



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

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

WCPFC-SC22-2026-SA-WP06_Rev02

1 October 2026

**Submitted by K. Kim¹, T. Teears¹, J. Hampton², N. Davies³, A. Magnusson¹, C. Castillo Jordan¹,
T. Peatman⁴, J. Scutt Phillips¹, P. Hamer¹**

¹ Oceanic Fisheries Programme of the Pacific Community

² Thunnidae Consulting

³ TeTakina Ltd

⁴ Shearwater Analytics

Revision 1

- Added results and tables for parameters on or close to bounds and the strongest estimated-parameter correlations.
- Updated diagnostic, model-fit, likelihood-profile and population-dynamics figures; added ASPM, Hessian and ensemble-projection figures while retaining the original figure files.
- Expanded the fishery-definition, eight-group Dirichlet-multinomial and MFCL tag-reporting-rate tables using the final model controls.
- Added the no-tagging and alternative-movement appendix and the interactive-resource index; aligned the concise interpretation, captions, values and cross-references with the revised evidence.

Revision 2 (1 October 2026)

- Corrected and replaced the regional and overall fishery-impact plots (Figure 66).

Table of contents

1	Executive Summary	6
1.1	Open and reproducible assessment resources	9
2	Introduction	10
3	Background	12
3.1	Stock Structure	12
3.2	Biological characteristics	13
3.3	Fisheries	14
4	Data compilation	15
4.1	General notes	15
4.2	Spatial stratification	15
4.3	Temporal stratification	16
4.4	Definition of fisheries	16
4.5	Catch and effort data	18
4.5.1	General characteristics	18
4.5.2	Purse seine	20
4.5.3	Longline	20
4.5.4	Other fisheries	21
4.6	Size data	21
4.7	Tagging data	22
5	Model description	23
5.1	General characteristics	23
5.2	Population dynamics	23
5.2.1	Recruitment	24
5.2.2	Initial population	24
5.2.3	Growth	25
5.2.4	Movement	25
5.2.5	Natural mortality	26
5.2.6	Reproductive potential	26
5.3	Fishery dynamics	27
5.3.1	Selectivity	27
5.4	Dynamics of tagged fish	27
5.4.1	Tag reporting	27
5.4.2	Tag mixing	28
5.5	Likelihood components	30
5.5.1	Index fishery CPUE likelihood	31
5.5.2	Length frequency	31
5.5.3	Tag data	31
5.5.4	Conditional age-at-length data (CAAL)	33
5.5.5	Parameter estimation and uncertainty	33
5.6	Diagnostic methods	34
5.6.1	Convergence	34

5.7	Other diagnostics	34
5.8	Stock assessment interpretation methods	36
5.8.1	Depletion and fishery impact	36
5.8.2	Reference points	37
5.8.3	Yield analysis	37
5.8.4	Kobe analysis and Majuro plots	38
5.8.5	Stock projections from the structural uncertainty grid	38
6	Model runs	38
6.1	Developments from the last assessment	38
6.1.1	Stepwise model development	38
6.2	Sensitivity analyses and uncertainty ensemble	40
6.2.1	Sensitivities	40
6.2.2	Uncertainty characterisation	41
7	Results	42
7.1	Fit of the diagnostic model to data sources	45
7.1.1	Standardized CPUE: index fisheries	45
7.1.2	Size composition data	45
7.1.3	Tagging data	46
7.1.4	Tag reporting rates	46
7.1.5	Conditional age-at-length	46
7.2	Diagnostic analysis for the diagnostic model	47
7.2.1	Convergence	47
7.2.2	Jitter	47
7.2.3	Self test	47
7.2.4	Retrospectives	48
7.2.5	Likelihood profiles	48
7.2.6	ASPM	48
7.2.7	No-tagging and alternative-movement comparisons	49
7.3	Population dynamics estimates	49
7.3.1	Selectivity	49
7.3.2	Movement	49
7.3.3	Natural mortality	49
7.3.4	Maturity	50
7.3.5	Growth	50
7.3.6	Recruitment: diagnostic model	50
7.3.7	Spawning potential and depletion: diagnostic model	50
7.3.8	Fished (SB) versus unfished ($SB_{F=0}$) spawning potential: diagnostic model	50
7.3.9	Fishing mortality: diagnostic model	51
7.4	Sensitivity analyses	51
7.4.1	Steepness	51
7.4.2	Natural mortality	51
7.4.3	CAAL data weighting	51

7.4.4	Effort creep on CPUE	52
7.4.5	CPUE time period for regional scaling	52
7.4.6	Tag mixing	52
7.4.7	Reporting rates adjustment in the mixing periods	52
7.4.8	Tau overdispersion parameter	52
7.5	Stock assessment results	52
7.5.1	Uncertainty ensemble	52
7.5.2	Analyses of stock status	53
7.5.3	Majuro and Kobe plots and comparisons with reference points	53
7.5.4	Fishing impact and time-dynamic MSY	54
7.5.5	Yield analysis	54
7.5.6	Catch-based projections	54
8	Discussion and conclusions	55
8.1	Stock Status from the uncertainty grid	55
8.2	Diagnostic interpretation	55
8.3	Recommendations for further work	57
9	Acknowledgements	58
10	Statement on the use of generative AI	58
11	Tables	59
12	Figures	68
A	Appendices	136
A.1	No-tagging and alternative-movement model comparisons	136
A.1.1	Conclusions	141
	References	141

1 Executive Summary

This paper describes the 2026 stock assessment of bigeye tuna *Thunnus obesus* in the western and central Pacific Ocean. An additional three years of data were available since the previous assessment in 2023, and the model extends through to the end of 2024. The assessment applies a simplified 5 region model structure compared to the 9 region structure used in the previous four bigeye assessments; 2017, 2018, 2020, and 2023. The spatial structure represents a step in simplifying this assessment considering comments from SC19 and the 2026 Pre-Assessment Workshop (PAW) (Hamer, 2026). Readers are also referred to SC22-SA-IP09 (Hampton et al., 2026b) which includes exploration of a simpler single region model. New developments to the stock assessment, many of which have been in response to comments on the 2023 assessment by SC19 and discussion at the PAW, include:

- Simplified spatial structure from 9 to 5 regions, and revised fisheries definitions.
- Conversion of longline weight data to lengths outside the model and addition of observed length data for longline fisheries (Peatman et al., 2026b)
- Stricter size composition data filtering and enhanced reweighting approach (Peatman et al., 2026b)
- Global sdmTMB CPUE analysis used for regional scaling, and new MFCL regional scaling penalty (Davies et al., 2026a; Teears, 2026)
- Separate sdmTMB CPUE analyses for each model region (x 10), removing the need for shared selectivity and catchability among CPUE indices (PAW and yellowfin tuna peer review recommendation) (Hamer, 2026; Punt et al., 2023; Teears, 2026)
- External M-scale assumption rather than internal estimation (PAW recommendation to improve model stability)
- Release-group-specific tag mixing periods based on an individual-based tag-mixing model (consistent with the yellowfin tuna peer-review recommendation in Punt et al. (2023)), and removal of Coral Sea tagging data (supported by PAW)
- Tag reporting rate penalties informed by increased tag seeding (Peatman et al., 2026a)
- Substantial increase in age at length data, including north Pacific samples (Hasegawa et al., 2026; Wakefield et al., 2026)
- Uncertainty ensemble rather than orthogonal grid
- Enhanced diagnostics

This assessment is supported by the analysis of catch and effort data for longline fisheries to provide regional abundance indices (Teears, 2026), revised analysis of tagger effects and tag reporting rates (Peatman et al., 2026c; Peatman et al., 2026a), enhanced and improved size composition data analyses and preparation, include the weight data to length conversion work (Peatman et al., 2026b), tag mixing simulation modeling (Scutt Phillips et al., 2026) and

developments to MFCL software (Davies et al., 2026a). An extensive process of stepwise model development was conducted to transition from the 2023 to the 2026 assessment model and new data.

The 2026 bigeye stock assessment estimated the median $SB_{recent}/SB_{F=0}$ across the model grid to be 0.277, where *recent* is the period 2021-2024. Overall, the current assessment estimates indicate a more pessimistic stock status, or lower $SB/SB_{F=0}$, than the 2023 assessment.

In addition to the diagnostic model, we report the results of one-off sensitivities to explore the impact of important data and model assumptions for the diagnostic model on the stock assessment results and conclusions. These sensitivity analyses informed the structure of a model ensemble used to characterise the uncertainty of key management quantities for considering the management advice. The most important uncertainties related to assumptions required for the modeling of the tag/recapture data. This was because there is conflict between the tag data and the other key data (CPUE abundance indices and size data) with regards to the population size of the bigeye stock in the model region. The model interprets the tag data as being consistent with a larger stock size and lower impact of fishing resulting in a more optimistic/less depleted status, whereas it interprets the CPUE abundance indices and the size data as more consistent with a lower stock size, higher impact of fishing and a more pessimistic/more depleted status. The sensitivity analysis has shown that models with greater emphasis on either of these data sources can produce acceptable diagnostics and similar trends and dynamics, but different outcomes for the estimates of the key management quantities due to their conflicting information on population scaling. Further discussion on this topic is provided below.

Across the 80 retained models, the natural-mortality scale and tag-mixing periods were the most influential axes for the spread in stock status. Stepwise development and one-off sensitivities also showed a strong response to tag-data overdispersion for τ values from 1.0 to approximately 1.4, above which estimates of depletion and population scale changed little. The global-CPUE period used to inform the regional-scaling penalty was another influential assumption.

The general conclusions of this assessment are as follows:

General trends

- The spawning potential of the stock has become more depleted across all model regions until around 2010, after which it has become more stable.
- Average fishing mortality rates for juvenile and adult age-classes have increased throughout the period of the assessment until around 2000, after which they have stabilised, but with high inter-annual variability for juveniles.
- Juveniles have experienced considerably higher fishing mortality than adults.
- The diagnostic model in this assessment was one of the most pessimistic models of the uncertainty ensemble used for management advice.

Stock status management quantities

- Median recent depletion (2021–2024; $SB_{recent}/SB_{F=0}$) is estimated at 0.277 (central 80% interval including structural and available estimation uncertainty: 0.172–0.395; central 95% interval: 0.136–0.442).
- Fifteen of the 80 models in the ensemble had $SB_{recent}/SB_{F=0} < \text{LRP } 0.2$, the uncertainty ensemble with estimation uncertainty estimates 18.6% probability on being below the LRP.
- CMM 2025-02 contains an objective to maintain the spawning biomass depletion ratio at or above the average $SB/SB_{F=0}$ for 2012–2015. Across the 80 ensemble models, excluding estimation uncertainty, this average was 28.8%. The corresponding average for 2021–2024 was 27.8%, giving a recent-to-2012–2015 spawning depletion ratio of 96.7%; stock status was therefore slightly below the objective (Table 9).
- Median recent fishing mortality (2020–2023), F_{recent}/F_{MSY} , was 0.838 (central 80% interval including structural and available estimation uncertainty: 0.592–1.206; central 95% interval: 0.460–1.351).
- Although the stock is assessed as more depleted and subject to higher fishing mortality than in 2023, the ensemble indicates that it is not overfished under the adopted LRP criterion (18.6% probability below the LRP) and is not undergoing overfishing based on the median $F_{recent}/F_{MSY} < 1$.
- Overall, the 2026 assessment presents a more pessimistic stock status than the previous assessment.

The conflict between tag data and other data on the model estimation was considered when constructing the uncertainty ensemble. Sensitivity analyses indicated the influence of tag data could be moderated through altering the value of a parameter referred to as tau (τ). In simple terms the parameter τ measures the degree of "overdispersion" in the tag recapture data. In other words, how much more variable are the recapture observations than the model would expect if all observed recaptures behaved identically, i.e. homogeneity in the processes influencing tag recaptures and reporting rates. A τ value set at 1 assumes this homogeneity. Larger values assume additional, unmodeled variability, or heterogeneity in the tag recapture and reporting processes, e.g. due to differences among individual releases, years, areas, reporting fisheries etc.. Some of this may be moderated by the adjustment to tag releases numbers (i.e. tagger effects adjustments), but most of these influences the model cannot capture. A τ value set at 1 is not considered defensible for tag/recapture data. Higher levels admit more uncertainty to the tag/recapture data allowing the model greater flexibility to fit the other data. We found setting τ at 1.4 or higher achieved the same outcomes in terms of management quantities, including the value of 2 used in the diagnostic model. There is no clear guidance on the τ setting, except that it should not be close to 1 for tag/recapture data, and other assessments have pointed to a value of 2, but with limited details. Other factors in the ensemble also vary the tag data influence but the τ overdispersion parameter appeared most important. Given the uncertainty over the value of τ and its influence on outcomes, and in order to allow the tag data to maintain an

influence in the uncertainty ensemble considering the conflict, we include models with τ of 1.2, 1.3, 1.4, noting 1.4 appears to heavily downweights the influence of tag data on the management quantities. A key area for further research is highlighted by the influence of tag assumptions on this bigeye model (see below).

The assessment has implemented important improvements in key area such as the simpler spatial structure, size data preparation and conversion to length data, increased age at length data, release group specific mixing periods, and CPUE methods, and driven a deeper focus on the modeling of the tag data and related assumptions. The latter area provided the most problematic and challenging aspect of this assessment. Population scaling is a key uncertainty in most assessments, and particularly tuna assessments that rely heavily on fishery dependent data. Tagging data are an important data type that can provide information on population scaling, movement, mortality and relative abundance. Hence it is a key data type used in several tuna assessments across the Pacific, and other important tuna research applications. With the challenges related to understanding the tag data conflict and the methods and assumptions for modeling these data within an integrated assessment framework, we recommend:

- An analysis of the tag data outside of the MFCL model to better understand the variability and statistical properties (i.e. overdispersion) of these data.
- A focused technical review of the approaches for use of tag data to inform stock assessment models, both as external and internal data.
- Support the continued development of the external skipjack tag data analysis proposed by Project 123 (SA-WP01), and consider adding a feasibility analysis for bigeye and yellowfin tagging data.
- The SC continue to explore options for alternative scheduling of benchmark tuna assessments to allow more time for their development and related data preparation and research work.

1.1 Open and reproducible assessment resources

The assessment results and supporting analyses can be opened directly:

- **Assessment:** [clean report](#).
- **Model development and fit:** [stepwise viewer](#), [diagnostic report](#) and [likelihood-profile viewer](#).
- **Model diagnostics and sensitivities:** [jitter](#), [self-test](#), [retrospective](#) and [one-off sensitivity viewer](#).
- **Management uncertainty and projections:** [ensemble viewer](#).

The corresponding repositories and pinned build environment are documented once in Table 1. These linked, versioned resources expose detail that cannot be shown legibly in static figures

and support transparent, reproducible exploration beyond that available for previous WCPO tuna assessments.

2 Introduction

This paper presents the 2026 stock assessment of bigeye tuna (*Thunnus obesus*; BET) in the western and central Pacific Ocean (WCPO; west of 150° W). Assessments of WCPO bigeye tuna have been conducted regularly since the late 1990s (Davies et al., 2011; Day et al., 2023; Ducharme-Barth et al., 2020a; S. Harley et al., 2009; Harley et al., 2010; Harley et al., 2014; Langley et al., 2008; McKechnie et al., 2017a; Vincent et al., 2018). As in previous assessments, the objectives of the 2026 bigeye tuna assessment are to estimate population level parameters which indicate the stock status and impacts of fishing, such as time series of recruitment, biomass, biomass depletion and fishing mortality. We summarize the stock status in terms of reference points adopted by the Western and Central Pacific Fisheries Commission (WCPFC). The methodology used for the assessment is based on the general approach of integrated modeling (D. Fournier & Archibald, 1982), which is carried out using the stock assessment framework MULTIFAN-CL (version 2.2.7.9)⁵ Fournier et al. (1998); Hampton & Fournier (2001); Davies et al. (2026b)). MFCL implements a size-based, age- and spatially-structured population model. Model parameters are estimated by maximizing an objective function, consisting of both likelihood (data) and “prior”⁶ information components (penalties).

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

The assessment uses an ‘uncertainty ensemble’ of models as the basis for management advice. The uncertainty ensemble is a group of models that are run to capture uncertainty in stock status estimation that relates to uncertainty in biological and structural assumptions, data inputs and data treatment. The uncertainty models are generally based on one-off sensitivity models of a diagnostic (or base case) model to indicate which uncertainties have notable effects on the

⁵<https://github.com/PacificCommunity/multifan-cl/>

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

estimates of key model parameters and management quantities. The variation in estimates of the management quantities across the uncertainty ensemble represents the uncertainty in stock status and should be considered carefully by managers. Structural or “model” uncertainty is usually more important than the uncertainty in the estimation of the parameters from individual models, referred to as “estimation uncertainty”, however both are taken into account when documenting the uncertainty in management quantities provided by the assessment.

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

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

- Simplified spatial structure from nine to five regions and revised fishery definitions (Figure 1)
- Conversion of weight data to lengths outside the model and addition of length data for longline fisheries (Peatman et al., 2026b)
- Stricter size composition data filtering and enhanced reweighting approach (Peatman et al., 2026b)
- Global sdmTMB CPUE analysis used for regional scaling, with new MFCL regional scaling penalty (Davies et al., 2026a; Teears, 2026)
- Separate sdmTMB CPUE analyses for each model region, removes need for shared selectivity and catchability among CPUE indices (PAW and yellowfin tuna peer review recommendation) (Hamer, 2026; Punt et al., 2023; Teears, 2026)
- External M-scale assumption rather than internal estimation (PAW recommendation to improve model stability)
- Release-group-specific tag mixing periods based on an individual-based tag-mixing model (consistent with the yellowfin tuna peer-review recommendation in Punt et al. (2023)), and removal of Coral Sea tagging data
- Tag reporting rate penalties informed by increased tag seeding (Peatman et al., 2026a)

- Substantial increase in age at length data, including north Pacific samples (Hasegawa et al., 2026; Wakefield et al., 2026)
- Uncertainty ensemble rather than orthogonal grid
- Enhanced diagnostics

3 Background

3.1 Stock Structure

Bigeye tuna are distributed throughout the tropical and sub-tropical waters of the Pacific, Indian and Atlantic Oceans. Genetic studies indicate that Atlantic bigeye have minimal mixing with bigeye in the Indo-Pacific regions (Chow et al., 2000; Gonzalez et al., 2008; Grewe et al., 2019). There is no strong evidence for genetic population structure in the Pacific Ocean (Grewe & Hampton, 1998; Moore et al., 2020; Natasha et al., 2022). While genetic studies are largely uninformative on the rates of mixing of bigeye tuna throughout the Pacific, they are broadly consistent with the results of tagging experiments conducted by SPC and the Inter-American Tropical Tuna Commission (IATTC) (Moore et al., 2020; K. Schaefer et al., 2015). These studies show that while the majority of tagged bigeye have been recaptured in the general regions of their release (i.e. within 1,500 nautical miles) some tagged bigeye move large distances across the Pacific Ocean (i.e. >4,000 nautical miles) Figure 11. These occasional large-scale movements and the continuous distribution of bigeye tuna across the Pacific would explain the lack of clear genetic structure. The cumulative information on movement from conventional and electronic tag-recapture data, particularly efforts focused in the equatorial central Pacific, have indicated more extensive longitudinal, particularly west to east, displacements with fewer movements between tropical tag release sites and temperate zones (McKechnie et al., 2017a; Moore et al., 2020; K. Schaefer et al., 2015). While there is apparent high movement throughout the equatorial Pacific and lack of genetic structure, this does not preclude spatial structure in population processes that are important in stock assessment (i.e., growth, mortality, recruitment) and regional fishing impacts/local depletion. Growth rates have been shown to differ between bigeye tuna in the WCPO and EPO, with smaller lengths-at-age observed for bigeye in the WCPO, and suggestion of a clinal increase in growth rates from the west to east Pacific (Aires-da-Silva et al., 2015; Farley et al., 2017; McKechnie et al., 2015a). This is also reflected in observations of increased size at maturity for bigeye in the EPO compared to the WCPO (Farley et al., 2017). Further, there is some evidence from otolith chemistry studies for structuring of recruitment sources. For example, Rooker et al. (2016) indicated that 1–2 year old fish sampled from the Marshall Islands were derived from local sources, whereas fish caught off Hawai'i were suggested to have originated more broadly from the central equatorial region. In turn, otolith chemistry and genetic studies of juvenile bigeye in the Indonesian archipelagic waters suggested juvenile residency and low mixing with the broader western Pacific for this life stage (Moore et al., 2019; Proctor et al., 2019).

The paper by Hamer et al. (2023) considers published information on genetic and non-genetics

indicators of population structure, plus larval distribution patterns, and also includes analyses of spatial heterogeneities in size composition data and CPUE time series for the Pacific longline fisheries. That paper suggests bigeye are likely one genetic stock in the Pacific, and that there is no clear boundary between the WCPO and EPO populations in the central Pacific region. However, there is indication of substructure of bigeye across the Pacific, and that the populations across the west and east Pacific, and the north and south Pacific are unlikely to be extensively mixed. The paper notes that the tropical region has several features that warrant it being considered as a separate spatial strata from northern and southern sub-tropical/temperate regions. Likewise the paper suggests that the Indonesia/Philippines/Vietnam region warrants being considered a separate spatial strata. The available information did not clearly imply a finer structure. The review recommended alternative simplified spatial stratification that could be considered for bigeye tuna assessments in the WCPO and this is discussed further in Section 4.2.

