHASE 5
———
MULTIFAN-CLo64 alb.frq 04.par 05.par # - PHASE5
#
1 1 5000
1 50 -6
Selectivity time block
-1 71 2
-2 71 2
-3 71 2
-4 71 2
-8 71 2
#-14 71 2
#-16 71 2
PHASE5
———
PHASE 6

```
# ———
MULTIFAN-CLo64 alb.frq 05.par 06.par # - PHASE6
#
1 50 -7
1 1 5000
-1 74 4 # Number of seasonal selectivity patterns
-2 74 4
-3 74 4
-4 74 4 # Number of seasonal selectivity patterns
-5 74 4
-6 74 4
-7 74 4
-8 74 1
-9 74 1
-10 74 1
-11 74 1
-12 74 1
-13 74 1
-14 74 2
-15 74 1
-16 74 1
-17 74 1
-18 74 1
-19 74 1
-20 74 1
PHASE6
# ———
# PHASE 7
# ———
MULTIFAN-CLo64 alb.frq 06.par 07.par # - PHASE7
#
1 15 1
PHASE7
# ———
# PHASE 7a
# ———
MULTIFAN-CLo64 alb.frq 07.par 07-CAL.par # - PHASE7a
#
1 240 1
```

```
PHASE7a
# ———
# PHASE 8
# ———
MULTIFAN-CLo64 alb.frq 07-CAL.par 08-CAL.par # - PHASE8
#
1 13 1
1 1 5000
PHASE8
```

17 Appendix 2

Table 10: Monte Carlo ensemble model results (models 1–35) with steepness (h), average natural mortality ($\overline{M}$), maximum gradient, number of negative eigen values (Neg vals), and negative log-likelihood (LL) for total, total LF (length-frequency), CPUE (catch-per-unit-effort), conditional-age-at-length (CAAL), Beverton-Holt steepness, longline WCPO CPUE, troll CPUE, EPO CPUE, LF troll samples number 14–16 (when age class modes are distinct), and LF troll.