Despite uncertainties in population connectivity across the Pacific a sensitivity analysis by McKechnie et al. (2015a) in 2015 showed that despite extensive longitudinal movements in the equatorial Pacific, the discrete WCPO stock assumption was capable of accurately capturing the dynamics and stock status indicators for WCPO bigeye. The current assessment continues the approach of a separate assessment for the WCPO region and the stock within the domain of the model (essentially the WCPO, west of 150° W) is considered to be a discrete stock unit (Davies et al., 2011; Day et al., 2023; Ducharme-Barth et al., 2020a; S. Harley et al., 2009; Harley et al., 2010; Harley et al., 2014; Langley et al., 2008; McKechnie et al., 2017a; Vincent et al., 2018).

3.2 Biological characteristics

Bigeye tuna are moderately fast growing, and have a maximum fork length (FL) of about 200 cm (Aires-da-Silva et al., 2015; Farley et al., 2017), with an estimated average maximum length of 157 cm in the WCPO (Eveson et al., 2020; Farley et al., 2020). The growth parameters for bigeye tuna are influential biological parameters in the stock assessment (McKechnie et al., 2017a; Vincent et al., 2018). Studies to improve growth information for bigeye have occurred in recent years, notably by Eveson et al. (2020); Farley et al. (2020); Wakefield et al. (2026); Hasegawa et al. (2026). Recent bomb radiocarbon age work has validated the interpretation of growth bands in bigeye otoliths as annual and confirmed the oldest ‘validated’ age for a bigeye tuna in the WCPO at least 13 years of age. This compares to the longest period at liberty for a recaptured bigeye tuna tagged in the WCPO of approximately 14 years, for a fish estimated at 1-2 years old at release (SPC, unpublished data). The oldest estimated fish aged from otoliths in the WCPO is approximately 15 years (Farley et al., 2020). The 2026 assessment will incorporate a substantial increase in the amount of conditional-age-at-length data, both for the equatorial region and the region north of 10° N for which no samples were available in the previous assessments (Hasegawa et al., 2026; Wakefield et al., 2026).

Bigeye tuna in the WCPO become reproductively active from about 80 cm FL, and nearly all individuals >120 cm FL are mature (Farley et al., 2017). There is some evidence for regional variation in maturity-at-length in the WCPO and EPO (Farley et al., 2017; Nicol et al., 2011; K. M. Schaefer et al., 2005). As with other tunas, the population sex ratio of bigeye tuna changes

at around the age/size of reproductive maturity to increasingly favor males at larger sizes, thought to relate to higher mortality of females due to the energetic stress of regular spawning (Sun et al. (2006); McKechnie et al. (2017b)). Sex ratio information is used to define spawning potential (rather than spawning biomass) as a product of maturity status, female sex-ratio in the population, and fecundity. Spawning typically occurs in water temperatures above 24° C, with an approximate daily cycle, and occurs all year in the tropical Pacific, but more seasonally in the warmer months in sub-tropical regions (Farley et al., 2017; Sun et al., 2006). Areas of higher larval densities have been found in the western Pacific from Papua New Guinea to Philippines, Chinese Taipei to the Okinawa Islands, also several areas in the central Pacific, and in the eastern Pacific off Panama and Colombia (Ijima & Jusup, 2023; Nishikawa et al., 1985).

The natural mortality (M) rate of bigeye tuna varies with size, with rates of $< 0.5 \text{ yr}^{-1}$ for bigeye $>40 \text{ cm FL}$ and >2 years age and higher rates for the youngest age-classes (<2 years) (Hampton, 2000). The 2020 bigeye assessment conducted a life-history based meta-analysis (Piner & Lee, 2011) of natural mortality for bigeye and indicated an envelope of potential average quarterly M rates of lower 95% confidence interval (0.109), mean (0.127) and upper 95% confidence interval (0.146) (Ducharme-Barth et al., 2020a). The 2023 assessment estimated M internally to the MFCL model using a Lorenzen function, adult (age 40 quarters) M was estimated at 0.078 per quarter (Day et al., 2023).

3.3 Fisheries

Bigeye tuna are an important component of tuna fisheries throughout the Pacific Ocean. Larger bigeye are targeted by longlines, predominantly in tropical regions of the WCPO. Smaller juveniles are caught incidentally by purse seine vessels in association with Fish Aggregation Devices (FADs). They are a principal target species in tropical waters for both the large, distant-water longline fleets of Japan, Korea, China and Chinese Taipei and the smaller, fresh sashimi longline fleets based in several Pacific Island countries and Hawai'i. The prices paid for both frozen and fresh product on the Japanese sashimi market are generally the highest of all the tropical tunas.

While they are not targeted by purse seine vessels, bigeye are taken almost exclusively from sets on natural and artificial floating objects (FADs), and comprise mostly small juveniles. Estimation of the bigeye tuna catch from associated sets has been the focus of considerable research over recent years (Peatman et al., 2019). This purse seine fishery expanded rapidly from the early 1980's and the estimated annual bigeye catches for this gear have been in the vicinity of 45,000-72,000 mt over the last 15 years in the WCPO (i.e., the region of this assessment which excludes the overlap area with IATTC) (McKechnie et al., 2026). This compares to the WCPO longline catch of between approximately 32,000-59,000 mt over the same period (Hare et al., 2025; McKechnie et al., 2026).

The highest estimated bigeye tuna total catch from the WCPO was approximately 175,000 mt in 2004 (195,000 mt in the same year for the entire WCPFC-CA). The WCPO bigeye catches have declined since around 2010, after which they have ranged between around 110,000–160,000

mt (135,000-175,000 mt for entire WCPFC-CA) (Hare et al., 2025, 2026; McKechnie et al., 2026; Williams & Ruaia, 2023). Over the recent period (2015-2024), approximately 50% of the catch by weight, but 10% by numbers was taken by longline, reflecting the selection of longline gear for larger fish. In contrast, 50% of the catch by numbers, but only 35% by weight was taken by purse seine sets associated with FADs. A small purse seine fishery also operates in the coastal waters off Japan with an annual bigeye catch of less than 500 mt in recent years. A slightly higher level of bigeye catch is taken by the coastal Japanese pole-and-line fishery relative to the coastal Japanese purse seine, however, overall these catches are negligible.

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

4 Data compilation

4.1 General notes

Data used in the bigeye tuna stock assessment using MFCL consist of catch, effort, length-frequency data for the fisheries defined in the analysis, and tag-recapture data. The fitted length compositions combine observed lengths with longline weight observations converted externally to length. Conditional age-at-length data are also used directly as data in the assessment model, as was recommended by the peer review of the 2020 yellowfin tuna assessment (Punt et al., 2023). Improvements in these data inputs are ongoing and readers should refer to the companion papers highlighted at the end of Section 2 for detailed descriptions of how the data and biological inputs were formulated as only brief overviews are provided below. A summary of the data available for the assessment is provided in (Tears, 2026), and in Figure 6.

4.2 Spatial stratification

The geographical area considered in the assessment corresponds to the WCPO (from 50°N to 20°S between 120°E and 150°W) and oceanic waters adjacent to the east Asian coast (110°E between 20°N and 20°S). The eastern boundary of the assessment excludes the WCPFC Convention area component that overlaps with the Inter American Tropical Tuna Commission (IATTC) area. The assessment area is stratified into 5 regions Figure 1, which is a simplification of the 9 region stratification used in the the previous four assessments. The simplified spatial structure was motivated by a general desire to reduce the complexity following SC reviews of the 2023 assessment and discussion at PAW, particularly by removing the small region 9 that was implemented in 2014, after the bigeye assessment peer review (Ianelli et al., 2012). This small region was included to better model the tagging data from the historical Coral Sea tagging cruises (Harley et al., 2014). Region 9 was maintained in the previous assessments, partly

because of tag-recaptures with long times at liberty that were thought to provide information on natural mortality (M) which was estimated internal to the model in the 2023 assessment. In the 2026 assessment, in an attempt to improve model stability, after discussion at the PAW it was decided to return to the approach of earlier assessments and apply M as an external parameter (Hamer, 2026). Related to this change the archival tag-recapture data from the Coral Sea Tagging program combined with conventional tags from the Regional Tuna Tagging Programme (RTTP) and the Pacific Tuna Tagging Programme (PTTP) was used in an external analysis to estimate an adult M and M -at-age (using previous growth curves) Lorenzen curve (Ducharme-Barth et al., 2026). The smaller Region 8 around Papua New Guinea and the Solomon Is was used in previous assessment to better model tag-recapture data in relation to assumed tag mixing. However, this was mostly related to skipjack and yellowfin that have high numbers or tag releases in that region, there are few bigeye tag releases in this region so it was deemed unnecessary to maintain this region for bigeye. This reduced the stratification to 7 regions. The negligible tag data to inform movement across the northern (old regions 1, 3) as southern (old regions 5, 6) regions, and no other strong justification for multiple regions to the north and south of the equatorial region, so they were each reduced to single regions, retaining some spatial fleet separation by using a fleets-as-areas approach. This resulted a 4 region stratification that was supported at the PAW, and was more consistent with the conceptual model and the objective of removing unnecessary spatial complexity (Hamer et al., 2023; Hamer, 2026). An alternative 5 region structure was also proposed at PAW with a separate central-east equatorial region to better accommodate the concentration of tag-recapture data in that area and to improve the level of within model region tag mixing. Ultimately based on the tag mixing modeling (Scutt Phillips et al., 2026), the 5 region structure was chosen for the 2026 assessment Figure 1. The PAW also recommended to explore a single region fleets-as-area model, excluding all tag data (Hamer, 2026). This exploration is provided in a companion paper to SC22 (Hampton et al., 2026b).

The correspondence between the former nine-region structure and the five regions used in 2026 is shown explicitly in Figure 2.

4.3 Temporal stratification

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

4.4 Definition of fisheries

MFCL requires “fisheries” to be defined that consist of relatively homogeneous fishing units. Ideally, the defined fisheries will have selectivity and catchability characteristics that do not vary greatly over time and space. For most pelagic fisheries assessments, fisheries are typically defined

according to combinations of gear type, fishing method and region, and for some, also flag or fleet. The simplified spatial structure of the 2026 assessment necessarily required a revision of the fisheries definitions. This revision resulted in a reduction of the total number of defined fisheries from 41 in the 2023 assessment to 33 in the current assessment (Table 2). There are two fishery types: “index fisheries”, that are used for generating indices of abundance for each model region (five index fisheries, see further below), and “extraction fisheries” that account for the catches removed from the stock (28 extraction fisheries). Extraction fisheries include longline, purse seine, pole and line, handline and various miscellaneous fisheries in Indonesia/Philippines/Vietnam region. Availability of data for each fishery used in the assessment model is provided in Figure 6.

For the 2026 assessment the large fish handline fisheries have been separated for Indonesia and Philippines to allow separate treatment of the data from these two fisheries which have different levels of catch and size composition uncertainty. Equatorial purse seine fishing activity (i.e. region 3 and 4) was aggregated over all nationalities, but stratified by region and set type, in order to sufficiently capture the variability in fishing operations and selectivity of different purse seine set types. Set types were grouped into associated (i.e. log, FAD, whale, dolphin, and unknown set types) and unassociated (free-school) sets. Additional fisheries were defined for pole-and-line fisheries and miscellaneous fisheries (gillnets, ringnets, hook-and-line, handlines etc.) in the western equatorial area/Indonesia/Philippines/Vietnam. At least one longline index fishery was defined in each region, although in regions 2, 3, 4 extraction longline fishing was separated into distant water and offshore components to account for the apparent differences in fishing practices and selectivity for these fleets in these regions.

Longline index fisheries: The catch-conditioned approach allows the specification of “index fisheries” that are used to provide standardised CPUE indices of abundance. Index fisheries are akin to “survey fisheries” as described for other software such as Stock Synthesis, and may be the same fisheries as the extraction fisheries, but when used as index fisheries they do not take any catch, and must have effort data to allow modeling of CPUE. For this assessment one index fishery is defined for each model region as a composite fishery composed of longline fisheries operating under various flags in each assessment region (Tears, 2026). In the 2023 assessment a single sdmTMB CPUE standardisation model was conducted across the entire assessment region (i.e. global model), the quarterly predictions were then aggregated and summed by model regions to generate the region specific abundance indices (Tears et al., 2023), that were assumed to have constant catchability and selectivity. This allowed MFCL to use the differences in relative abundance indices among model regions as information on the relative size of the vulnerable population among model regions (i.e. regional scaling). The assumption of shared catchability and selectivity is problematic and can create problems with adequately fitting the size compositions for each index fishery. The yellowfin assessment peer review recommended freeing up these assumptions of shared catchability and selectivity among CPUE abundance indices (Punt et al., 2023). This was discussed in depth at the PAW, where it was recommended to run a global sdmTMB model using the long-term Japanese operational data, use this global model for generating regional abundance indices that are assumed to have shared catchability and selectivity, but only use these for information on relative vulnerable abundance among

model regions (i.e. only for regional scaling). For the region specific abundance indices, conduct individual sdmTMB CPUE models for each model region. Separate regional CPUE models could perhaps better account for regional differences in CPUE covariate relationships, and the resultant abundance indices can have their own selectivity without assuming constant catchability among model regions (Hamer, 2026). This approach was adopted and required considerable additional work in the CPUE analyses (Teears, 2026), and also modifications/enhancement to the MFCL code to allow one set of CPUE (i.e. global model) to be used to generate a regional abundance scaling penalty, and the other set of region specific CPUE indices to provide the time series abundance indices for the model to fit to (Davies et al., 2026a). Furthermore, specification of the time period from the global indices to use to calculate the regional scale penalty is required. The PAW recommended using a time period where data coverage was most broad across the global model domain, this period was 1965-1969 (Teears, 2026). The MFCL code was also modified to input global CPUE time series of different start and end dates so the scaling penalty could be determined for different time periods, noting the the penalty is applied only to that time period Section 4.5.3.

The regular longline extraction fisheries are based on the same data set, but are disaggregated into the longline fisheries defined in Table 2. The length-composition data for the extraction fisheries are assumed to represent the actual composition of the removed fish for any space-time strata, and in the data preparation process are weighted by the catch in order to represent the fisheries extractions at the spatial (region) and temporal (quarter) resolution of the model (Peatman et al., 2026b). However, for the index fisheries, while the same aggregation process is conducted, the size data are weighted by standardised CPUE (rather than by catch) so that the size data are more representative of the abundance of the underlying population in each region and time period. Further, because the size data for the index and extraction fisheries are effectively being used twice (but weighted differently), the likelihood weighting for the size composition is adjusted such that the original intended weight (effective sample size) in the likelihood is preserved.

4.5 Catch and effort data

4.5.1 General characteristics

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

In the catch conditioned model effort is not essential but is required (at least for a recent period of time) for projection analyses involving fisheries managed under effort rather than catch controls.

The effort data are necessary to derive recent estimates of catchability for running the effort based projections. In this case the main industrial purse seine fisheries operating in the tropical region (i.e., regions 3, 4) are managed under effort control. Effort data for these purse seine fisheries are defined as number of sets specified by set type (associated or unassociated), and are included for the last 12 quarters to facilitate projections. The period of 12 quarters is consistent with previous projections using catch errors models. For this assessment several other fisheries also have effort included to allow effort based projection for management purposes, these are the longline extraction fisheries, with effort measured as numbers of hooks per set, and the Japanese pole and line fisheries with effort measured in vessel fishing days.

Total annual catches by major gear categories for the WCPO are shown in Figure 3 and a regional breakdown is provided in Figure 4. Catches by fishery groups are provided in Figure 7, Figure 8 and Figure 9. The spatial distribution of catches over the past ten years is provided in Figure 5. Discarded catches are estimated to be minor and were not included in the analysis. Catches in the northern region are highly seasonal and the annual catch has been relatively stable over much of the assessment period. Most of the catch occurs in the tropical regions (2, 3 and 4).

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

- The steady increases in total bigeye catch over most of the assessment period in the equatorial regions (regions 2, 3, 4), and steady declines since around 2010, with relatively stable/slightly increasing catches over the last 5 years.
- The relatively stable catches of bigeye in the northern temperate region by longline vessels (regions 1).
- The development of the equatorial purse-seine fisheries from the mid-1970s and the widespread use of FADs since the mid-1990s, allowing an expansion of the purse-seine fishery, and corresponding increases in catch of bigeye, particularly in equatorial regions.
- Large changes in the purse seine fleet composition and increasing size and likely efficiency of the fleet.
- The steady increase in catch for the domestic fisheries of Indonesia and the Philippines (region 2) since 1970.
- The apparent stabilisation of catches of bigeye for most gears after the mid 2000's, and declines in overall catches since around 2010.
- The notable increase in handline catch reported for 2024. This resulted from a change in how Indonesia provided its 2026 catch statistics and allocated bigeye catch between small- and large-fish handline fisheries. Indonesia determined that a large portion of the catch previously reported as smaller-fish handline gear (part of the 'other' category) should be allocated to the larger-fish handline category (see Vidal et al. (2026)). A historical adjustment has not yet been made and is a priority under the WPEA project. The terminal-year increase is expected to have only a minor influence on the current assessment, but

the historical series should be corrected before the next assessment.

4.5.2 Purse seine

For the industrial purse seine fisheries predominantly operating in tropical regions 3, and 4, catch by species within each set type (associated or unassociated) is determined by applying estimates of species composition from observer-collected samples to total catches estimated from raised logsheet data (Hampton & Williams, 2016; Peatman et al., 2021; Peatman et al., 2023). For the Japanese (JP) fleet for which there is greater confidence in species-based reporting, reported catch by species is used. Purse seine catch for Philippines (PH) and Indonesian (ID) domestic purse seine fisheries, predominantly operating in region 2, was derived from raised port sampling data provided by these countries. We note that the COVID-19 pandemic resulted in low observer coverage of the purse seine fleets from mid 2020 to 2022 and purse seine catch estimates are less reliable over these years.

4.5.3 Longline

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

The index fishery CPUE time series for the 2026 assessment were derived from the operational longline dataset for the Pacific region Figure 10. This dataset is an amalgamation of operational level data from the distant-water fishing nations (DWFN), United States, Australian, New Zealand and Pacific-Island countries and territories (PICTs) longline fleets operating in the Pacific basin. It represents the most complete spatiotemporal record of longline fishing activity in the Pacific, spanning from 1952 through to the present and is the result of collaborative ongoing data-sharing efforts from many countries. This data-set was first created in 2015 in support of the Pacific-wide bigeye tuna stock assessment (McKechnie et al., 2015b), and has subsequently been used in each bigeye and yellowfin tuna assessment (Day et al., 2023; Ducharme-Barth et al., 2020a; McKechnie et al., 2017a; Vincent et al., 2018). Since 2017 spatiotemporal approaches (Thorson et al., 2015) have been used for CPUE modeling in WCPFC stock assessments (Ducharme-Barth et al., 2020b; Teears et al., 2023; Tremblay-Boyer et al., 2017). Consistent with the 2023 assessment, the CPUE modeling was conducted with sdmTMB (Anderson et al., 2022). A detailed description of the methods for generating the spatiotemporal abundance indices is provided in Teears (2026). However, some aspects are worth noting here. Firstly, the global analysis used for the regional scaling penalty takes account of habitat viability by not including predictions in cells when SST is below a specified thresholds. The threshold was determined as 17°C based on the probability of positive CPUE observations being greater than 10% (Teears, 2026). Applying the SST thresholds prevents inflating the relative abundance in the northern and southern regions by including spatial cells that have predictions of positive abundance that are unlikely in reality. The approach was a recommendation from the PAW.

The CPUE modeling for the 2026 assessment made some key improvements over the 2023 version, largely following detailed discussion and recommendations from PAW (Hamer, 2026). The 2026 analysis implemented encounter and positive component sub-models separately to allow for

differences in model formulas; specifically to allow the number of hooks to be included in the encounter sub-model. The inclusion of vessel ID as a random effect to account for vessel effects was another important improvement in the analysis, and the number of samples included in the sub-sampling of longline CPUE data for region specific indices was increased. Finally effort creep (positive increase in effective effort over time) was included as an adjustment on predicted regional CPUE indices. The decision at PAW was that zero effort creep was implausible, so a range of positive effort creep is explored based on the recommendation from PAW, and no zero options are included in the assessment models. Please refer to Teears (2026) for a full description of the CPUE modeling for generating the global and region specific abundance indices.

4.5.4 Other fisheries

Effort data for the ID, PH, and VN surface fisheries and Japanese research longline fisheries are unavailable. However, as these fisheries are not part of the index fisheries, with the catch-conditioned approached effort data are not necessary for these extraction fisheries. Catch estimates for the ID/PH/VN fisheries are derived from various port sampling programmes dating back to the 1960s for ID and the PH, and early 2000s for VN (Vidal et al., 2026; Williams & Ruaia, 2023).

4.6 Size data

Available length-frequency data for each of the fisheries were compiled into 95 x 2cm size classes from 10–12 cm to 198–200 cm. Weight data were compiled into 200 x 1 kg size classes from 0–1 kg to 199–200 kg. Most weight data were recorded as processed weights (usually recorded to the nearest kilogram). Processing methods varied between fleets requiring the application of conversion factors to convert the available processed weight data to whole weight equivalents. Details of the various processed weight to whole weight conversion factors are provided in Hamer et al. (2026). Note that these conversions occur in the database prior to data extraction for stock assessments (Hamer et al., 2025). Data were either collected onboard by fishers, through observer programs, or through port sampling, see descriptions of data sources in Hamer et al. (2025); Hamer et al. (2026). Each size-frequency record in the model consisted of the actual number of bigeye tuna measured and Figure 6 provides details of the temporal availability of length and weight-frequency data (also see Teears (2026)). Note that a maximum effective sample size of 1,000 is implemented in MFCL when using the robust normal likelihood for size composition data. The effective sample size was further down-weighted as explained in Section 5.5.2. Summaries of the available size composition data by year and fishery are provided in Teears (2026) and the processed data in Peatman et al. (2026b).

An extensive process of preparatory size-data processing is required to improve the reliability and representativeness of size compositions. This includes aggregating data by model region and fishery, filtering unreliable sources and small samples, and reweighting size frequencies to represent catches in extraction fisheries and relative abundance in index fisheries. Only length data are available for purse-seine, pole-and-line and other miscellaneous fisheries in the Indonesia–Philippines–Vietnam region; both weight and length observations are available for

longline fisheries. For the 2026 assessment, following earlier recommendations (Day et al., 2023; Teears et al., 2023) and PAW discussion (Hamer, 2026), longline weight observations were converted externally to length and combined with observed length data (Peatman et al., 2026b). The final assessment therefore used length-frequency compositions only; no weight-frequency compositions were fitted in MFCL. For each fishery and period, observed length data replaced converted-weight data when length coverage was greater, avoiding simultaneous use of both representations. The full preparation, conversion, filtering and reweighting process is described by Peatman et al. (2026b).

The Philippines domestic handline fishery in region 2 (Fishery 15.HL.PH.2) consistently catches the largest individuals in the WCPO. Handline fishing often takes place on mixed-gear trips with other gears such as hook-and-line targeting smaller fish. To avoid “contaminating” the length-frequency data for this fishery with fish that were mis-reported as being caught using a handline, fish smaller than 70 cm were excluded, and selectivity set to zero below 70 cm, consistent with previous assessments.

4.7 Tagging data

Of the three main tropical tuna species, bigeye tuna has the least amount of tagging data available. Information on the bigeye tag data characteristics, the process of constructing the MFCL tagging file, and estimating tag release correction factors (tagger effects) and reporting rate priors (from tag seeding experiments) can be found in Peatman et al. (2026c); Peatman et al. (2026a); Teears (2026). A summary of the tagging data and maps displaying tag displacements are in Teears (2026). Data were available from the Regional Tuna Tagging Project (RTTP) during 1989–92 (including affiliated in-country projects in the Solomon Islands, Kiribati, Fiji and the Philippines), the ongoing Pacific Tuna Tagging Programme (PTTP) that began in 2006, see Figure 12 and Figure 11. The Japanese Tagging Programme (JPTP) conducted by NRIFSF and the Ajinomoto Co. Inc, over the period 2000–2020 is also included. The new tagging data for the 2026 assessment comes from PTTP. As discussed previously historical (1995, 1999–2001) data from the Coral Sea tagging cruises by CSIRO (Evans et al., 2008) are no longer included.

After tagged fish are recaptured, there is often a delay before the tag is reported and the data are entered into the tagging databases. If this delay is significant then reported recapture rates for recent release events will be biased low and will impact estimates of fishing mortality and population scale in the terminal time periods of the assessment. For this reason, tag releases from the PTTP are included only up until the end of 2022. The 2023 assessment included tag data until 2019. Significant numbers of tagged bigeye have been released in recent (2020 and 2021) tagging cruises focused on bigeye and yellowfin in the central Pacific (model region 4). Recaptures from these releases are available for the 2026 assessment (Teears, 2026). Tags were released using standard tuna tagging equipment and techniques by trained scientists and technicians. Tags have been returned from a range of fisheries, having been recovered onboard or via processing and unloading facilities throughout the Asia-Pacific region, archival and double tagged fish are not included.

The numbers of tag releases input to the assessment model are adjusted for a number of sources of tag loss, unusable recaptures due to lack of adequately resolved recapture data, estimates of tag loss (shedding and initial mortality) due to variable skill of taggers (i.e., tagger effects), and estimates of base levels of tag shedding and tag mortality. These adjustments are described in more detail in Peatman et al. (2026c); Teears (2026). An additional issue for the bigeye assessment is that there are a considerable number of tag returns that were released within the WCPO but recaptured to the east of longitude 150°W, outside the WCPO assessment region. The adjustment or rescaling of releases for recaptures in the EPO preserves the recovery rates of tags from individual tag groups that would otherwise be biased low given that a considerable proportion of recaptures cannot be attributed to a recapture category in the assessment. These procedures were first described in Berger et al. (2014) and McKechnie et al. (2016). However, the releases are not adjusted for emigration outside of the model domain to the EPO. This is an important issue that requires further consideration as to options for informing adjustments to release numbers to account for emigration to the EPO.

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

5 Model description

5.1 General characteristics

The model can be considered to consist of several components, (i) the dynamics of the fish population; (ii) the fishery dynamics; (iii) the dynamics of tagged fish populations (iv) the observation models for the data; (v) the parameter estimation procedure; and (vi) stock assessment interpretations. Detailed technical descriptions of components (i)–(iv) are given in Hampton & Fournier (2001) and Davies et al. (2026b), and summarised below. In addition, we describe the procedures followed for estimating the parameters of the model and the way in which stock status conclusions are drawn relative to a series of reference points.

5.2 Population dynamics

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

population dynamics processes are as follows.

5.2.1 Recruitment

Recruitment is defined as the appearance of age-class 1 quarter fish (i.e. fish averaging ~ 24 cm FL) in the population. Bigeye tuna spawning does not typically follow a strong seasonal pattern, especially in the tropical waters where most of the spawning occurs. Rather it occurs sporadically all year and is thought to be influenced by food availability (Itano, 2000; K. M. Schaefer et al., 2005; Sun et al., 2006). The assessment model assumed that recruitment occurs instantaneously at the beginning of each quarter. This is a discrete approximation to continuous recruitment, but provides sufficient flexibility to allow a range of variability to be incorporated into the estimates as appropriate. The proportion of total recruitment occurring in each region was initially set relative to the variation in average regional catch and then estimated during the later phases of the fitting procedure. Time-series variability in this proportion was estimated within the model and allowed to vary in a relatively unconstrained fashion.

In recent assessments of tuna in the WCPO, the last few recruitment deviations have little information to inform them and these are usually fixed at the mean of the estimated recruitment deviations, which also reduces problems with retrospective analyses. This approach has been continued here with the last six recruitment deviations fixed at the geometric mean of the estimated recruitment deviations, which is appropriate for a log-normally distributed random variable.

Spatially-aggregated (over all model regions) recruitment was assumed to have a weak relationship with annual mean spawning potential via a Beverton and Holt stock-recruitment relationship (SRR) with a fixed value of steepness (h). Steepness is defined as the ratio of the equilibrium recruitment produced by 20% of the equilibrium unexploited spawning potential to that produced by the equilibrium unexploited spawning potential (Francis, 1992; Harley, 2011). Typically, fisheries data are not very informative about steepness (ISSF, 2011); hence, it was fixed at 0.90 in the diagnostic model and sensitivity analyses used values of 0.65, 0.80 and 0.95. The high CV (2.2) of the annual log-recruitment deviations ensured that the SRR had negligible influence on recruitment and other model estimates, as recommended by Ianelli et al. (2012).

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

5.2.2 Initial population

The population age structure in the initial time period in each region was assumed to be in equilibrium, and the initial fishing mortality was assumed to be zero ($Z = M$). As noted above,

the population is partitioned into quarterly age-classes with an aggregate class for the maximum age (plus-group, 40 quarters). The aggregate age class makes it possible for accumulation of old fish, which is likely in the early years of the fishery when exploitation rates were very low.

5.2.3 Growth

The standard assumptions for WCPO assessments fitted in MFCL were made concerning age and growth: i) the lengths-at-age are normally distributed for each age-class; ii) the standard deviations of length for each age-class are a log-linear function of the mean lengths-at-age; and 3) the probability distributions of weights-at-age are a deterministic function of the lengths-at-age and a specified weight-length relationship⁷. These processes are assumed to be regionally and temporally invariant.

In previous assessments several approaches to growth estimation were explored, including external otolith based growth curves (Farley et al., 2020), a growth curve based on the otoliths with the addition of tag-recapture growth increment data (Eveson et al., 2020), an internal MFCL growth estimation using a conditional-age-at-length (CAAL) dataset (Farley et al., 2020), which also made use of the modal progressions apparent in various size composition data (Ducharme-Barth et al., 2020a). Results were sensitive to approaches used for growth. The peer review of the yellowfin assessment (Punt et al., 2023) carefully considered the approaches to estimating growth. They concluded that based on the otolith sampling protocols, with otoliths selected according to length (i.e., attempting to achieve sufficient numbers of otolith samples across the full length range, plus bias in otolith readability with age), that any external growth curves would likely be biased. They strongly recommended that for this type of otolith sampling regime that age data should be conditional on length and the growth estimated internally, using conditional-age-at-length data within MFCL. This approach was applied in the 2023 assessment and is continued for this assessment to estimate the von Bertalanffy growth curve internal to the MFCL model, considering also information on growth from size composition data. We also note the substantial enhancement of the CAAL data set through the efforts of SPCs recently established schlerochronology lab (Wakefield et al., 2026) and the contributions of Japan’s FRA (Hasegawa et al., 2026).

5.2.4 Movement

Movement was assumed to occur instantaneously at the beginning of each quarter via movement parameters that connect regions sharing a common boundary. Note that fish can move between non-contiguous regions in a single time step due to the “implicit transition” computational algorithm employed (see Hampton & Fournier (2001) and Davies et al. (2026b) for details). Movement is parameterized by a pair of bi-directional parameters at each region boundary. Movement is therefore possible in both directions across each regional boundary in each of the four quarters. Each of these coefficients is estimated independently resulting in 56 estimated movement parameters for the 5 region spatial structure ($2 \times \text{no. region boundaries} (7) \times 4 \text{ quarters}$).

⁷The length-weight relationship has been updated for the current assessment based on an analysis of current and historical port sampling data under WCPFC Project 90 with bias correction applied (Hamer et al., 2026)

There are limited data from which to estimate long-term, annual variation in movement or age-specific movement rates. As such, the estimated seasonal pattern is assumed to be fixed across years and the movement parameters are invariant with respect to age. A “prior” of 0 is assumed for all movement parameters, and a low penalty is applied to deviations from the “prior”. We ran a preliminary model with fixed movement parameters, using outputs from the SEAPODYM model, but this model could not be written into this paper due to insufficient time to explore the results and run the required diagnostics.

5.2.5 Natural mortality

The instantaneous rate of natural mortality (M) consists of an average over age classes and age-specific deviations from the average M . Average M can be estimated internally by the model or specified as a fixed value. Internal estimation can be constrained by providing a prior value and a penalty weight for deviations from the prior. Age-specific deviations can also be estimated or entered as a fixed M -at-age function. The latter approach was taken in the 2020 bigeye and yellowfin assessments (Ducharme-Barth et al., 2020a; Vincent et al., 2020), using an approach applied to other tunas in the WCPO and EPO (Harley & Maunder, 2003; S. D. Hoyle, 2008; Hoyle & Nicol, 2008). That external function depends on empirical and assumed inputs for sex ratio, growth, male natural mortality and critical lengths, all of which have uncertainty, particularly for large and old fish. It also requires recalculation under each alternative growth curve. Following the 2023 yellowfin peer review and the CAPAM Tuna Stock Assessment Good Practices Workshop, the 2023 bigeye assessment instead estimated a Lorenzen M -at-age function internally (Day et al., 2023). However, that model was not stable to jittering, and SC19 raised concern that jointly estimating growth and M may contribute to instability. The 2026 PAW therefore recommended fixing the adult/asymptotic M scale, based on assumed M at age 40 quarters, initially at the 2023 diagnostic-model estimate, while retaining the inverse relationship between the Lorenzen shape and the internally estimated growth curve (Hamer, 2026).

In this assessment we continue to use the Lorenzen functional form for M -at-age (with M -at-age being inversely proportional to the mean length at age) but to fix the asymptotic value of M , i.e. the M for age 40 quarters, and treat this fixed M as an uncertainty in the uncertainty ensemble used for management advice. Further, as noted previous, a companion paper estimated M based on tag-recapture data (Ducharme-Barth et al., 2026), and the results of that study are included in the M uncertainty range for this assessment.

5.2.6 Reproductive potential

The reproductive potential ogive is an important component of the assessment structure as it translates model estimates of total population biomass to the relevant management quantity, spawning potential biomass (SB). Assumptions about maturity do not affect the process of fitting the model, only the reference point values. The approach for calculating the reproductive potential at length ogive in this assessment is the same as the 2023 assessment (Day et al., 2023). MFCL estimates the reproductive potential at age ogive internally from the externally calculated reproductive potential ogive at length. This length-based ogive is converted internally to the

reproductive potential-at-age using a smooth-spline approximation (Davies et al., 2019). This allows for a more natural definition of reproductive potential as the product of three length-based processes: proportion females-at-length⁸ (sex-ratio), proportion of females mature-at-length⁹, and the fecundity-at-length of mature females¹⁰ (Figure 14). Another added benefit is that this reproductive potential ogive is growth invariant. As in the 2023 assessment, we have not included spawning fraction in the reproductive potential calculation, as these data are still not adequate for bigeye in the WCPO.

5.3 Fishery dynamics

5.3.1 Selectivity

Selectivity has typically been estimated as a function of age within MFCL. Recent developments in MFCL, based on recommendations from the yellowfin peer review (Punt et al., 2023), have developed an option to estimate length-based selectivity, which was applied in the recent skipjack assessment (Teeares et al., 2025). We attempted the length-based selectivity option, however, despite various configurations and model runs we could not achieve better/satisfactory results in terms of data fits than the traditional age-based selectivity. We ran out of time for deeper exploration and decided to continue with age-based selectivity for the 2026 assessment.

Estimated selectivity curves at age are included in Figure 40 and their equivalent length-based curves are shown in Figure 41. Selectivity was modelled using cubic splines, which provide more flexibility than a fixed functional relationship with age and require fewer parameters than separate age-specific estimates. The number of spline nodes was chosen to represent the required complexity: five nodes were used for most fisheries and seven for the two associated purse-seine fisheries spanning Regions 3 and 4 (Table 2). Selectivity was time-invariant and fishery-specific in this assessment.

5.4 Dynamics of tagged fish

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

5.4.1 Tag reporting

In theory, tag-reporting rates can be estimated internally within the model. In practice, experience has shown that independent information on tag-reporting rates for at least some fisheries is required for reasonable model behavior to be obtained. We provided reporting rate priors (penalties) for reporting groups (combinations of tagging programme and fishery) that reflect independent estimates of the reporting rates and their variances from tag seeding (Peatman et al., 2026a). For the RTTP and PTTP, relatively informative “priors” were formulated for

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

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

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

the equatorial purse seine fisheries given that tag seeding experiments were focused on purse seiners. We also make some assumptions regarding fisheries that have similar characteristics to those with independent estimates, but increase the prior variance for these fisheries. For others where there was very little information to inform priors and variance, uninformative priors were allocated, or in cases where there were 5 or less tag returns, tag reporting rates for these groups were not estimated. In these cases, the small numbers of tag returns were removed from the tagging data input file and the reporting rate for these groups set to a fixed value of zero.

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

All reporting rates were assumed to be time-invariant and have an upper bound of 0.99. Tag recapture and reporting rate groupings are provided in Table 5. Previous assessments have assumed fishery-specific reporting rates are constant over time. This assumption was reasonable when most of the tag data were associated with a single tagging program. However, tag reporting rates may vary considerably between tagging programs due to changes in the composition and operation of individual fisheries, and different levels of awareness and follow-up. Consequently, fishery-specific tag reporting rates that are also specific to individual tagging programs were estimated.

5.4.2 Tag mixing

The population dynamics of the fully recruited tagged and untagged populations are governed by the same model structures and parameters. The populations differ in respect of the recruitment process, which for the tagged population is the release of newly tagged fish, i.e., an individual tag and release event is akin to a recruitment event for the tagged population. Implicitly, we assume that the probability of recapturing a given tagged fish is the same as the probability of catching any given untagged fish in the same region and time period. For this assumption to be valid, either the distribution of fishing effort must be random with respect to tagged and untagged fish

and/or the tagged fish must be randomly mixed with the untagged fish. The former condition is unlikely to be met because fishing effort is almost never randomly distributed in space. The second condition is also unlikely to be met soon after release due to the time required for mixing with the untagged population to take place.

Depending on the distribution of fishing effort in relation to tag release sites, the probability of capture of tagged fish soon after release may be different to that for the untagged fish. It is therefore desirable to designate a time period after release as “pre-mixed” and compute fishing mortality for the tagged fish during this period based on the actual recaptures, corrected for tag reporting (noting the previous discussion on this) and tagging effects, rather than use fishing mortalities based on the general population parameters. This, in effect, desensitizes the likelihood function to tag recaptures in the specified “pre-mixed” periods while correctly removing fish from the tagged population after the “pre-mixed” period. Herein we refer to the “pre-mixed” period as the “mixing period”. In short, tags recaptured within a specified mixing period do not contribute to the tagging likelihood component while they can still be considered in relation to reporting rate estimations.

A key aspect of fitting to tag-recapture data is the assumptions on mixing periods. The allocation of appropriate “mixing periods” in stock assessments that use tag-recapture data is problematic as model estimations are sensitive to this assumption and misspecification can bias estimation of management quantities (Kolody & Hoyle, 2015). Mixing rates may vary depending on release locations and times depending on various factors, including; oceanographic dynamics, contexts of releases (e.g., FADs versus free schools, archipelagic waters versus oceanic), fishing effort distribution and environmental/food conditions that influence movements.

Previous bigeye assessments applied fixed one- or two-quarter mixing periods to all tag-release groups, based on qualitative comparisons of post-release recaptures and purse-seine catch maps (Tearns et al., 2023). The 2023 yellowfin peer review noted the individual-based tag-mixing approach developed for skipjack (Castillo Jordan et al., 2022; Scutt Phillips et al., 2022). It simulates dispersal and exposure to fishing mortality for each release group and a contemporaneous untagged population, accounting for the location, timing, environment and fishing context of each release event. The Ikamoana Lagrangian model tracks individual particles using forcings and parameters from SEAPODYM (Scutt Phillips et al., 2018; Senina et al., 2020). The peer-review panel recommended exploring this approach for yellowfin and bigeye, and release-group-specific mixing periods were subsequently estimated for this assessment (Scutt Phillips et al., 2026).

With the individual based modeling approach, mixing is estimated to have occurred for each tag release group at a point in time post-release when the distribution of recapture probabilities for tagged fish is similar to that of the untagged population. The similarity of the recapture probabilities for tagged and untagged fish is gauged using the Kolmogorov-Smirnov K-Statistic, which indicates the level of dissimilarity between the two distributions. Lower values of the K-statistic indicate the distributions are more similar (i.e. more well mixed). The dissimilarity K-statistic was estimated for each release group (noting not all groups could be simulated) under

four assumed mixing periods (zero, one, two and three quarters), using the results from the individual-based Ikamoana tagging simulations (SC22-SA-IP-10), giving four values that describe how the representativeness of the tagged group’s experienced fishing pressure changes as more time is allowed for mixing. To translate these into a single, release group specific mixing period for use in the assessment, a cutoff or threshold value of K is chosen to define the level of dissimilarity considered acceptable, where a lower cutoff is a stricter requirement for representativeness. For a given release group, the assigned mixing period is then the shortest of the four assumed periods at which K first falls below this cutoff. In other words, the tagged group is allowed the least possible time to mix that is consistent with reaching the required level of representativeness, so that as many recaptures as possible are retained for inclusion within the MFCL likelihood calculation. If none of the four assumed mixing periods achieves a K below the chosen cutoff, suggesting that the group never becomes sufficiently representative within the simulated range, then a default mixing period of four quarters is assigned, which in practice excludes the great majority of that group’s recaptures from the likelihood. This choice of cutoff or threshold value therefore acts as a single control on how conservatively tagging data are treated across the assessment: lower K -statistic cutoffs impose stricter representativeness, generally lengthening the assigned mixing periods, and admitting fewer recaptures, providing continuous values that can be varied as a structural sensitivity or in the uncertainty ensemble.