Model	h	$\overline{M}$	Max gradient	Neg vals	Total LL	total LF	CPUE	CAAL	BH Steep	CPUE WCPO	CPUE Troll	CPUE EPO	LF troll 14-16	LF troll
1	0.91	0.28	8.13E-07	0	-41,816	-42,454	-166	787	10.3	-85.4	-7.8	-73.1	-131	-4,712
2	0.85	0.42	5.59E-07	0	-41,831	-42,467	-176	796	8.5	-85.6	-14.4	-75.7	-130	-4,704
3	0.87	0.26	1.45E-05	0	-41,810	-42,451	-164	787	10.7	-85.0	-6.6	-72.5	-131	-4,713
4	0.78	0.44	5.75E-07	0	-41,717	-42,555	-179	1,005	7.6	-83.7	-19.0	-76.2	-133	-4,767
5	0.90	0.29	1.36E-05	0	-41,818	-42,455	-167	787	10.2	-85.6	-8.3	-73.3	-131	-4,712
6	0.93	0.45	1.52E-06	0	-41,725	-42,542	-179	985	7.5	-84.6	-18.4	-76.4	-133	-4,758
7	0.87	0.45	7.59E-07	0	-41,724	-42,542	-179	985	7.6	-84.6	-18.2	-76.4	-133	-4,758
8	0.80	0.31	1.33E-06	0	-41,822	-42,456	-169	787	9.9	-85.8	-9.4	-73.9	-131	-4,712
9	0.91	0.35	2.26E-06	0	-41,828	-42,459	-172	786	9.4	-86.1	-11.0	-74.5	-131	-4,711
10	0.84	0.29	2.29E-06	0	-41,817	-42,454	-167	787	10.3	-85.4	-7.9	-73.2	-131	-4,712
11	0.94	0.24	6.98E-07	0	-41,807	-42,450	-163	787	10.9	-84.8	-5.9	-72.1	-131	-4,713
12	0.89	0.33	1.15E-06	0	-41,821	-42,463	-171	797	9.5	-85.4	-11.1	-74.3	-131	-4,705
13	0.97	0.28	2.37E-06	0	-41,816	-42,454	-166	787	10.3	-85.5	-7.9	-73.1	-131	-4,712
14	0.93	0.48	9.10E-07	0	-41,835	-42,469	-177	797	8.1	-85.5	-15.7	-76.2	-130	-4,704
15	0.88	0.33	5.12E-07	0	-41,825	-42,457	-170	786	9.7	-85.9	-10.1	-74.1	-131	-4,711
16	0.84	0.44	2.33E-06	0	-41,838	-42,463	-176	786	8.6	-86.3	-13.8	-75.9	-131	-4,709
17	0.97	0.33	4.09E-06	0	-41,820	-42,463	-170	797	9.6	-85.4	-10.9	-74.2	-131	-4,705
18	0.90	0.36	6.40E-07	0	-41,830	-42,460	-172	786	9.3	-86.1	-11.4	-74.7	-131	-4,711
19	0.90	0.36	5.64E-07	0	-41,830	-42,460	-172	786	9.3	-86.1	-11.4	-74.7	-131	-4,711
20	0.91	0.51	5.93E-06	0	-41,842	-42,465	-178	787	8.1	-86.2	-15.3	-76.6	-131	-4,708
21	0.87	0.44	4.07E-06	0	-41,724	-42,542	-179	985	7.6	-84.6	-18.1	-76.3	-133	-4,758
22	0.84	0.54	4.54E-05	0	-41,843	-42,466	-179	787	7.9	-86.1	-15.7	-76.9	-131	-4,707
23	0.85	0.41	9.59E-07	0	-41,722	-42,542	-178	984	7.9	-84.6	-17.2	-75.8	-133	-4,759
24	0.91	0.52	8.09E-06	0	-41,842	-42,465	-178	787	8.1	-86.2	-15.4	-76.7	-131	-4,708
25	0.84	0.39	2.02E-06	0	-41,833	-42,461	-174	786	9.0	-86.2	-12.3	-75.1	-131	-4,710
26	0.74	0.36	8.49E-07	0	-41,829	-42,459	-172	786	9.3	-86.1	-11.2	-74.7	-131	-4,711
27	0.82	0.50	7.52E-07	0	-41,726	-42,543	-181	985	7.2	-84.5	-19.3	-77.0	-133	-4,757
28	0.92	0.20	6.20E-06	0	-41,794	-42,452	-158	798	11.7	-83.8	-3.4	-70.9	-130	-4,705
29	0.89	0.30	2.76E-06	0	-41,815	-42,460	-168	797	10.0	-85.1	-9.3	-73.6	-131	-4,705
30	0.94	0.41	7.88E-07	0	-41,722	-42,542	-178	984	7.8	-84.6	-17.3	-75.8	-133	-4,758
31	0.79	0.51	2.04E-06	0	-41,728	-42,531	-182	972	7.3	-85.9	-18.7	-77.5	-133	-4,755
32	0.82	0.23	4.06E-07	0	-41,804	-42,449	-161	787	11.1	-84.5	-5.0	-71.8	-131	-4,713
33	0.82	0.35	1.68E-06	0	-41,823	-42,464	-172	797	9.3	-85.4	-11.7	-74.6	-131	-4,705
34	0.83	0.24	8.54E-07	0	-41,808	-42,450	-163	787	10.8	-84.8	-5.9	-72.2	-131	-4,713
35	0.90	0.29	8.40E-07	0	-41,817	-42,454	-167	787	10.3	-85.5	-8.0	-73.2	-131	-4,712

Table 11: Monte Carlo ensemble model results (models 36–70) with steepness (h), average natural mortality ($\overline{M}$), maximum gradient, number of negative eigen values (Neg vals), and negative log-likelihood (LL) for total, total LF (length-frequency), CPUE (catch-per-unit-effort), conditional-age-at-length (CAAL), Beverton-Holt steepness, longline WCPO CPUE, troll CPUE, EPO CPUE, LF troll samples number 14–16 (when age class modes are distinct), and LF troll.