To note the mixing simulations could only be conducted on release group pre-2019 as the SEAPODYM model forcings did not extend beyond that time period. This means the recent tag releases (not included in the 2023 assessment) could not have simulated mixing periods. Further, no releases were simulated from JPTP in region 1 (see (Scutt Phillips et al., 2026)). For release groups that were not directly simulated, due to either computational reasons or a lack of forcing data from SEAPODYM, a mixing period was assigned using averaged Kolmogorov dissimilarity values from simulated groups. Specifically, for each mixing scenario (zero, one, two and three quarters), the mean Kolmogorov dissimilarity K was taken across the simulated release groups either across all groups, or across all groups within a region from which an indication of region-specific mixing was required, producing a single representative K value per scenario. The cutoff selection rule was then applied to this averaged set of four values exactly as it was for an individual group: the assigned mixing period is the shortest scenario whose mean K falls below the chosen cutoff, or four quarters if none does. It is important to note that the averaging is performed on the dissimilarity metric, at each mixing scenario, and not on the resulting mixing periods themselves. Averaging the mixing periods would combine values that are already the outcome of a threshold rule, discarding information about how close each group was to its cutoff and giving a non-integer result requiring arbitrary rounding; averaging the underlying K values instead preserves the full recapture probability distribution information from simulations, and applies the selection rule once in the same way as for a simulated group.

5.5 Likelihood components

There are four data components that contribute to the log-likelihood function for the bigeye stock assessment: the index fishery CPUE data; the length frequency data, the tagging data and

the conditional-age-at-length data.

5.5.1 Index fishery CPUE likelihood

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

5.5.2 Length frequency

In the 2023 assessment, size-frequency compositions were fitted with a robust-normal likelihood, with nominal sample sizes divided by 10, 20 (diagnostic model) or 40 to allow for overdispersion. A Dirichlet–multinomial (DM) alternative was also explored, but an effective-sample-size upper asymptote of 1,000 allowed the composition data to receive high weight and degraded the fits to abundance indices, so it was not retained (Thorson et al., 2017).

The 2026 diagnostic model used the DM likelihood without random effects and grouped the 33 fisheries into eight composition groups (Table 4). In this parameterisation, MFCL calculates the incident effective sample size as $N_{\text{eff}} = N_{\text{max}}(1 + \lambda)/(N_{\text{max}} + \lambda)$; consequently, N_{max} is an asymptotic upper bound rather than an effective sample size assigned to every composition. Preliminary fits evaluated values including 10, 25, 40, 100 and the MFCL default of 1,000. From 40 upward, length-frequency fit improved slightly but CPUE fit deteriorated as the composition component gained influence. As an independent empirical reference, the 95th percentile of composition-level effective sample sizes obtained by Francis reweighting ranged from 22.22 to 23.81 across 2,399 positive length-frequency compositions (Francis, 2011). We therefore set $N_{\text{max}} = 25$, immediately above that upper-tail range, while assessing the joint CPUE and composition fits rather than treating the Francis estimates as DM weights.

The final controls are summarised in Table 3. The common log-concentration intercept (`fish_pars(22)`) had reached its upper bound of 7 in preliminary estimation and was fixed at 7 (fish flag 69 = 0); it was not re-estimated in the diagnostic model. The relative-sample-size exponent (`fish_pars(23)`) was estimated from Phase 2 for each of the eight groups (fish flag 89 = 1). This retained internal estimation of relative composition information while avoiding renewed estimation of an already boundary-limited concentration intercept.

5.5.3 Tag data

We applied the same approach as the previous assessments which set a fixed tau (τ) across all release groups/recapture fisheries. An alternative to this approach is to allow τ to be estimated (either overall or by groups of recapture fisheries) as done in the 2025 skipjack assessment (Tears et al., 2025), this effectively allows τ to vary somewhat to achieve a better overall objective function within the joint likelihood. The τ estimation will be impacted by all the data in the

model, and through the negative binomial assumptions it is influenced by the overdispersion in the tag data that it can detect. τ is not simply a penalty weight on the neg-binomial term. It is implicit in the likelihood formulation. We began experimenting with estimating τ late in the assessment process, but needed more time to understand it, and found, in the limited exploration, it tend towards a value of 1. The estimation of τ maybe a reasonable approach, but requires further time to consider more thoroughly for this assessment.

As discussed in Section 7.2.5, the tag data was in conflict with the CPUE indices and to a lesser extent the length data. This needed to be considered within the uncertainty characterisation. There are several options for this, the most practical is to set fixed higher levels of the τ . τ determines the variance of the negative binomial distribution in relation to the Poisson distribution, for which the variance is equal to the mean. Therefore, the τ parameter value influences the weighting of the tagging data in the overall model joint likelihood. As noted above, similar to McKechnie et al. (2017a), initial attempts to estimate τ , assumed common for all fisheries and tag release groups, tended to estimate a value close to 1, i.e. Poisson, assumes little or no over dispersion. Tag recapture data is likely to have a degree of over dispersion due to clustering of recaptures, although it likely depends on the level of dispersal and mixing after release. Recaptures of tag groups shortly after release (that might fall in the mixing period) could be expected to have higher overdispersion than those recaptured after longer periods at liberty. Assigning a fixed τ to all fisheries and tag release groups is perhaps not ideal, but a pragmatic approach given the limited time to statistically explore the tag recapture overdispersion properties outside the model.

Late in this assessment process when exploring the τ flag setting and the MFCL code in depth we found that the last assessments settings, unbeknownst to those analysts, actually resulted in a fixed τ of only slightly >1 (1.02), close to Poisson, rather than the value of 2 as noted in those previous assessment reports (Day et al., 2023; Ducharme-Barth et al., 2020a). This issue was not obvious and required digging into MFCL code to discover. This meant that the current diagnostic model that we had already fully diagnosed had a τ of 1.02. This is not desirable for tag recapture data, as noted above. We therefore ran an additional model with τ set to 2, as this value appears to be considered suitable in previous assessment methods, as it approaches a negative binomial with variance of twice the Poisson, which can better account for overdispersion, but could substantially downweight the tag data. This setting did allow a comparison between a model that placed a high weight on the tag data versus a model that downweighted the tag data. The value of τ may also interact with the mixing period assumptions in the uncertainty ensemble, such that scenarios with longer tag mixing time, and therefore less tags remaining in the joint likelihood estimation, would be subject to less overdispersion, but be even more severely downweighted with an increased τ . The use of τ , along with other assumptions required for fitting tag recapture data in the joint likelihood requires further technical review and guidance for future MFCL tuna assessments. Ultimately, we added an additional step to the stepwise and chose the τ of 2 as the diagnostic model setting.

5.5.4 Conditional age-at-length data (CAAL)

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

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

5.5.5 Parameter estimation and uncertainty

The parameters of the model were estimated by minimising the sum of the negative log-likelihoods associated with each of the data components plus the negative log of the probability density functions of the priors and penalties specified in the model. The optimisation is continued either until a pre-specified minimum gradient is achieved or until a maximum number of iterations is reached. This optimisation is performed by an efficient algorithm using exact derivatives with respect to the model parameters (auto-differentiation, Fournier et al. (2012)). Estimation was conducted in a sequence of phases, the first of which uses somewhat arbitrary starting

parameter values. A bash shell script, or “doitall”, implements the phased sequence for gradually estimating more and more parameters for the model. Some parameters were assigned starting values consistent with available biological information. The values of these parameters are provided in the .ini input file or within the MFCL code.

5.6 Diagnostic methods

5.6.1 Convergence

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

1. 1. Minimum gradient sufficiently small, preferably ≤ 0.0001 .
2. 2. Hessian matrix was positive definite.
3. 3. Models were not overly sensitive to jitter analyses (i.e. minor sensitivity of management quantities and -ve log-likelihood).
4. 4. The number of parameters estimated on or close to bounds was small, preferably zero.
5. 5. The number and composition of pairs of highly correlated estimated parameters ($|r| > 0.90$).

Obtaining PDH solutions in models with high numbers (i.e. $>1,000$) of estimated parameters can be challenging, and failure is often due to one or a few very small negative eigenvalues that have little influence on the estimations of key management quantities. MFCL has methods for computing estimation uncertainty for Hessians with negative eigenvalues, allowing the parameter estimation uncertainty to be calculated for Hessians with a small numbers of small negative eigenvalues. It has been shown by comparing pairs of similar models with the same structure, but with slightly different solutions, one with and one without a PDH, that the estimation of parameter uncertainty is very similar (e.g., appendix 15.3 Castillo Jordan et al. (2022)). While a PDH solution is an important diagnostic to meet for a diagnostic model, this criterion may not be essential (or practicable) for all models in the uncertainty ensemble. While it is possible to simply filter out any model that does not achieve a PDH from the uncertainty ensemble, it is worth also considering the number of -ve eigenvalues, what parameters these relate to, the maximum gradient and the plausibility of the estimated management quantities. The purpose of the uncertainty ensemble is to provide a distribution of plausible management quantities that captures a range of structural, biological and or data uncertainties, rather than to fine tune each individual model to achieve the best possible diagnostics, which in any case would not be possible in the time available. Further, the structural, biological and data uncertainties are typically much more consequential than the parameter estimation uncertainties.

5.7 Other diagnostics

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

Model fits to data Plots of observed and predicted index, LF and tagging data were examined, including residuals plots.

Jitter analysis Thirty phase-aware jitter fits were used to test sensitivity to starting values and local minima (Carvalho et al., 2021; Subbey, 2018). This implementation was more rigorous than applying jitter only to the final maximum-likelihood estimate: it began from the ordinary `makepar`-generated fitting path, used the completed diagnostic fit to identify which parameters were estimated and in which phase they became active, and then repeated the full fitting schedule.

1. Thirty runs used a perturbation CV of 0.1 across the 12 estimation phases. Only parameters estimated in the completed diagnostic model were included.
2. After Phase 1, parameters already active were perturbed; parameters first estimated in later phases were perturbed when their phase began. This avoided imposing inappropriate fitted values on parameters that were not yet active and allowed each run to explore from the same phase-specific initialisation used in the original fit.
3. The stated CV controlled perturbation size on a parameter-appropriate scale: mean-preserving proportional changes for positive parameters, additive normal changes for unconstrained or near-zero parameters based on their parameter-family scale, and bound-preserving changes for bounded parameters.
4. The maximum gradient component (MGC) is the largest absolute component of the objective-function gradient with respect to estimated parameters; smaller values indicate closer approach to a stationary point. Runs with $\text{MGC} \leq 0.0001$ were retained, and their objective values and derived quantities were compared with the unperturbed diagnostic fit.

Results for Jitter models are shown in Section 7.2.2.

Self tests This assessment included the first formal simulation self-test in the WCPO tuna assessment series, following the general simulation–estimation framework of Deroba et al. (2015) and the practical-estimability interpretation used by Kim et al. (2024). The diagnostic model was treated as the operating model and 50 independent pseudo-data sets were generated with replicate-specific seeds. CPUE, length-composition, age-at-length and post-mixing tag observations were simulated; catch and empirical pre-mixing tag observations remained conditioned. Each data set was refitted with the recorded MFCL `doitall` schedule, and all 50 refits met the archived MGC criterion of 10^{-4} . Relative recovery error was calculated against the generating quantity for each replicate. These empirical distributions test internal consistency and practical estimability under the configured observation models; they are not parameter confidence intervals and do not determine whether the diagnostic model is an unbiased representation of the unknown population. Results are presented in Section 7.2.3.

ASPM A strict age-structured production model (ASPM) diagnostic (Carvalho et al., 2017; Carvalho et al., 2021; Maunder & Piner, 2015; Minte-Vera et al., 2021) was fitted by conditioning on the observed catches and retaining only the five CPUE indices in the observation likelihood. Tagging, length- and weight-frequency, catch-at-age, conditional age-at-length and penalty

components were excluded. Biology, movement, selectivity, regional recruitment allocation and the initial age–region structure were fixed at their diagnostic-model values. Temporal relative recruitment was set to one, regional recruitment deviations were set to zero and the stock–recruit relationship was inactive; the only estimated parameter family was mean recruitment, which determines the overall population scale. Comparison with the integrated model therefore shows the biomass scale and trend supported by the catch-conditioned CPUE signal under constant recruitment (Section 7.2.6).

Retrospective Retrospective analyses (Hurtado-Ferro et al., 2015) were conducted as a general test of the stability of the model, as a robust model should produce similar output when rerun with data for the terminal quarters sequentially excluded (Cadigan & Farrell, 2005). Retrospective analyses with 7 peels for the diagnostic model are presented in Section 7.2.4. The Mohn’s rho statistic (Mohn, 1999), a measure of the average relative bias of retrospective estimates, was computed for recruitment, SB and spawning potential depletion ($SB_t/SB_{F=0[t]}$) and annual F , to indicate whether significant retrospective bias was present in the model.

Likelihood profiles For highly complex population models fitted to large amounts of often conflicting data, absolute estimates of total abundance can be unstable. Therefore, a likelihood profile analysis was undertaken of the marginal penalized likelihood in respect of population scaling, following the procedure outlined by McKechnie et al. (2017a) and Tremblay-Boyer et al. (2017). The results of these procedures are presented for the diagnostic model Section 7.2.5. Likelihood profiling is initially conducted on the main data components, i.e. CPUE indices, size data, tag data, CAAL data. If significant conflict is found among the main data components, additional profiling is conducted within each data component to explore the sources of conflict.

5.8 Stock assessment interpretation methods

5.8.1 Depletion and fishery impact

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

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

5.8.2 Reference points

The unfished spawning potential ($SB_{F=0}$) in each time period was calculated given the estimated recruitments and the Beverton–Holt SRR. For management reporting, annual dynamic depletion compares spawning biomass in year y with mean unfished spawning biomass over the preceding ten years, excluding year y .

Two depletion quantities were used for management reporting. The WCPFC-adopted limit reference point (LRP) for bigeye is $SB_{recent}/SB_{F=0} = 0.20$, where the numerator is mean spawning biomass during 2021–2024 and the denominator is mean unfished spawning biomass during 2014–2023. CMM 2025-02 paragraph 10 separately states that, pending agreement on a target reference point, the spawning-biomass depletion ratio is to be maintained at or above its 2012–2015 average. This comparison used the 80 equal-weight central models and excluded estimation uncertainty; the result is reported in Table 9, with the other management and reference-point quantities in Table 8 and Table 11.

The other key reference point, F_{recent}/F_{MSY} , is the estimated average fishing mortality at the full assessment area scale over a recent period of time (F_{recent} ; 2020–2023 for this stock assessment) divided by the fishing mortality producing MSY which is a product of the yield analysis and is detailed in Section 5.8.3.

Several ancillary analyses using the converged model/s were conducted in order to interpret the results for stock assessment purposes. The methods involved are summarized below and the details can be found in Davies et al. (2026b).

5.8.3 Yield analysis

The yield analysis consists of computing equilibrium catch (or yield) and spawning potential, conditional on a specified basal level of age-specific fishing mortality (F_a) for the entire model domain, a series of fishing mortality multipliers ($fmult$), the natural mortality-at-age (M_a), the mean weight-at-age (w_a) and the SRR parameters. All of these parameters, apart from $fmult$, which is arbitrarily specified over a range of 0–50 (in increments of 0.1), are available from the parameter estimates of the model. The maximum yield with respect to $fmult$ can be determined using the formulae given in Davies et al. (2026b), and is equivalent to the MSY. Similarly, the spawning potential at MSY SB_{MSY} can be determined from this analysis. The ratios of the current (or recent average) levels of fishing mortality and spawning potential to their respective levels at MSY are determined for all models of interest. This analysis was conducted for all models in the structural uncertainty grid and thus includes alternative values of steepness assumed for the SRR.

Fishing mortality-at-age (F_a) for the yield analysis was determined as the mean over the recent period 2020–2023, consistent with the definition of F_{recent} used in this assessment. We do not include 2024 in the average as fishing mortality tends to have high uncertainty for the terminal data year of the analysis and the catch and effort data for this terminal year are potentially incomplete. Additionally, recruitments for the terminal year of the model are constrained to be the geometric mean across the entire time series, which affects the F for the youngest age classes.

MSY was also computed using the average annual F_a from each year included in the model (1952–2024). This enabled temporal trends in MSY to be assessed and a consideration of the differences in MSY levels under historical patterns of age-specific exploitation.

5.8.4 Kobe analysis and Majuro plots

For the standard yield analysis (Section 5.8.3), the fishing mortality-at-age, F_a , is determined as the average over some recent period of time (2020–2023). In addition to this approach the MSY-based reference points (F_t/F_{MSY}), and SB_t/SB_{MSY} are computed by repeating the yield analysis for each year in turn. This enabled temporal trends in the reference points to be estimated and a consideration of the differences in MSY levels under historical patterns of age-specific exploitation. This analysis is presented in the form of dynamic Kobe and Majuro plots, which have been presented for all stock assessments in recent years.

5.8.5 Stock projections from the structural uncertainty grid

Thirty-year stochastic projections were conducted within MFCL for 2025–2054, following the general framework in Pilling et al. (2016). Each of the 80 retained assessment models was paired with ten future recruitment sequences, giving 800 model–recruitment realisations. Future recruitment deviations were sampled from the model-specific estimates for 1972–2023, excluding the first 20 assessment years and terminal 2024. All 33 fisheries were catch-conditioned separately by fishery and quarter at their observed 2022–2024 mean catches; missing catch observations were treated as zero when calculating these means. Projection intervals combine structural uncertainty and future recruitment variability but do not include Hessian parameter draws. Results are reported in Section 7.5.6.

6 Model runs

6.1 Developments from the last assessment

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

6.1.1 Stepwise model development

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

1. **[01]** 2023 diagnostic-model refit: Provide a reproducible reference of the previous diagnostic model for subsequent comparisons.
2. **[02]** New MFCL executable updated from archived 2.2.2.0 (2023 assessment) to the recent 2.2.7.9 version used in this assessment, plus new format INI 1007: Isolates the new executable effect
3. **[03]** Lorenzen natural-mortality scaling fixed to the 2023 diagnostic-model estimate: Improve model stability, recommended at the PAW.
4. **[04]** Bias-corrected length-weight parameters: Update the biomass conversion for the 2026 assessment.
5. **[05]** Convert to 5-region, 33-fishery structure (28 extraction, 5 index fisheries): Represents revised spatial and fishery definitions, supported by PAW.
6. **[06]** Size-composition inputs replaced with the reweighted weight data converted to lengths for longline fisheries, data until 2021: Evaluate the conversion of weight data to lengths, and all size compositions input as length frequencies.
7. **[07]** Weight converted to lengths plus observed-length compositions for longline fisheries (i.e. adds the observed length data, noting previous bigeye assessment only used weight data for longline fisheries): Evaluate the effect of using observed lengths where their catch coverage exceeded that of weight samples (Peatman et al., 2026b)
8. **[08]** Update data through to 2024, except for CAAL data: Extend the temporal coverage of the assessment.
9. **[09]** Remove data <70cm for PH and ID handline fisheries, remove data >90 cm for domestic mixed gear fisheries in ID and PH: replicate this data truncation applied in previous assessments to deal with non-representative data.
10. **[10]** Switch to separate sdmTMB regional CPUE indices and apply the MFCL regional scaling penalty feature based on a global sdmTMB model: Penalize relative abundance differences among model regions but allow region specific CPUE indices to have individual selectivity and catchability, strong recommendation from PAW and yellowfin assessment peer review (Hamer, 2026; Punt et al., 2023).
11. **[11]** Time-varying CPUE uncertainty: Account for temporal variation in the relative precision of the CPUE indices.
12. **[12]** Regional scale CPUE log-scale observation-error SDs fixed at 0.35, 0.24, 0.21, 0.24 and 0.23: Set the five regional log-scale observation-error standard deviations to values obtained by maximum-likelihood estimation and retain them in subsequent steps.
13. **[13]** Add the new conditional age-at-length data (weighting factor = 0.75): Use the 2023 assessment weighting as the reference treatment for the new age data (i.e. no regional reweighting of the data contributions)

14. [14a] Apply the regional CAAL data reweighting by all 5 regions (i.e. reweight observed sample sizes based on relative mean CPUE among regions from global sdmTMB model, as discussed at PAW and in (Hasegawa et al., 2026): Test region-level spatial weighting.
15. [14b] As previous but combining CAAL data from regions 3 and 4 (sub-basin CAAL reweighting by CPUE): Sub-basin reweighting of the CAAL data.
16. [15] Selectivity modifications, flexible fishery-specific selectivity plus weak F10 non-decreasing penalty.
17. [16] Apply release-group-specific tag-mixing periods with K-statistic cut-off 0.2: Assign release-group-specific mixing periods (see appendix Tables in Scutt Phillips et al. (2026)
18. [17] Tag reporting rate adjustment exclusion (MFCL tag-flag02=1) in pre-mixing window: Avoid applying poorly determined or assumed reporting rates within the pre-mixing windows close to release events, post-mixing reporting-rate activation is unchanged.
19. [18] Effort-creep adjustment added to CPUE indices: Account for gradual changes in fishing efficiency.
20. [19] Apply Dirichlet-multinomial (DM) composition weighting: Estimate length-composition overdispersion internally using eight fishery groups and an effective-sample-size upper asymptote of 25.
21. [20] Increase tau (τ) from 1 to 2, a value noted in previous assessments to account for assumed overdispersion in tag recapture data using negative binomial rather than Poisson.
22. [21] Add weak non-decreasing penalty on F33 (region 5 index)
23. [22] Steepness = 0.9 (mode of ensemble distribution and more defensible for tropical tuna)
24. **Diagnostic2026**

We note that many additional models were explored in the development of the diagnostic model, these are not described in this report.