Model	h	$\overline{M}$	Max gradient	Neg vals	Total LL	total LF	CPUE	CAAL	BH Steep	CPUE WCPO	CPUE Troll	CPUE EPO	LF troll 14-16	LF troll
36	0.93	0.33	1.20E-06	0	-41,820	-42,463	-171	797	9.5	-85.4	-11.0	-74.2	-131	-4,705
37	0.86	0.21	3.10E-06	0	-41,799	-42,447	-159	788	11.5	-84.1	-3.7	-71.0	-131	-4,713
38	0.96	0.40	3.43E-06	0	-41,834	-42,461	-174	786	9.0	-86.3	-12.5	-75.2	-131	-4,710
39	0.90	0.29	7.89E-07	0	-41,819	-42,455	-167	787	10.1	-85.6	-8.4	-73.4	-131	-4,712
40	0.91	0.42	8.50E-07	0	-41,837	-42,463	-175	786	8.7	-86.3	-13.3	-75.6	-131	-4,709
41	0.90	0.39	3.18E-06	0	-41,833	-42,461	-173	786	9.1	-86.2	-12.2	-75.1	-131	-4,710
42	0.88	0.33	2.43E-06	0	-41,826	-42,458	-170	786	9.6	-86.0	-10.3	-74.2	-131	-4,711
43	0.85	0.29	1.10E-06	0	-41,817	-42,454	-167	787	10.3	-85.5	-8.0	-73.2	-131	-4,712
44	0.82	0.39	7.63E-07	0	-41,721	-42,541	-177	984	8.0	-84.6	-16.8	-75.6	-133	-4,759
45	0.86	0.35	1.01E-06	0	-41,823	-42,464	-172	797	9.3	-85.5	-11.9	-74.6	-131	-4,705
46	0.84	0.35	7.37E-07	0	-41,828	-42,459	-172	786	9.4	-86.1	-10.9	-74.5	-131	-4,711
47	0.81	0.25	3.16E-06	0	-41,809	-42,450	-163	787	10.8	-84.8	-6.2	-72.4	-131	-4,713
48	0.84	0.39	1.63E-06	0	-41,721	-42,541	-177	984	8.0	-84.6	-16.8	-75.6	-133	-4,759
49	0.91	0.34	9.41E-07	0	-41,822	-42,463	-171	797	9.4	-85.4	-11.5	-74.4	-131	-4,705
50	0.88	0.50	2.42E-06	0	-41,842	-42,465	-178	787	8.2	-86.3	-15.0	-76.5	-131	-4,708
51	0.86	0.36	4.25E-06	0	-41,829	-42,459	-172	786	9.3	-86.1	-11.2	-74.7	-131	-4,711
52	0.76	0.49	2.48E-06	0	-41,726	-42,543	-181	985	7.3	-84.6	-19.1	-76.9	-133	-4,757
53	0.87	0.40	1.25E-06	0	-41,721	-42,541	-177	984	8.0	-84.6	-16.9	-75.6	-133	-4,759
54	0.88	0.45	2.74E-05	0	-41,839	-42,464	-176	786	8.5	-86.3	-14.0	-76.0	-131	-4,709
55	0.93	0.36	7.11E-07	0	-41,825	-42,464	-173	797	9.2	-85.5	-12.3	-74.8	-131	-4,705
56	0.93	0.41	9.81E-07	0	-41,722	-42,542	-178	984	7.9	-84.6	-17.2	-75.8	-133	-4,759
57	0.71	0.28	1.13E-06	0	-41,815	-42,453	-166	787	10.4	-85.3	-7.4	-73.0	-131	-4,712
58	0.84	0.39	9.69E-07	0	-41,834	-42,461	-174	786	9.0	-86.3	-12.5	-75.2	-131	-4,710
59	0.91	0.33	5.46E-06	0	-41,824	-42,457	-170	786	9.7	-85.9	-10.0	-74.1	-131	-4,711
60	0.89	0.45	1.15E-06	0	-41,724	-42,542	-179	985	7.5	-84.6	-18.3	-76.4	-133	-4,758
61	0.91	0.40	2.71E-06	0	-41,722	-42,542	-177	984	7.9	-84.6	-17.1	-75.7	-133	-4,759
62	0.89	0.26	4.59E-06	0	-41,812	-42,452	-165	787	10.6	-85.1	-7.0	-72.7	-131	-4,713
63	0.89	0.40	1.37E-06	0	-41,722	-42,542	-177	984	7.9	-84.6	-17.1	-75.7	-133	-4,759
64	0.72	0.44	9.46E-07	0	-41,717	-42,555	-179	1,005	7.6	-83.6	-19.0	-76.2	-133	-4,767
65	0.76	0.32	1.56E-06	0	-41,819	-42,462	-170	797	9.6	-85.3	-10.6	-74.2	-131	-4,705
66	0.85	0.30	4.35E-06	0	-41,819	-42,455	-167	787	10.1	-85.6	-8.4	-73.4	-131	-4,712
67	0.86	0.53	1.53E-06	0	-41,843	-42,466	-179	787	8.0	-86.2	-15.7	-76.8	-131	-4,707
68	0.87	0.31	9.33E-07	0	-41,822	-42,456	-169	787	9.9	-85.7	-9.2	-73.8	-131	-4,712
69	0.91	0.51	3.88E-06	0	-41,842	-42,465	-178	787	8.1	-86.3	-15.2	-76.6	-131	-4,708
70	0.74	0.41	9.78E-07	0	-41,722	-42,542	-178	984	7.9	-84.6	-17.2	-75.8	-133	-4,759

Table 12: Monte Carlo ensemble model results (models 71–100) with steepness (h), average natural mortality ($\overline{M}$), maximum gradient, number of negative eigen values (Neg vals), and negative log-likelihood (LL) for total, total LF (length-frequency), CPUE (catch-per-unit-effort), conditional-age-at-length (CAAL), Beverton-Holt steepness, longline WCPO CPUE, troll CPUE, EPO CPUE, LF troll samples number 14–16 (when age class modes are distinct), and LF troll.

Model	h	$\overline{M}$	Max gradient	Neg vals	Total LL	total LF	CPUE	CAAL	BH Steep	CPUE WCPO	CPUE Troll	CPUE EPO	LF troll 14-16	LF troll
71	0.88	0.40	1.19E-06	0	-41,721	-42,542	-177	984	7.9	-84.6	-17.0	-75.6	-133	-4,759
72	0.90	0.31	2.33E-06	0	-41,818	-42,462	-169	797	9.8	-85.3	-10.2	-73.9	-131	-4,705
73	0.90	0.38	8.23E-07	0	-41,832	-42,461	-173	786	9.1	-86.2	-12.1	-75.0	-131	-4,710
74	0.86	0.29	9.60E-07	0	-41,818	-42,455	-167	787	10.2	-85.5	-8.3	-73.4	-131	-4,712
75	0.96	0.31	6.03E-06	0	-41,821	-42,456	-169	787	9.9	-85.7	-9.1	-73.7	-131	-4,712
76	0.92	0.52	2.50E-06	0	-41,842	-42,465	-178	787	8.1	-86.2	-15.4	-76.7	-131	-4,708
77	0.79	0.34	2.04E-06	0	-41,821	-42,463	-171	797	9.4	-85.4	-11.3	-74.4	-131	-4,705
78	0.85	0.46	1.44E-05	0	-41,839	-42,464	-176	786	8.5	-86.3	-14.1	-76.0	-131	-4,709
79	0.84	0.38	3.21E-07	0	-41,720	-42,541	-176	984	8.1	-84.5	-16.5	-75.4	-133	-4,759
80	0.95	0.28	6.68E-07	0	-41,816	-42,453	-166	787	10.4	-85.4	-7.7	-73.0	-131	-4,712
81	0.88	0.37	8.00E-07	0	-41,831	-42,460	-173	786	9.2	-86.2	-11.6	-74.8	-131	-4,711
82	0.70	0.43	6.02E-07	0	-41,837	-42,463	-176	786	8.7	-86.3	-13.6	-75.9	-131	-4,709
83	0.92	0.36	9.52E-07	0	-41,824	-42,464	-172	797	9.2	-85.5	-12.2	-74.7	-131	-4,705
84	0.94	0.40	1.25E-06	0	-41,722	-42,542	-177	984	7.9	-84.6	-17.1	-75.7	-133	-4,759
85	0.85	0.31	2.45E-06	0	-41,821	-42,456	-169	787	9.9	-85.7	-9.2	-73.8	-131	-4,712
86	0.91	0.36	2.21E-06	0	-41,829	-42,459	-172	786	9.3	-86.1	-11.3	-74.7	-131	-4,711
87	0.90	0.43	2.82E-07	0	-41,723	-42,542	-178	985	7.7	-84.6	-17.7	-76.1	-133	-4,758
88	0.86	0.33	2.38E-06	0	-41,825	-42,457	-170	786	9.7	-85.9	-10.0	-74.1	-131	-4,711
89	0.81	0.39	4.53E-06	0	-41,833	-42,461	-174	786	9.0	-86.2	-12.3	-75.2	-131	-4,710
90	0.93	0.37	2.42E-06	0	-41,830	-42,460	-173	786	9.2	-86.2	-11.6	-74.8	-131	-4,711
91	0.88	0.28	7.24E-07	0	-41,816	-42,454	-166	787	10.3	-85.4	-7.8	-73.1	-131	-4,712
92	0.97	0.37	8.13E-07	0	-41,830	-42,460	-172	786	9.3	-86.2	-11.5	-74.7	-131	-4,711
93	0.86	0.35	9.62E-07	0	-41,823	-42,464	-172	797	9.3	-85.5	-11.9	-74.6	-131	-4,705
94	0.94	0.44	2.47E-06	0	-41,717	-42,555	-179	1,005	7.6	-83.7	-19.0	-76.2	-133	-4,767
95	0.89	0.37	6.06E-07	0	-41,831	-42,460	-173	786	9.2	-86.2	-11.7	-74.8	-131	-4,711
96	0.86	0.33	7.34E-07	0	-41,825	-42,457	-170	786	9.7	-85.9	-10.1	-74.1	-131	-4,711
97	0.95	0.31	3.75E-06	0	-41,822	-42,456	-169	787	9.9	-85.8	-9.2	-73.7	-131	-4,712
98	0.87	0.33	2.77E-06	0	-41,825	-42,457	-170	786	9.7	-85.9	-10.1	-74.1	-131	-4,711
99	0.94	0.30	8.16E-07	0	-41,816	-42,461	-169	797	9.9	-85.2	-9.7	-73.7	-131	-4,705
100	0.96	0.28	8.85E-07	0	-41,816	-42,454	-166	787	10.3	-85.5	-7.9	-73.1	-131	-4,712