6.2 Sensitivity analyses and uncertainty ensemble

6.2.1 Sensitivities

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

We primarily focused on the key sensitivities below:

1. Steepness: 0.65, 0.95 (diagnostic model 0.9)
2. Tag mixing: K-statistic cut-off values of 0.1, 0.3 (diagnostic 0.2)

3. Conditional-age-at-length data weighting: 1, 0.5 (diagnostic 0.75)
4. Adult natural mortality (M and at age 40 quarter) value: Tag based = 0.062/quarter (Ducharme-Barth et al., 2026), Higher M = 0.10/quarter (diagnostic 2023: 0.078/quarter)
5. Effort creep on longline CPUE indices: ‘maximum’ scenario suggested from PAW (Hoyle, 2024) 3% per year pre 1976, 1.5% year post 1976, (diagnostic: min 1% per year pre 1976, 0.5% year post 1976).
6. Time period of the global CPUE to use for regional scaling penalty: full time period (diagnostic: 1965-1969)
7. Reporting rates during mixing period: reporting rates activated in mixing period (diagnostic: reporting rates not activated)
8. Tau - 1.2, 1.4, 1.6, 1.8 (1.0 was used in the stepwise stages up until step 19, diagnostic = 2.0)

6.2.2 Uncertainty characterisation

Typically, three types of uncertainty could be incorporated into the estimates of stock status used for management advice. One involves the statistical uncertainty of the estimates produced by individual models, often referred to as ‘estimation’ uncertainty. The second involves ‘model’ uncertainty, which is the uncertainty in the structural and fixed-parameter assumptions underpinning individual models, (e.g., fixed biological parameter such as M, steepness, tag mixing etc.). The third involves data inputs, such as alternative abundance indices with effort creep, size data treatments or other data inputs. A key uncertainty in models fitting to tag recapture data is tag mixing. Implementing the release group specific mixing periods based on the tag mixing simulations, allows a more continuous treatment of this uncertainty by integrating across a range of K-statistic cut-off values. We selected a range of K-statistic cut-off scenarios from 0.05 to 0.35, to cover assumptions where by tag mixing assumptions are mostly assigned as 4 quarters ($K=0.05$), effectively removing most of the tag data from the likelihood, to mixing assumptions that including zero or 1 quarter mixing periods ($K=0.35$), retaining much of the data in the likelihood. The extremes of the range of K-statistics applied are reasonable to assume as less plausible than the in between values and a flat unweighted treatment of this uncertainty is unrealistic. How to treat this uncertainty distribution needs further discussion among technical experts. This was not discussed at the PAW and we chose the pragmatic option of a symmetrical normal distribution centered on the diagnostic K-statistic of 0.2. The relationship between these K-statistic cut-off values and the assigned mixing periods can be viewed in Scutt Phillips et al. (2026). Finally, the sensitivity analysis showed that once τ increased beyond 1.4 the key results were insensitive, so τ above 1.4 seems somewhat redundant in further downweighting the tag data influence. However there was sensitivity between 1 and 1.4. Acknowledging that the basis for setting a τ value is not clear, other than it should be above 1, we include three values for τ in the uncertainty ensemble: 1.2, 1.3, 1.4.

The ‘uncertainty ensemble’ provides a distribution of plausible management quantities that

captures a range of structural, biological and or data uncertainties. For this assessment we have followed the Monte Carlo ensemble approach applied in the most recent skipjack (Tearns et al., 2025) and south Pacific albacore assessments (Tearns et al., 2024). The choice of uncertainties to included in the ensemble models is informed by the sensitivity analyses. For the uncertainty ensemble we include a combination of parameter distributions from which values are randomly drawn which implicitly weights those draws according to the shape of the distribution, and fixed values of alternative plausible options for other factors which are drawn at random with equal probability. 100 separate combinations of these parameters/factors are drawn and each combination forms a separate model run for the uncertainty ensemble.

The parameters and fixed factors included in the uncertainty ensemble:

1. Steepness [beta distribution]
2. Tag mixing [range of K stat values 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35 - normal distribution centred on 0.2]
3. M at age 40 quarters [beta distribution]
4. Effort creep – 5 x options
5. Reporting rates during mixing period: reporting rates activated, reporting rates not activated
6. Tau = 1.2, 1.3, 1.4

Estimation uncertainty The inclusion of estimation uncertainty is recommended to form the basis for overall uncertainty in the management quantities. Hessian-based estimates of the standard deviations for the key management quantities are obtained using the variance-covariance matrix of the model parameters and the Delta Method.

Estimation uncertainty is provided for the terminal management quantities: $SB_{recent}/SB_{F=0}$, F_{recent}/F_{MSY} , SB_{recent}/SB_{MSY} . The estimation uncertainty is combined across all 100 models in the ensemble using a parametric bootstrap similar to the approach used in the 2023 bigeye assessments (Day et al., 2023). We generate 100 random draws from normal distributions with mean and standard deviation specified for each ensemble model, in the case of $SB_{recent}/SB_{F=0}$, which was estimated on the log scale, transform the random deviates to normal space by taking their exponent. The management quantities with the full uncertainty are summarised with the mean, median, and 10, 25, 75 and 90 percentiles of the 10,000 values of each management quantity.

7 Results

The progression of model development from the 2023 diagnostic model to the 2026 diagnostic model is described in Section 6.1 and displayed in Figure 16 and Figure 17. Each step below is compared with the preceding step. These changes should be regarded as indicative because the intermediate models did not undergo the full diagnostic and tuning process. The diagnostic

model represents the culmination of a number of steps that incorporated changes from the previous assessment, recommendations from the Pre-assessment Workshop (PAW) and Scientific Committee, new data and biological information, and methodological developments. It provides a fully examined reference model for sensitivity analyses and construction of the uncertainty ensemble; management advice is based on that ensemble rather than on the diagnostic model alone.

1. [01] 2023 diagnostic-model refit: no change in $SB/SB_{F=0}$ or SB .
2. [02] New MFCL executable, and ini update (1007): very minor changes in early recruitment, no other changes $SB/SB_{F=0}$ or SB .
3. [03] Lorenzen natural-mortality scaling fixed to the 2023 diagnostic-model estimate: very minor changes in $SB/SB_{F=0}$ and SB .
4. [04] Minor update to length-weight conversion factors: very minor changes in $SB/SB_{F=0}$ and SB .
5. [05] Convert to the 5 region structure, and new fishery definitions 33-fishery structure (28 extraction, 5 index): reductions in historical levels pre-1990s, but very minor changes in terminal $SB/SB_{F=0}$ and SB .
6. [06] Size-composition inputs replaced with the reweighted weight data converted to lengths for longline fisheries, data until 2021: very minor change in $SB/SB_{F=0}$ and SB .
7. [07] Weight-as-length plus more recent observed-length compositions for longline fisheries, data until 2021: very minor change $SB/SB_{F=0}$ and SB .
8. [08] Update all data sets to 2024, except CAAL: **first notable change**, some increase (less depleted) in $SB/SB_{F=0}$ and SB , particularly since mid 1990s.
9. [09] Apply the size data truncations to the ID and PH fisheries size data as per previous assessments: no change in $SB/SB_{F=0}$ and SB .
10. [10] Switch to separate sdmTMB regional CPUE indices and apply the MFCL regional scaling penalty feature based on a global sdmTMB model, time period 1965-1969 (max spatial coverage): **second notable change**, shift to lower (more depleted) $SB/SB_{F=0}$ and SB .
11. [11] Time-varying CPUE uncertainty: very minor change in $SB/SB_{F=0}$ or SB .
12. [12] Fixed regional CPUE SDs: very minor change in $SB/SB_{F=0}$ or SB .
13. [13] Update CAAL data, raw unweighted data update: **third clear but not major change**, reduction (more depleted) in $SB/SB_{F=0}$ and SB .
14. [14a] Reweighted CAAL data based on the model region global CPUE sdmTMB relative mean density: **fourth clear but not major change**, reversed the previous step with an increase (less depleted) in $SB/SB_{F=0}$ and SB .

15. [14b] Reweighted CAAL data based on the model region global CPUE sdmTMB relative mean density, but with region 3 and 4 merged: minor reduction in (more depleted) $SB/SB_{F=0}$ and SB .
16. [15] Selectivity modifications, flexible fishery-specific selectivity plus weak F10 non-decreasing penalty
17. [16] Implement release-group-specific mixing periods, K-statistic cut-off 0.2: **fifth clear but not major change**, with lower $SB_{F=0}$ and SB , and greater depletion.
18. [17] Deactivate the reporting rate adjustments during mixing: **sixth and most notable scaling impact**, substantial increase in historical and terminal (less depleted) $SB/SB_{F=0}$ and SB .
19. [18] Add effort-creep adjustment to the regional CPUE abundance indices: increase in terminal (less depleted) $SB/SB_{F=0}$ and SB , with higher early SB and $SB_{F=0}$ and a steeper decline in SB .
20. [19] Implement Dirichlet multinomial on size data, maximum effective sample size 25: minor decrease (more depleted) in $SB/SB_{F=0}$ and SB .
21. [20] Tau (τ)=2: **seventh and largest scaling impact**, caused a reduction (more depleted) $SB/SB_{F=0}$ just below 0.2, and an approximate halving of the population scale SB .
22. [21] Add weak non-decreasing penalty on F33 (region 5 index): minor change to $SB/SB_{F=0}$ and SB .
23. [22] Steepness increased from 0.8 - 0.9, more defensible steepness for a tropical tuna, consistent with mode of ensemble distribution: minor change to $SB/SB_{F=0}$ and SB .
24. **Diagnostic 2026**

The stepwise model development can be explored in the [interactive HTML viewer](#).

Overall, only minor to moderate changes in terminal estimates, relative to the preceding step, were observed when updating the data to 2024 (less depleted and higher spawning potential), implementing the new CPUE approach (more depleted and lower spawning potential), implementing release-group-specific mixing periods (moderately more depleted and slightly lower spawning potential), applying effort creep (slightly less depleted and higher spawning potential), and implementing the Dirichlet-multinomial likelihood (slightly more depleted and slightly lower spawning potential). The increased initial biomass, steeper decline and slightly higher terminal biomass after applying effort creep were counter-intuitive. Each effort-creep scenario nevertheless represents a separate set of CPUE indices to which the model is fitted, so the starting biomass can shift. The alternatives were therefore treated as independent hypotheses, although methods that separate starting-scale effects from effort-creep effects warrant further investigation.

The most notable changes in the stepwise were related to tag data, initially, (step 17) deactivating the application of reporting rate adjustments in the mixing period. This change was as expected

and noted earlier, higher spawning potential and less depleted status if reporting rate adjustments are not made. This confirms the importance of this decision on the influence of tag data on population scaling, and it being included in the uncertainty ensemble. It should be noted that while the direction of this effect will likely be consistent across mixing period assumptions, the degree of impact will be variable depending on the mixing period applied to the individual release groups and observed tag recaptures. We also suggest both options for dealing with reporting rates in the mixing period are not ideal, and this issue requires further consideration. By far the biggest stepwise result was from increasing the fixed value of τ from 1 (Poisson) to 2 (negative binomial) (step 20) Figure 17. This resulted in the population scaling approximately halving and the terminal depletion dropping to $SB/SB_{F=0}$ just below 0.2. This major shift reflected the downweighting of the tag data in the joint likelihood, and signals that in the current assessment the tagging data has a strong influence on increasing the biomass scale and reducing the stock depletion (more optimistic stock status). This is discussed further below in relation to the likelihood profiles and sensitivity analysis results.

7.1 Fit of the diagnostic model to data sources

7.1.1 Standardized CPUE: index fisheries

The model reproduced the main trends in the standardised CPUE indices for all five regions (Figure 18). The fit was weaker in Region 1 during approximately the final 15 years, when predicted CPUE tended to be higher than most observations. In Region 3 the model tended to overpredict some early observations and underpredict some later observations. The paired fit and residual panels distinguish the observation-error intervals from parameter-estimation uncertainty.

7.1.2 Size composition data

Fishery-level fits, conditional predictive intervals and the model-implied incident effective sample sizes are shown in Figure 19. The complementary regional summary pools fitted and observed counts over fisheries and years (Figure 20). Fits reproduced the main regional modes. Region 1 was dominated by large fish, whereas Regions 2–4 contained clearer small- and large-fish modes. Region 5 had the smallest ΣESS and less information on small fish. Fishery- and year-specific selections can also be explored in the [interactive diagnostic report](#).

Longline fisheries: Fits for most longline extraction fisheries (fisheries 1–11) and longline index fisheries (fisheries 29–33) were acceptable. The Region 5 index fishery had particularly poor size-data coverage, and the model overestimated larger length classes while underestimating small-to-medium length classes. The LL.ALL extraction fishery in Region 5 had more data and a better fit.

The time series plots of observed versus predicted mean lengths for the longline fisheries were acceptable for most fisheries, particularly for those fisheries with consistent data coverage and narrower confidence intervals across time (Figure 21). The index fisheries were acceptable, although for region 1 the model struggled to predict the observed extent of the decline in mean

length, although still estimating a long-term decline.

Other fisheries: Length-composition fits were also relatively good for most non-longline and purse-seine fisheries. Time-series fits to mean length were clearest for fisheries with more observations, such as the Philippines purse-seine fishery in Region 2 and the important associated purse-seine fishery in Region 3, which accounts for a large amount of the bigeye catch. The model predicted the mean-length dynamics of this fishery relatively well (Figure 22).

7.1.3 Tagging data

The programme-specific and aggregate tag-attrition predictions fitted the broad pattern of observed recaptures relatively well (Figure 23), although the aggregate prediction overestimated some recaptures between 10 and 15 quarters at liberty. Tag-attrition predictions were closer to the observations when $\tau = 1$ (not shown), consistent with the greater likelihood weight assigned to tag data in that model.

There was notable variation in recaptures among tag-recapture groups and through time (Figure 24). Eight of the 16 plotted groups dominated the recaptures, and the diagnostic-model predictions generally matched the observations for those dominant groups.

7.1.4 Tag reporting rates

The 12 estimated MFCL reporting-rate groups are shown together in Figure 25 and their fitted values are matched to release programmes, fisheries and release groups in Table 5. Estimates varied among recapture-fishery groups and programmes. The more informative RTTP and PTTP penalties based on tag-seeding data constrained groups 7, 10, 14, 17 and 18, although groups 17 and 18 favoured higher reporting rates. Group 23 (JPTP, fisheries 12–13) reached the 0.99 upper bound because predicted recaptures remained below observations. The inactive Group 31 assignment for Index fisheries F29–F33 is required to complete the MFCL matrix but contributes no tag-recapture likelihood and is neither estimated nor plotted.

7.1.5 Conditional age-at-length

The estimated growth curve for the diagnostic model, which is informed by both the CAAL and size-composition data, is shown in Figure 13. Regional fits to the conditional-age-at-length (CAAL) data are shown in Figure 26. For length bins from around 70 to 120 cm, which contain most observations, predicted median ages generally tracked the data well. At smaller sizes the model tended to slightly overpredict median age, whereas median age was generally underpredicted from around 120 to 150 cm; predictions were closer to the observations at the largest sizes.

At the model-region level, Region 1 had a reasonable amount of data after the addition of samples from north of 10° N and showed the best fit, with model estimates tracking observed mean age at length well. Regions 2 and 4 had relatively few observations; the fit was weaker in Region 2 but reasonable in Region 4. Regions 3 and 5 had more data and, as in the aggregate

data, predictions were reasonable to approximately 120 cm, above which mean age at length tended to be underpredicted (Figure 26).

7.2 Diagnostic analysis for the diagnostic model

7.2.1 Convergence

The diagnostic model reached a maximum gradient component of 9.68×10^{-5} and had a positive-definite Hessian, with all 1,997 eigenvalues positive. MFCL flagged 38 estimated parameters as on or close to a bound (Table 6). Most were movement diffusion coefficients estimated effectively as zero. The other bound cases were six regional recruitment-distribution deviations, two selectivity spline coefficients and one tag-reporting-rate parameter. These cases did not prevent a positive-definite Hessian.

The inverse-Hessian correlation matrix contained 22 parameter pairs with $|r| > 0.90$, involving 29 unique parameters; nine pairs had $|r| > 0.95$. The nine strongest pairs are shown in Table 7. Six involved the jointly estimated mean and standard deviation of length at age, and three involved regional recruitment distribution in the same quarter. The remaining 13 correlations (0.90–0.95 in absolute value) were also confined to growth-variability parameters and regional recruitment-distribution deviations.

7.2.2 Jitter

The phase-aware jitter analysis explored a broader set of initial values than the final-MLE jitter used in earlier assessments. All 30 fits completed and 25 met the maximum-gradient-component criterion of 10^{-4} ; ten retained fits had a lower objective than the diagnostic fit. The lowest objective was 89,469.7, which was 1,345.2 units below the diagnostic objective of 90,814.9 (Figure 27). This indicates local structure in the likelihood surface. Recruitment and some regional trajectories varied among retained solutions, whereas terminal depletion and other management quantities were comparatively stable (Figure 28 and Figure 29). The result is therefore a warning about parameter identifiability and local minima, but not evidence of equivalent instability in the principal management quantities.

The complete jitter results can be viewed in the [browser report](#).

7.2.3 Self test

All 50 pseudo-data refits completed and met the recorded convergence criterion, with maximum gradient components from 5.13×10^{-5} to 9.97×10^{-5} . Median whole-series pseudo-data errors ranged from -0.23% to 0.07%, and the generating trajectory was inside the pointwise central 95% empirical refit range at 385 of 438 annual quantity–year points (87.9%; Figure 31). Median relative errors for recent and native assessment summaries ranged from -2.1% to 3.5%; across the selected fitted recruitment-scale, growth and stock–recruitment parameters they ranged from -3.85% to 2.00% (Figure 30). Consistent with the purpose of a self-test (Deroba et al., 2015; Kim et al., 2024), these results show that the configured model can recover its principal derived

quantities from simulated observations without consistent large bias. They do not test structural realism or form confidence intervals for the assessment.

The complete self-test results can be viewed in the [browser report](#).

7.2.4 Retrospectives

The retrospective analysis showed acceptable results with no strong bias for depletion (Mohn's ρ -0.195), recruitment (Mohn's ρ -0.062) and spawning potential (Mohn's ρ -0.235) all within or close to the acceptable range of ± 0.2 to ± 0.3 , fishing mortality had Mohn's ρ of 0.308, but this is not unexpected as F is very uncertain in the final years of the assessment (Figure 32).

The complete retrospective results can be viewed in the [browser report](#).

7.2.5 Likelihood profiles

The likelihood profile over mean total biomass showed conflict among data sources. CPUE, and to a lesser extent length-frequency data, favoured lower biomass, whereas tag data favoured higher biomass (Figure 33). CAAL data favoured higher biomass than CPUE and length-frequency data but had a comparatively flat profile across much of the range containing the overall minimum. The component figures show the CPUE indices (Figure 34), length-frequency contributions by region (Figure 35), CAAL contributions by region (Figure 36) and tag programmes (Figure 37). This conflict was represented in the uncertainty ensemble through alternative tag-mixing periods, reporting-rate treatment during mixing and tag overdispersion (τ).

The likelihood profiles can be explored in the [interactive HTML viewer](#).

Among tagging programmes, the tropically focused PTTP made the strongest contribution favouring higher biomass. RTTP was less influential but also tended towards higher biomass, whereas JPTP in Region 1 favoured lower biomass. The programmes are spatially separated and represent different parts of the WCPO stock. Release-group profiles did not reveal a single group driving the programme-level pattern. The static figure is therefore limited to programme level, while programme and release-group curves can be selected in the interactive viewer.

7.2.6 ASPM

The constant-recruitment age-structured production model (ASPM), conditioned on observed catches and fitted only to the five CPUE indices, provided a reduced comparison with the integrated diagnostic model (Figure 38). In 2024 the ASPM estimated $SB/SB_{F=0} = 0.480$, spawning potential of 985.6 thousand t, recruitment of 116.6 million fish and annual fishing mortality of 0.146. Corresponding diagnostic-model estimates were 0.162, 233.6 thousand t, 69.3 million fish and 0.433. The ASPM retained the broad historical decline but produced a substantially larger and less depleted terminal population. This contrast is consistent with the population-scaling conflict identified by the likelihood profiles; it does not replace the integrated-model results used to construct the uncertainty ensemble.

7.2.7 No-tagging and alternative-movement comparisons

Three additional comparisons removed tagging data and either fixed movement at the diagnostic MFCL estimates, fixed movement at SEAPODYM estimates, or used a single-region fleets-as-areas structure. The model retaining fixed diagnostic movement was closest to the diagnostic fit: average recruitment, spawning biomass and recent depletion differed by +14%, +9% and +9%, respectively. The SEAPODYM-movement and single-region models differed more in scale, but the broad depletion history and headline estimates remained reasonably consistent. These comparisons support the robustness of the main temporal signal while leaving the tagging and movement mechanisms as important sources of structural uncertainty; full methods, figures and results are provided in Section A.1.

Overall, from these diagnostics, acknowledging the tag data conflict would need to be considered as uncertainty, we were comfortable that diagnostic was satisfactory for conducting the sensitivity analysis as the basis for the uncertainty ensemble.

7.3 Population dynamics estimates

A summary of diagnostic-model dynamic spawning depletion, spawning potential, recruitment and fishing mortality relative to F_{MSY} is provided in Figure 39.

7.3.1 Selectivity

A range of selectivity curves are estimated for the different fisheries in the model and can be largely classified by gear type. The age-specific selectivity patterns are shown in Figure 40 and length-specific selectivity patterns are shown in Figure 41.

7.3.2 Movement

The model estimated very low movement rates among model regions (Figure 42). There was virtually no movement from region 1 to any other regions and very minor movement from region 2, 3 and 4 into region 1. Region 3, 4 and 5 showed the most connection with other regions. Region 2 showed some movement to region 3.

7.3.3 Natural mortality

The diagnostic model fixed the Lorenzen natural-mortality scale at $M_{40} = 0.07814$ per quarter. The longevity-based prior followed the meta-analytical approach of Hamel (2015) and the updated longevity formulation of Hamel & Cope (2022); transforming its adult-mortality median to the assessment-model M_0 scale gave $M_{40} = 0.08082$. Applying that value to the same 2026 growth schedule and Lorenzen exponent produces a curve very close to the diagnostic one, whereas the external tag estimate ($M_{40} = 0.0624$; 90% CI 0.0500–0.0749) produces lower mortality throughout the age range (Figure 15). All three curves decline steeply for young fish and approach their respective fixed scales at age 40 quarters; the diagnostic choice is represented as an uncertainty axis rather than estimated internally.

7.3.4 Maturity

Maturity at age was derived by applying the estimated 2026 growth curve to the fixed maturity-at-length relationship used in the 2023 diagnostic model. The resulting maturity-at-age curve differs slightly from that in the 2023 diagnostic model because the growth curves differ (Figure 14).

7.3.5 Growth

The diagnostic-model growth curve is shown in Figure 13. Estimated mean length was 25.68 cm at the youngest age (SD 4.03 cm) and 156.43 cm at the oldest age (SD 14.32 cm); the von Bertalanffy growth-rate parameter was $\kappa = 0.0956$.

7.3.6 Recruitment: diagnostic model

Total annual recruitment fluctuated widely but was generally between approximately 30 and 120 million individuals (Figure 43). Estimated recruitment was lower from the 1960s to the 1990s and subsequently fluctuated mostly between 40 and 120 million individuals per year. Regional patterns also varied substantially through time. Region 1 declined sharply from the 1950s to around 1980 and then stabilised, whereas Regions 2 and 3 generally increased after 1980. Regions 1 and 3 contributed the largest shares of recruitment; the seasonal regional allocation is shown in Figure 44.

The fit to the stock recruitment relationship for the 2026 diagnostic model is shown in Figure 45.

7.3.7 Spawning potential and depletion: diagnostic model

The 2026 diagnostic model estimated a steep initial decline in spawning potential from the late 1950s to the mid-1980s, followed by a more stable period to the late 1990s and a further decline to approximately 2010. Spawning potential was comparatively stable thereafter, with a small decline in the terminal year (Figure 39). The long-term pattern is consistent with the CPUE indices (Figure 18) and the catch history, which declined and then stabilised after approximately 2010 (Figure 3).

The 2026 diagnostic model shows a decline in overall spawning potential depletion, $SB/SB_{F=0}$ through to the mid 1960s when there is a more stable period, followed by decline through to about 2010, after which it levels off just below 0.2 Figure 39. As with spawning potential, the dynamic trend mirrors the catch history, with the more recent stability reflecting more stable average catches.

The relative levels of total biomass, spawning potential and recruitment among the model regions for the diagnostic model is shown in Figure 46.

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

The regional total-biomass comparison under the fitted fishing history and the dynamic no-fishing counterfactual shows a distinctive pattern in Region 1 (Figure 47). Both trajectories declined through most of the assessment, whereas no-fishing biomass generally increased over the

later period in Regions 2–5. This contrast is considered with the regional length-frequency and recruitment information in Section 8; it should not be interpreted as a direct test of causation.

7.3.9 Fishing mortality: diagnostic model

Overall fishing mortality (F) increases since 1970s before stabilising with interannual fluctuations between approximately 0.3 to 0.6 per year since 2010 Figure 39.

Differences over time in the estimated adult and juvenile fishing mortality are shown in Figure 48, with a gradual increase in overall adult fishing mortality up until around 2000, with relatively stable adult mortality since then. The uptick in the final year for adult mortality likely relates to the increase in catch attributed to the large fish handline fishery due to Indonesia’s change in how they allocate catch between the small fish (part of other gears domestic fishery) and large fish handline, which requires an historical adjustment based on Indonesia’s advice before the next bigeye assessment. Fishing mortality of juveniles starts to increase above that of adults from around 1980 with the expansion of the purse seine associated catch fisheries, and continues above adult mortality stabilising from 2010, with a recent peak in 2022-2023. The recent peak appears related to increased juvenile fishing mortality by the fisheries operating in regions 2 and 3, which along with region 4 are the highest contributors to both adult and juvenile fishing mortality Figure 48.

7.4 Sensitivity analyses

The results of all sensitivity analysis can be explored in the [interactive HTML viewer](#). All the sensitivity models achieved a suitable gradient and had positive definite Hessian status.

7.4.1 Steepness

The sensitivity on steepness had some impacts on $SB/SB_{F=0}$, with more depleted stock when applying lower steepness values Figure 49.

7.4.2 Natural mortality

The sensitivity on M at age 40 quarters had notable impacts $SB/SB_{F=0}$, with more depleted (pessimistic) stock status when applying the diagnostic value, and less depleted (more optimistic) status when applying both the higher and the lower values Figure 50. The more optimistic depletion status and notably higher spawning potential scale for the lowest value was unexpected and requires further consideration. The highest M value produced the highest recruitment and lowest fishing mortality mortality.

7.4.3 CAAL data weighting

The sensitivity on the CAAL data weighting had minor effects on the terminal $SB/SB_{F=0}$, and SB . The highest downweighting of 0.5 produced a more optimistic depletion status, higher recruitment, and also higher historic SB (Figure 51).

7.4.4 Effort creep on CPUE

The sensitivity on the level of effort creep had no impact on the terminal levels of $SB/SB_{F=0}$, SB , fishing mortality or recruitment. The higher effort creep scenario produced higher initial SB and recruitment and steeper declining trends (Figure 52).

7.4.5 CPUE time period for regional scaling

The choice of the global-CPUE period used for the regional scaling penalty affected $SB/SB_{F=0}$, SB , fishing mortality and recruitment. Applying the full CPUE series rather than the shorter period of maximum spatial coverage used in the diagnostic model (1965–1969, as recommended by PAW) produced a higher population scale, lower fishing mortality, higher recruitment and less depletion (Figure 53).

7.4.6 Tag mixing

The sensitivity on the choice of the K-statistic for assigning release specific tag mixing periods was impactful on $SB/SB_{F=0}$, SB , recruitment and fishing mortality. The lower value of 0.1 which results in more conservative longer tag mixing periods produced higher SB scale and a less depleted, more optimistic $SB/SB_{F=0}$, also higher recruitment and lower fishing mortality (Figure 54).

7.4.7 Reporting rates adjustment in the mixing periods

Activating or deactivating tag-reporting-rate adjustments during the mixing period had little effect on $SB/SB_{F=0}$, SB , recruitment or fishing mortality. This was unexpected and may reflect the downweighting of tag data when $\tau = 2$; the stepwise analysis showed a larger effect when $\tau = 1$ (Figure 55).

7.4.8 Tau overdispersion parameter

The sensitivity on the choice of the overdispersion parameter (τ) was impactful on $SB/SB_{F=0}$, SB , recruitment and fishing mortality when it was set below 1.4, settings above 1.4 were very similar. At lower than 1.4, SB scale increased, $SB/SB_{F=0}$ increased to more optimistic levels, recruitment increased and fishing mortality decreased (Figure 56).

7.5 Stock assessment results

7.5.1 Uncertainty ensemble

Of the 100 planned ensemble configurations, 80 met the maximum-gradient-component (MGC) criterion $\leq 1.0 \times 10^{-4}$, of which 62 had positive-definite Hessians. The 80 models were retained in the ensemble for management advice; ten did not meet the MGC criterion and ten were incomplete. For the 62 positive-definite Hessian fits, 100 correlated inverse-Hessian draws were propagated jointly by the multivariate delta method; 18 Near-PDH fits contribute structural uncertainty only. Estimation uncertainty is therefore incorporated into the statistics for management advice, but is not complete for all models.

The retained models and uncertainty results can be explored in the [interactive HTML viewer](#).

The planned and retained model numbers by uncertainty category are shown in Figure 57. More models were retained at the two higher τ values, which assign less weight to tag data, and when reporting-rate adjustments were excluded during the mixing period. The balanced design contained 50 models per reporting-rate treatment, but the retained set contained 34 inclusion and 46 exclusion models. Under inclusion, MFCL reconstructs pre-mixing tag catches by dividing by the estimated reporting rate; values approaching zero can create very large intermediate values and numerical overflow (Davies et al., 2026b). The imbalance therefore reflects differential numerical stability, not greater scientific support for exclusion. Every retained model received equal model weight, and the unequal retention is reported explicitly rather than used as an additional weighting criterion.

The influence of each uncertainty component on the management quantities across the ensemble of models is displayed in Figure 58, Figure 59 and Figure 60. Natural mortality and tag mixing were most influential on the spread of the ensemble.

The trajectories of $SB/SB_{F=0}$, spawning potential, fishing mortality and recruitment for each of the 80 models are displayed in Figure 61. Fifteen central model estimates had recent $SB/SB_{F=0} < 0.2$. Ribbon plots combining structural uncertainty and the available estimation uncertainty are provided in Figure 62.

7.5.2 Analyses of stock status

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

We do not present every ancillary result for every model. This section relies mainly on the equal-model-weight ensemble summaries in Table 8, Table 9, Table 10, Table 11 and Table 12, the time-series uncertainty in Figure 62, and the dynamic Majuro and Kobe plots in Figure 64.

7.5.3 Majuro and Kobe plots and comparisons with reference points

Most of the combined uncertainty distribution lies in the Kobe and Majuro areas corresponding to a stock that is not overfished and not undergoing overfishing, although the distribution extends into less favourable categories (Figure 63). The equal-model-weight mixture assigns an 18.6% probability to recent depletion below the LRP, a 17.3% probability to $SB_{recent}/SB_{MSY} < 1$, and a 24.1% probability to $F_{recent}/F_{MSY} > 1$ (Table 10). The category probabilities are printed directly within the corresponding plot areas to make the uncertainty allocation explicit, following the general panel-percentage presentation used by Kim & Neubauer (2025); that study is cited for the visualisation approach, not as evidence for the tuna assessment results.

CMM 2025-02 contains an objective to maintain the spawning biomass depletion ratio at or above the average $SB/SB_{F=0}$ for 2012–2015. Across the 80 ensemble models, excluding estimation uncertainty, this average was 28.8%. The corresponding average for 2021–2024 was 27.8%, giving a recent-to-2012–2015 spawning depletion ratio of 96.7%; stock status was therefore slightly below

the objective (Table 9). This comparison is distinct from the LRP and MSY-based quantities.

The dynamic Majuro and Kobe trajectories for both the diagnostic model and the annual ensemble median show declining stock status and increasing fishing mortality since the 1950s, with fishing mortality increasing more strongly from the late 1960s (Figure 64). The ensemble median remains more optimistic than the diagnostic model, consistent with the diagnostic model being near the pessimistic edge of the ensemble.

7.5.4 Fishing impact and time-dynamic MSY

Dynamic MSY declined from relatively high values early in the assessment as the fishery expanded and the age composition of removals changed, and was broadly between approximately 120 and 150 thousand t in recent decades (Figure 65). Recent catches were generally near or below the corresponding dynamic-MSY series, although both quantities vary annually and this comparison is not a replacement for the ensemble reference-point results.

Estimated fishery impact increased in all regions, with the greatest terminal impact in Region 2 and substantial impacts in Regions 3 and 4 (Figure 66). At the stock-wide scale, associated purse-seine fisheries made the largest single contribution to the increase since the 1990s; longline and miscellaneous fisheries also contributed materially. The attribution depends on the diagnostic-model spatial and fishery structure and should be interpreted as a decomposition of that model's estimated impact.

7.5.5 Yield analysis

The diagnostic-model equilibrium yield curve was relatively flat near its maximum (Figure 67). The fishing-mortality multiplier associated with recent fishing was close to, but slightly above, the multiplier producing the maximum equilibrium yield, while the corresponding equilibrium yield remained close to MSY. Management quantities derived model by model from the yield analysis are summarised in Table 8 and Table 11; the ensemble median F_{recent}/F_{MSY} was 0.838 after incorporating structural and available estimation uncertainty.

7.5.6 Catch-based projections

Under fishery- and quarter-specific catches fixed at their 2022–2024 means, the ensemble projected recovery in both the LRP depletion metric and spawning potential after 2024 (Figure 68). Across the 800 model–recruitment realisations, the median LRP metric increased to 0.355 in 2030, 0.517 in 2040 and 0.545 in 2054; 5.2% of realisations were below the 0.20 LRP in 2030 and none were below it in 2040 or 2054 (Table 13). Median spawning potential increased from 523.1 thousand t in 2030 to 720.6 thousand t in 2054. These are conditional projections under the specified recent-catch scenario, not forecasts of future catches or management action, and their intervals include recruitment variability but not Hessian parameter uncertainty.

8 Discussion and conclusions

8.1 Stock Status from the uncertainty grid

The 2026 assessment is a substantial revision of the WCPO bigeye tuna assessment while retaining continuity with the 2023 analysis. Moving from nine to five regions, revising fishery definitions and fixing the Lorenzen natural-mortality scale reduced the number of estimated parameters from 3,067 to 1,997, a reduction of 1,070 (34.9%). This materially improved the instability seen in the 2023 jitter analysis. Nevertheless, the 2026 diagnostics show that a smaller model is not necessarily a fully identified model.

8.2 Diagnostic interpretation

- **Optimisation and practical estimability.** This was the first assessment in the series to use phase-aware jittering from the original fitting path rather than perturbing only the final MLE. The lower-objective solutions show that recruitment deviations, regional recruitment allocation and some movement parameters remain weakly determined. However, terminal depletion and the principal management quantities changed much less than the objective function or individual parameters. These parameters are therefore important diagnostics of over-parameterisation, but their instability did not translate directly into similar instability in management advice.
- **Regional recruitment information.** Regional recruitment deviations, particularly in Region 5, were weakly determined where length-frequency data provided limited recruitment information. A related issue is evident in Region 1: its length compositions contain relatively few small fish, and both fished and unfished biomass decline almost monotonically, unlike the later increases in unfished biomass in other regions. This suggests that the model may assign substantial recruitment information to the appearance of small length classes. Because regional recruitment is allocated jointly, weak information in one region can affect estimates elsewhere. This pattern warrants targeted simulation and data-conflict analysis rather than a causal interpretation of the plotted trajectories.
- **Length-composition weighting.** The selected DM setting balanced a small improvement in length-frequency fit against the deterioration in CPUE fit observed from $N_{\max} \geq 40$. Many incident compositions nevertheless approached the $N_{\max} = 25$ asymptote. This is plausible when one group-level parameterisation must represent fisheries and samples that remain heterogeneous through space and time: no single shared relationship can reproduce the information content of every sample. The boundary pattern therefore should not be interpreted as evidence that every composition contains exactly 25 effective observations. The Francis effective-sample-size distribution supplied an empirical scale for the cap, while the integrated fit determined the final choice.
- **Self-test.** The first formal self-test in this assessment series recovered the principal time series and management quantities with small median error, despite weak practical estimation of some recruitment and movement parameters. This is consistent with the

jitter result: not all individual parameters were well determined, but their net effect on the main derived quantities was limited. The self-test establishes internal consistency under the configured observation models, not realism of the operating model.

- **Tagging and population scale.** Population scaling remained the dominant data conflict. CPUE and length-frequency data favoured a lower biomass scale, whereas tagging data, particularly PTTP, favoured a higher scale. With near-Poisson tag variance ($\tau \approx 1$), tagging observations received greater weight and results approached the more optimistic 2023 assessment; allowing overdispersion with $\tau = 2$ reduced that weight and allowed the other likelihood components to support a lower population scale. Changing the K-statistic cutoff similarly changes the number of post-mixing recaptures and hence the effective tag weight. Reporting-rate Group 23 reached its 0.99 upper bound because predicted recaptures remained below observations. Together, these results point to limitations in the current treatment of tag mixing, overdispersion and reporting rather than to a single problematic release group.
- **Alternative movement and no-tagging models.** Removing tagging data while fixing movement at either MFCL or SEAPODYM estimates, and fitting a single-region model without tags, produced reasonably consistent headline trajectories (Section A.1). The closest comparison was the no-tagging model with diagnostic MFCL movement fixed. This is reassuring about the broad assessment signal, but it does not resolve whether tagging primarily affects scale directly or through estimated movement.

The ASPM comparison produced a larger and less depleted terminal population than the integrated diagnostic model, reinforcing the same population-scaling uncertainty. These diagnostics do not identify a single preferred scale; they explain why tag mixing, reporting-rate treatment, tag overdispersion, movement and natural mortality were represented in the ensemble. The diagnostic model was near the pessimistic edge of the retained ensemble, whereas management quantities were calculated from an equal-model-weight mixture that incorporated available Hessian estimation uncertainty.

The ensemble median $SB_{recent}/SB_{F=0}$ was 0.277, with a central 80% interval of 0.172–0.395, and the probability of this LRP quantity being below 0.20 was 18.6% (Table 8 and Table 10). Median F_{recent}/F_{MSY} was 0.838, although 24.1% of the uncertainty distribution exceeded one. For the separate CMM 2025-02 objective, mean spawning depletion was 27.8% in 2021–2024 and 28.8% in 2012–2015 across the 80 central models, excluding estimation uncertainty, giving a recent-to-2012–2015 spawning depletion ratio of 96.7% (Table 9). Stock status was therefore slightly below this objective. The central assessment outcome is neither below the adopted LRP nor above F_{MSY} , while the uncertainty around those benchmarks remains important for management interpretation.

The constant-recent-catch projections provide a conditional view of future dynamics rather than a forecast. Under catches fixed at their 2022–2024 fishery- and quarter-specific means, the median LRP depletion metric increased to 0.355 in 2030 and 0.517 in 2040 (Table 13). This result depends on the assumed catches, model ensemble and resampled recruitment process; it

does not account for implementation uncertainty, future changes in fishing patterns or Hessian parameter draws.

8.3 Recommendations for further work

The assessment identifies the following priorities:

- Develop and simulation-test a tag-recapture likelihood that treats overdispersion, mixing, spatial heterogeneity and reporting processes coherently; assess reporting rates during mixing rather than choosing between two imperfect switches.
- Investigate the lower-objective jitter solutions and weakly determined recruitment and movement parameters, including whether recruitment deviations can be represented as random effects in next-generation software. MFCL's large number of fixed-effect recruitment deviations remains a recurring source of over-parameterisation.
- Test the sensitivity of regional recruitment to the amount and size coverage of length-frequency data, especially in Regions 1 and 5, and examine how information is transferred through the joint regional-allocation process.
- Re-evaluate the eight DM groups and $N_{\max}$ with composition-level predictive checks or finer hierarchical structure, so that persistent spatial and temporal heterogeneity is not forced into a single group-level relationship.
- Continue evaluation of the regional CPUE-scaling penalty and its sensitivity to the global-index period, and reconstruct the historical Indonesian handline catch allocation before the next assessment.
- Extend self-tests to structural cross-tests and the tag-likelihood alternatives, and repeat fixed-movement likelihood profiles to distinguish the direct scaling effect of tags from their indirect effect through movement.
- Establish a separate MFCL continuous-integration project that exercises assessment-critical flags and their interactions on small verified models, including τ , pre-mixing reporting-rate treatment and mixing-period controls. Automated checks of expected likelihood components, fitted outputs and numerical failures would reduce the risk that complex or poorly documented settings affect an assessment unnoticed.

The general conclusions of this assessment are as follows:

- The spawning potential of the stock has become more depleted across all model regions until around 2010, after which it has become more stable.
- The diagnostic model in this assessment is one of the most pessimistic models of the uncertainty ensemble used for management advice.
- Average fishing mortality rates for juvenile and adult age-classes have increased throughout the period of the assessment until around 2000, after which they have stabilised, but with

high inter-annual variability for juveniles.

- Juveniles have experienced considerably higher recent fishing mortality than adults.
- Median recent depletion (2021–2024; $SB_{recent}/SB_{F=0}$) is 0.277 (central 80% interval including structural and available estimation uncertainty: 0.172–0.395; central 95% interval: 0.136–0.442).
- Fifteen of the 80 central model estimates had $SB_{recent}/SB_{F=0} < 0.2$; after including available estimation uncertainty, the probability below the LRP is 18.6%.
- CMM 2025-02 contains an objective to maintain spawning depletion at or above its 2012–2015 average. Across the 80 central models, excluding estimation uncertainty, mean depletion was 27.8% in 2021–2024 and 28.8% in 2012–2015, giving a recent-to-2012–2015 spawning depletion ratio of 96.7%; stock status was therefore slightly below the objective (Table 9).
- Median recent fishing mortality (2020–2023), F_{recent}/F_{MSY} , is 0.838 (central 80% interval: 0.592–1.206; central 95% interval: 0.460–1.351).
- Although the stock is assessed as more depleted and subject to higher fishing mortality than in 2023, the ensemble indicates that it is not overfished under the adopted LRP criterion and is not undergoing overfishing based on median $F_{recent}/F_{MSY} < 1$.
- Overall, the 2026 assessment presents a more pessimistic stock status than the 2023 assessment.

9 Acknowledgements

We thank WCPFC members, participating territories and other providers of fisheries data, and the observers, port samplers, tagging teams and fishers who collected the biological and tagging observations. We also acknowledge the 2026 Pre-assessment Workshop participants; the Pacific Community data, assessment and computing-support teams; and the SPC sclerochronology laboratory team, which aged many additional otoliths for this assessment. We thank the Japan Fisheries Research and Education Agency (FRA) for providing additional age–length data from the North Pacific. This assessment was supported through the WCPFC regular Scientific Services and Additional Resourcing to SPC budget lines, including funding for the additional stock-assessment scientist who led this assessment. CPUE data preparation was also supported by the European Union’s Pacific–European Union Marine Partnership Programme Phase 2. K. Kim especially thanks his family for their understanding of the long hours required to complete the assessment.

10 Statement on the use of generative AI

GitHub Copilot, a generative artificial-intelligence assistant, was used to support code development and checking and the preparation of selected visualisations for this report.

11 Tables

Table 1: Public assessment outputs, source repositories and pinned build resource.

Resource	Browser output	Source repository
Assessment report	Clean PDF	ofp-sam-bet-2026-report
Stepwise development	Interactive model viewer	ofp-sam-bet-2026-stepwise
Diagnostic fit and profiles	Diagnostic report; Likelihood-profile viewer	ofp-sam-bet-2026-diagnostic
Jitter	Jitter report	ofp-sam-bet-2026-jitter
Self-test	Self-test report	ofp-sam-bet-2026-selftest
Retrospective analysis	Retrospective report	ofp-sam-bet-2026-retrospective
One-off sensitivities	Sensitivity viewer	ofp-sam-bet-2026-sensitivity
Uncertainty ensemble and projections	Ensemble viewer	ofp-sam-bet-2026-ensemble
Pinned build environment	<code>tuna-flow-private:v2.7</code>	Report workflow and full digest (<code>sha256:4fee4c40...9c8852</code>)

Table 2: Definition of fisheries for the 2026 bigeye tuna stock assessment in the western and central Pacific Ocean. The tag-recapture group identifies the recapture-fishery aggregation, whereas the selectivity group identifies the fitted selectivity curve.

Fishery	Name	Region	Data group	Tag-recapture group	Selectivity group	Selectivity form	Nodes
1	01.LL.WEST.1	1	LL	1	1	Cubic spline	5
2	02.LL.EAST.1	1	LL	2	2	Cubic spline	5
3	03.LL.US.1	1	LL	3	3	Cubic spline	5
4	04.LL.ALL.2	2	LL	4	4	Cubic spline	5
5	05.LL.OS.2	2	LL	5	5	Cubic spline	5
6	06.LL.ARCH.3	3	LL	6	6	Cubic spline	5
7	07.LL.WEST.3	3	LL	7	7	Cubic spline	5
8	08.LL.EAST.4	4	LL	8	8	Cubic spline	5
9	09.LL.OS.3	3	LL	9	9	Cubic spline	5
10	10.LL.ALL.5	5	LL	10	10	Cubic spline	5
11	11.LL.AU.5	5	LL	11	11	Cubic spline	5
12	12.PS.JP.1	1	PS	12	12	Cubic spline	5
13	13.PL.JP.1	1	PL	13	13	Cubic spline	5
14	14.HL.ID.2	2	HL	14	14	Cubic spline	5
15	15.HL.PH.2	2	HL	14	15	Cubic spline	5
16	16.PL.ALL.2	2	PL	15	16	Cubic spline	5
17	17.PS.ID.2	2	PS	14	17	Cubic spline	5
18	18.PS.PH.2	2	PS	14	18	Cubic spline	5
19	19.PS.ASS.2	2	PS ASS	16	19	Cubic spline	5
20	20.PS.UNA.2	2	PS UNASS	16	20	Cubic spline	5
21	21.DOM.ID.2	2	DOM	14	21	Cubic spline	5
22	22.DOM.PH.2	2	DOM	14	22	Cubic spline	5
23	23.DOM.VN.2	2	DOM	17	23	Cubic spline	5
24	24.PL.ALL.WEST.3	3	PL	18	24	Cubic spline	5
25	25.PS.ASS.WEST.3	3	PS ASS	19	25	Cubic spline	7
26	26.PS.ASS.EAST.4	4	PS ASS	20	26	Cubic spline	7
27	27.PS.UNA.WEST.3	3	PS UNASS	19	27	Cubic spline	5
28	28.PS.UNA.EAST.4	4	PS UNASS	20	28	Cubic spline	5
29	29.Index R1	1	Index	21	29	Cubic spline	5
30	30.Index R2	2	Index	21	30	Cubic spline	5
31	31.Index R3	3	Index	21	31	Cubic spline	5
32	32.Index R4	4	Index	21	32	Cubic spline	5
33	33.Index R5	5	Index	21	33	Cubic spline	5

Table 3: Dirichlet–multinomial configuration used for length-composition weighting in the 2026 diagnostic model. $N_{\max}$ is an asymptotic upper bound on incident effective sample size, not a fixed weight assigned to each composition.

Component	Diagnostic-model setting	Interpretation
Likelihood	DM without random effects (parest flag 141 = 11)	Allows extra-multinomial variation in length compositions.
Fishery grouping	Eight groups covering all 33 fisheries (fish flag 68; Table 4)	Shares information among fisheries with similar gear or data roles.
Concentration intercept	fish_pars (22) fixed at 7 (fish flag 69 = 0)	The intercept had reached its upper bound in preliminary estimation and was not re-estimated.
Relative-sample-size exponent	Group-specific fish_pars (23) estimated from Phase 2 (fish flag 89 = 1)	Allows effective composition information to change with relative observed sample size within each group.
ESS upper asymptote	$N_{\max} = 25$ (parest flag 342; MFCL default = 1,000)	Selected just above the Francis ESS 95th-percentile range of 22.22–23.81 across 2,399 positive compositions and to retain acceptable CPUE fit.
Length-bin support	Minimum span of five positive bins (parest flag 320 = 5); parest flag 313 = 0	Uses the DM support rule rather than the robust-normal tail control.

Table 4: Eight fishery groups used to estimate group-specific relative-sample-size effects in the Dirichlet–multinomial length-composition likelihood. Fishery numbers and series names follow the fishery-definition table.

DM group	Fisheries	Fishery series	Grouping basis
G1	F1–F4, F6–F8, F10–F11	LL.WEST.1; LL.EAST.1; LL.US.1; LL.ALL.2; LL.ARCH.3; LL.WEST.3; LL.EAST.4; LL.ALL.5; LL.AU.5	Main longline composition process
G2	F5, F9	LL.OS.2; LL.OS.3	Offshore longline series with a distinct sampling history
G3	F12, F17–F18	PS.JP.1; PS.ID.2; PS.PH.2	Purse-seine series without associated/unassociated set-type separation
G4	F19, F25–F26	PS.ASS.2; PS.ASS.WEST.3; PS.ASS.EAST.4	Associated purse-seine series
G5	F20, F27–F28	PS.UNA.2; PS.UNA.WEST.3; PS.UNA.EAST.4	Unassociated purse-seine series
G6	F14–F15	HL.ID.2; HL.PH.2	Handline series
G7	F13, F16, F21–F24	PL.JP.1; PL.ALL.2; DOM.ID.2; DOM.PH.2; DOM.VN.2; PL.ALL.WEST.3	Other extraction fisheries pooled for stable estimation
G8	F29–F33	Index R1; Index R2; Index R3; Index R4; Index R5	Regional indices sharing the relative-abundance reweighting procedure

Table 5: Tag-reporting-rate groups, release coverage, estimation status and fitted reporting rates. Group is the MFCL reporting-rate group number. The pooled-matrix-row column indicates whether the same group number is also assigned to MFCL’s extra pooled row; pooled is not a tagging programme. Group 31 is assigned to Index fisheries F29–F33 to complete the matrix, but is an inactive technical placeholder with no tag-recapture likelihood and is omitted.

Group	Tag programmes	Pooled matrix row	Fisheries	Release groups	Release years	Estimated	Fitted RR
1	RTTP/PTTP	Yes	1–2, 4–10	1–61	1989–1992, 2006–2013, 2015–2021	Yes	0.258
2	RTTP/PTTP	Yes	3	1–61	1989–1992, 2006–2013, 2015–2021	No	–
3	RTTP/PTTP	Yes	11	1–61	1989–1992, 2006–2013, 2015–2021	No	–
4	RTTP	No	12	1–15	1989–1992	No	–
5	RTTP	No	13	1–15	1989–1992	No	–
6	RTTP	No	14–18, 21–22	1–15	1989–1992	Yes	0.637
7	RTTP	No	19–20	1–15	1989–1992	Yes	0.625
8	RTTP	No	23	1–15	1989–1992	No	–
9	RTTP	No	24	1–15	1989–1992	No	–
10	RTTP	No	25, 27	1–15	1989–1992	Yes	0.567
11	RTTP	No	26, 28	1–15	1989–1992	No	–
12	PTTP	Yes	12–13	16–61	2006–2013, 2015–2021	No	–
13	PTTP	Yes	14–18, 21–22	16–61	2006–2013, 2015–2021	Yes	0.461
14	PTTP	Yes	19–20	16–61	2006–2013, 2015–2021	Yes	0.503
15	PTTP	Yes	23	16–61	2006–2013, 2015–2021	No	–
16	PTTP	Yes	24	16–61	2006–2013, 2015–2021	No	–
17	PTTP	Yes	25, 27	16–61	2006–2013, 2015–2021	Yes	0.732
18	PTTP	Yes	26, 28	16–61	2006–2013, 2015–2021	Yes	0.843
19	JPTP	No	1	62–98	2000–2010, 2012, 2014, 2017	Yes	0.665
20	JPTP	No	2, 4–10	62–98	2000–2010, 2012, 2014, 2017	Yes	0.159
21	JPTP	No	3	62–98	2000–2010, 2012, 2014, 2017	No	–
22	JPTP	No	11	62–98	2000–2010, 2012, 2014, 2017	No	–
23	JPTP	No	12–13	62–98	2000–2010, 2012, 2014, 2017	Yes	0.990
24	JPTP	No	14–15, 17–18, 21–22	62–98	2000–2010, 2012, 2014, 2017	No	–
25	JPTP	No	16	62–98	2000–2010, 2012, 2014, 2017	No	–
26	JPTP	No	19–20	62–98	2000–2010, 2012, 2014, 2017	No	–
27	JPTP	No	23	62–98	2000–2010, 2012, 2014, 2017	No	–
28	JPTP	No	24	62–98	2000–2010, 2012, 2014, 2017	No	–
29	JPTP	No	25, 27	62–98	2000–2010, 2012, 2014, 2017	Yes	0.173
30	JPTP	No	26, 28	62–98	2000–2010, 2012, 2014, 2017	No	–

Table 6: Estimated parameters that MFCL flagged as on or close to a bound in the diagnostic model. Parameter names are descriptive rather than MFCL internal variable names.

Parameter type	Bound location(s)	Number
Movement diffusion coefficient	Lower bound (0); 29 season–movement-path coefficients	29
Regional recruitment-distribution deviation	Lower bound (-3); Region 2 in 1980 Q1, 1982 Q1–Q2, 1984 Q1 and Q4, and 1985 Q1	6
Fishery selectivity spline coefficient	Lower bound (-20); spline node 5 for fishery 16 (PL.ALL.2) and fishery 24 (PL.ALL.WEST.3)	2
Tag reporting rate	Upper bound (0.99); reporting-rate group 23 (JPTP, fisheries 12–13)	1

Table 7: Strongest estimated-parameter correlations in the diagnostic model ($|r| > 0.95$), calculated from the inverse Hessian.

Parameter 1	Parameter 2	Correlation
Mean length of oldest age class	von Bertalanffy growth-rate parameter (κ)	-0.9997
Mean length of youngest age class	von Bertalanffy growth-rate parameter (κ)	-0.9976
Mean length of youngest age class	Mean length of oldest age class	0.9965
Mean length of oldest age class	Age-dependency parameter for the standard deviation of length at age	-0.9868
von Bertalanffy growth-rate parameter (κ)	Age-dependency parameter for the standard deviation of length at age	0.9868
Mean length of youngest age class	Age-dependency parameter for the standard deviation of length at age	-0.9856
Regional recruitment-distribution deviation, Region 1 (2024 Q3)	Regional recruitment-distribution deviation, Region 5 (2024 Q3)	-0.9773
Regional recruitment-distribution deviation, Region 1 (2024 Q1)	Regional recruitment-distribution deviation, Region 5 (2024 Q1)	-0.9611
Regional recruitment-distribution deviation, Region 4 (1965 Q4)	Regional recruitment-distribution deviation, Region 5 (1965 Q4)	-0.9603

Table 8: Management quantities from the equal-model-weight structural mixture with available Hessian estimation uncertainty. Each of 62 PDH models contributes 100 joint Hessian draws; each of 18 Near-PDH models contributes its central estimate with equal total model weight. Values are the median and central 50%, 80% and 95% equal-tailed intervals. Estimation uncertainty is unavailable for the Near-PDH fits.

Quantity	Period	Median	50% interval	80% interval	95% interval
$SB_{\text{recent}}/SB_{F=0}$	2021–2024 / 2014–2023	0.277	0.217–0.336	0.172–0.395	0.136–0.442
$SB_{\text{recent}}/SB_{\text{MSY}}$	2021–2024 / equilibrium SB_{MSY}	1.532	1.136–2.024	0.822–2.572	0.635–3.519
$F_{\text{recent}}/F_{\text{MSY}}$	2020–2023 / equilibrium F_{MSY}	0.838	0.702–0.983	0.592–1.206	0.460–1.351

Table 9: Recent spawning depletion relative to the 2012–2015 average across the 80 equal-weight central models, excluding estimation uncertainty. For each year, depletion is spawning biomass divided by mean unfished spawning biomass over the preceding ten years. The period values are arithmetic means across models; the final row is their ratio.

Quantity	Value
D_{recent} (2021–2024)	0.278
$D_{2012-2015}$	0.288
$D_{\text{recent}}/D_{2012-2015}$	0.967

Table 10: Stock status probabilities in relation to the adopted limit and MSY-based reference points.

Criterion	Probability
$SB_{\text{recent}}/SB_{F=0} < 0.20$	18.6%
$SB_{\text{recent}}/SB_{\text{MSY}} < 1$	17.3%
$F_{\text{recent}}/F_{\text{MSY}} > 1$	24.1%

Table 11: Reference-point quantities across the 80 retained central model estimates, with equal weight. All ratios and reference points are calculated within each model before summarizing; periods and units are listed in the table. Intervals describe structural uncertainty only. Quantities without propagated estimation uncertainty are omitted.

Quantity	Period	Unit	Min	10%	Median	Mean	90%	Max
F_{mult}	2020–2023	multiplier	0.695	0.837	1.192	1.253	1.682	3.026
$F_{\text{recent}}/F_{\text{MSY}}$	2020–2023	ratio	0.330	0.594	0.839	0.860	1.195	1.440
$SB_{F=0}$	2014–2023 mean	thousand MT	874.7	969.7	1625.0	1851.3	3222.4	3990.9
$SB_{\text{latest}}/SB_{F=0}$	2024 / 2014–2023	ratio	0.113	0.151	0.242	0.248	0.350	0.428
$SB_{\text{latest}}/SB_{\text{MSY}}$	2024 / equilibrium SB_{MSY}	ratio	0.657	0.842	1.348	1.395	1.994	2.477
$SB_{\text{recent}}/SB_{F=0}$	2021–2024 / 2014–2023	ratio	0.126	0.175	0.278	0.280	0.384	0.503
$SB_{\text{recent}}/SB_{\text{MSY}}$	2021–2024 / equilibrium SB_{MSY}	ratio	0.737	0.997	1.507	1.581	2.236	2.882
SB_{MSY}	Equilibrium	thousand MT	112.6	152.6	289.0	346.1	661.8	1019.0
$SB_{\text{MSY}}/SB_{F=0}$	Equilibrium / 2014–2023	ratio	0.128	0.156	0.176	0.179	0.200	0.255

Table 12: Estimation-uncertainty sensitivity and Monte Carlo results for the three core stock status quantities. The all-model column gives 62 retained PDH models with 100 joint draws plus 18 retained Near-PDH point masses at equal model weight; the PDH-only column excludes the point masses. The final column is the largest absolute change in q10, median or q90 when the first 50 rather than all 100 draws per model are used, with matching Near-PDH point-mass replication.

Quantity	All models	PDH only	Maximum 50–100 difference
$SB_{\text{recent}}/SB_{F=0}$	0.277 (0.172–0.395)	0.275 (0.173–0.393)	0.0005
$SB_{\text{recent}}/SB_{\text{MSY}}$	1.532 (0.822–2.572)	1.522 (0.828–2.752)	0.0396
$F_{\text{recent}}/F_{\text{MSY}}$	0.838 (0.592–1.206)	0.838 (0.577–1.201)	0.0034

Table 13: Projected management quantities at selected years across 80 models and ten recruitment sequences per model. The LRP metric is the four-year mean spawning biomass divided by the preceding ten-year mean unfished spawning biomass; the LRP is 0.20. Catches are fixed by fishery to their 2022–2024 means, and intervals exclude Hessian parameter uncertainty.

Year	<i>D</i> 10%	<i>D</i> median	<i>D</i> 90%	Below LRP	Spawning median (10 ³ MT)	<i>C</i> /MSY 10%	<i>C</i> /MSY median	<i>C</i> /MSY 90%
2030	0.225	0.355	0.514	5.2%	523.1	0.638	0.811	1.003
2040	0.409	0.517	0.628	0.0%	690.5	0.642	0.842	1.052
2054	0.455	0.545	0.641	0.0%	720.6	0.648	0.847	1.053

12 Figures

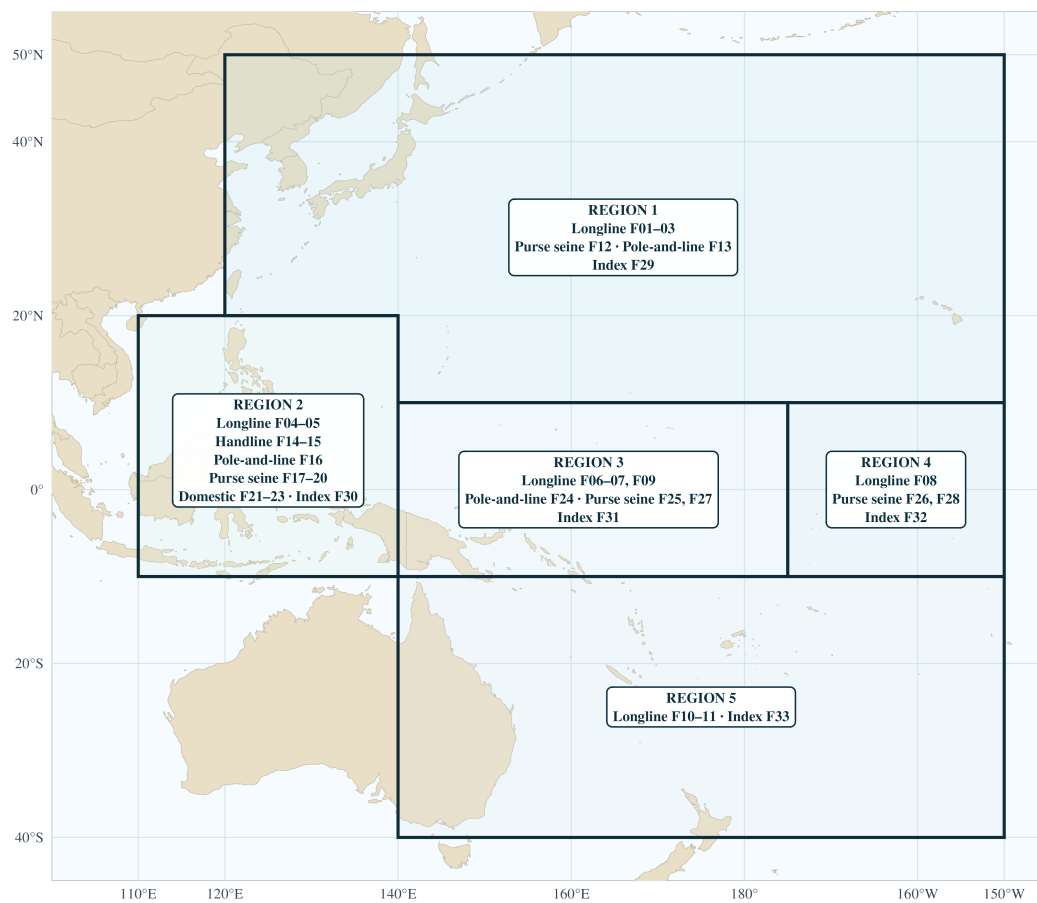


Figure 1: Five-region spatial structure used by the 2026 diagnostic model. Labels identify the fitted fisheries in each region by gear and fishery ID; the IDs correspond to Table 2.

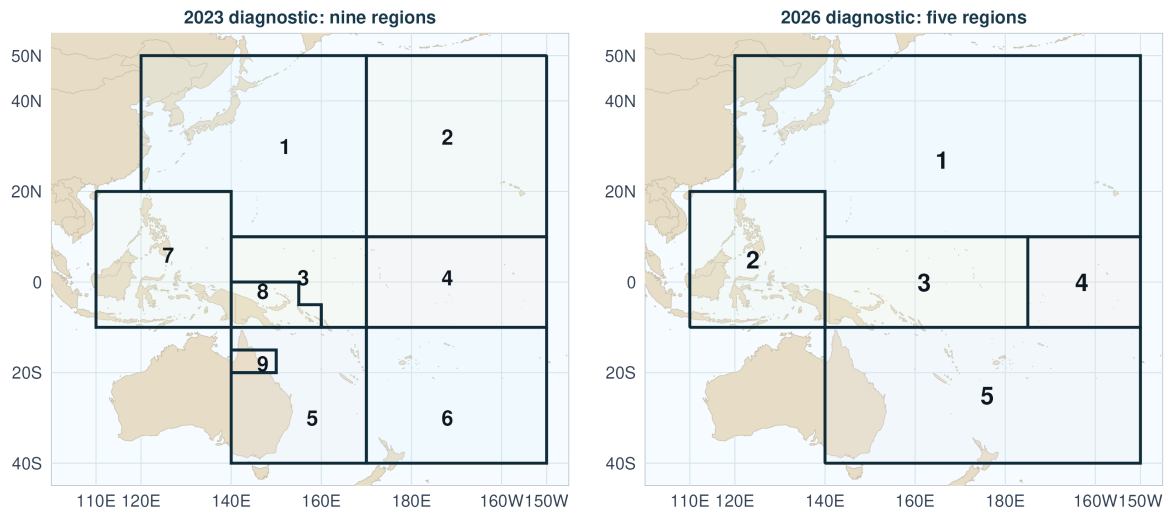


Figure 2: Comparison of the spatial structures used in the 2023 and 2026 diagnostic models. The 2026 structure consolidates the previous nine regions into five while retaining an eastern equatorial region for the principal tag-recapture information.

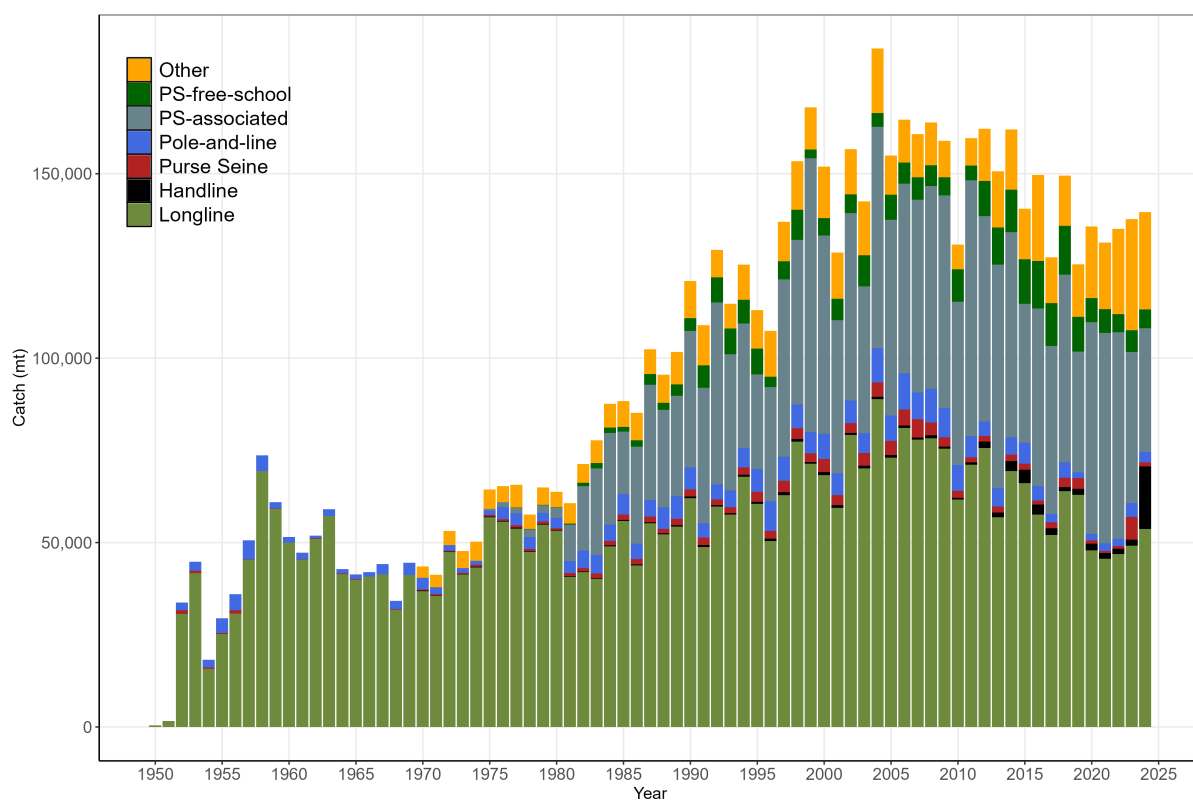


Figure 3: Annual catches of bigeye by gear type in the WCPO area covered by the assessment.

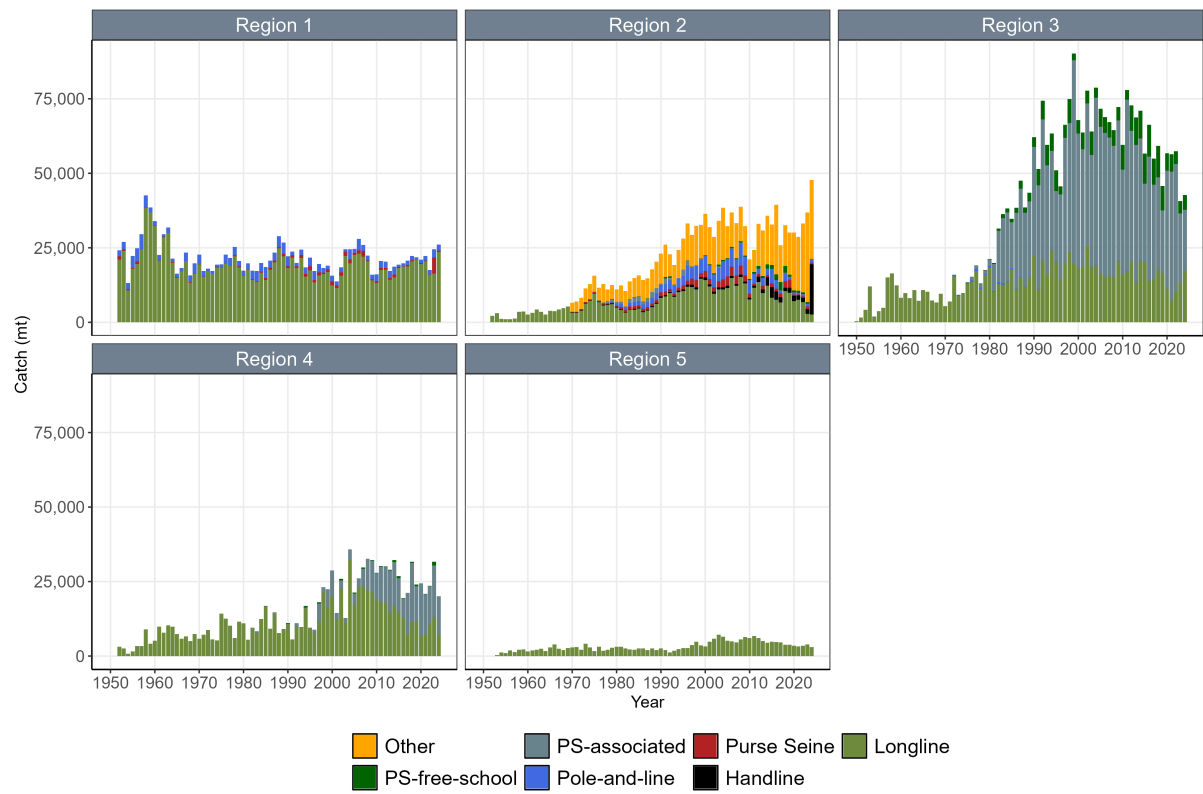


Figure 4: Annual catches of bigeye tuna by gear type for each of the five model regions.

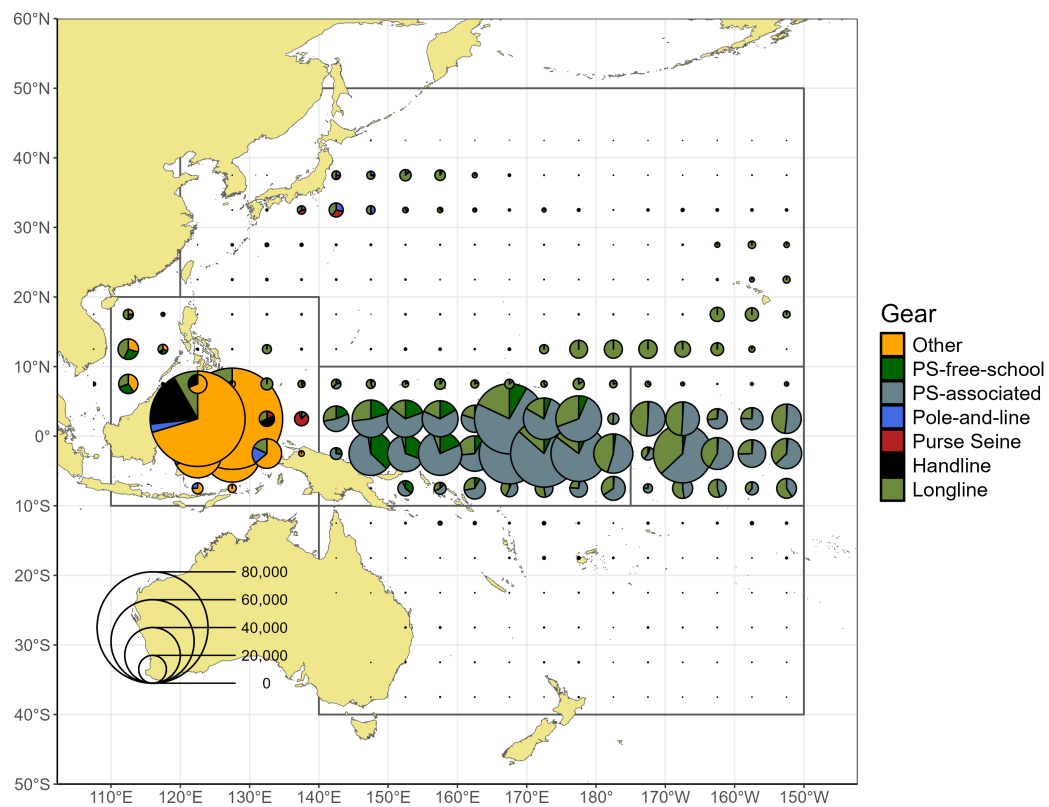


Figure 5: Distribution and magnitude of bigeye catches (mt) by gear type summed over the last 10 years (2015-2024) for 5 x 5 degree cells.

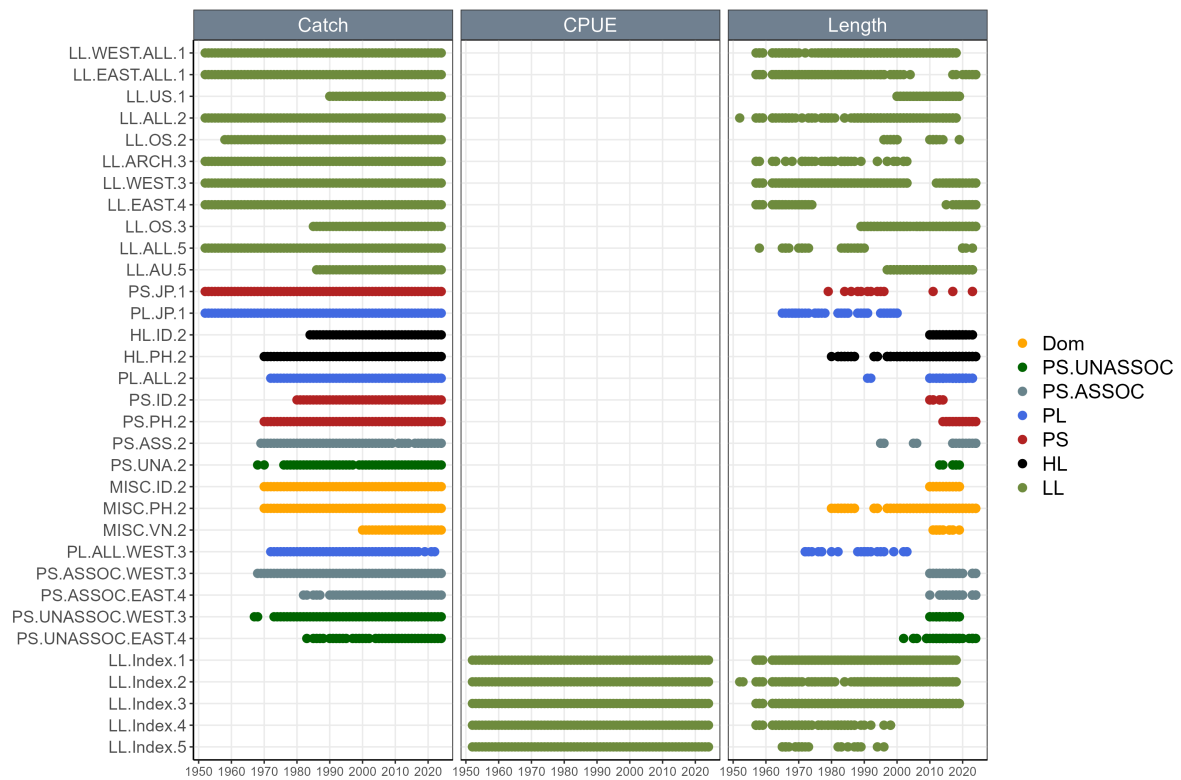


Figure 6: Summary of data coverage by fishery for the 2026 WCPO bigeye tuna assessment. I denotes index fisheries, L longline, P pole-and-line, S purse seine (unspecified), SA associated purse seine, SU unassociated purse seine and Z miscellaneous gears.

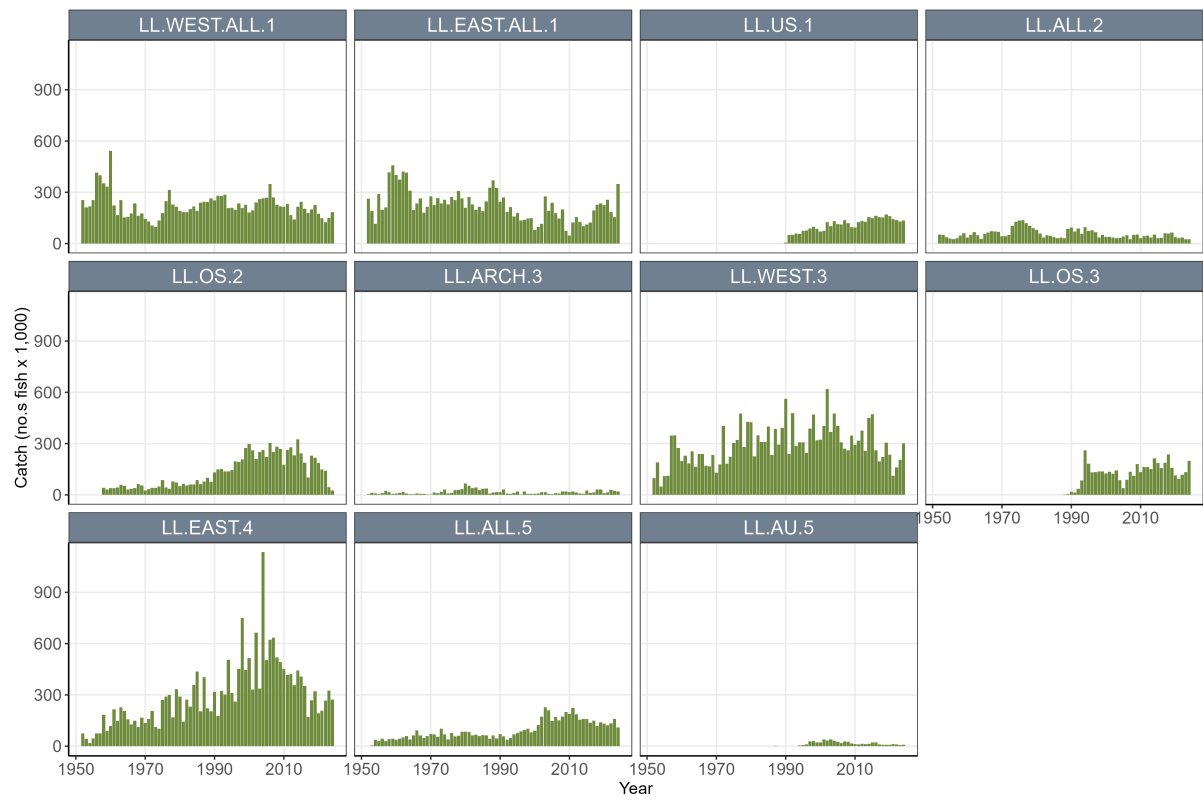


Figure 7: Time series of catches (numbers of fish) by fishery: longline.

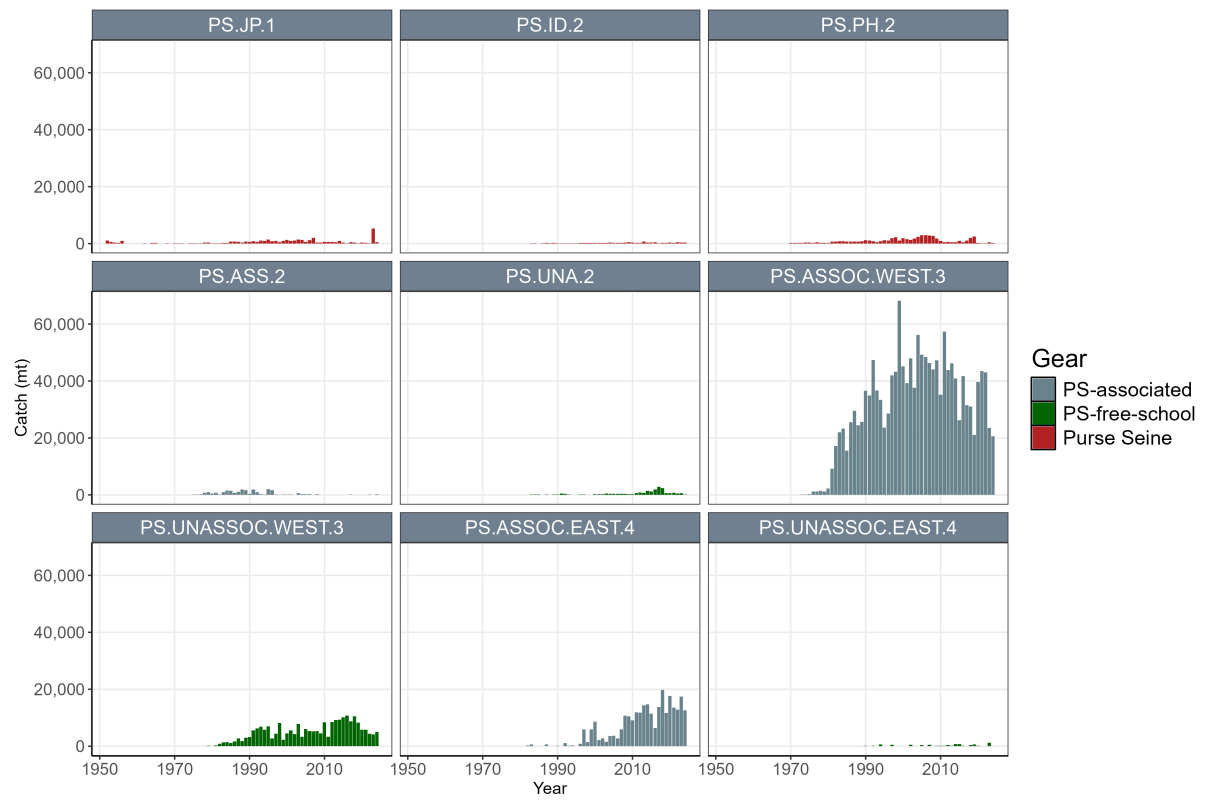


Figure 8: Time series of catches (mt) by fishery: purse seine.

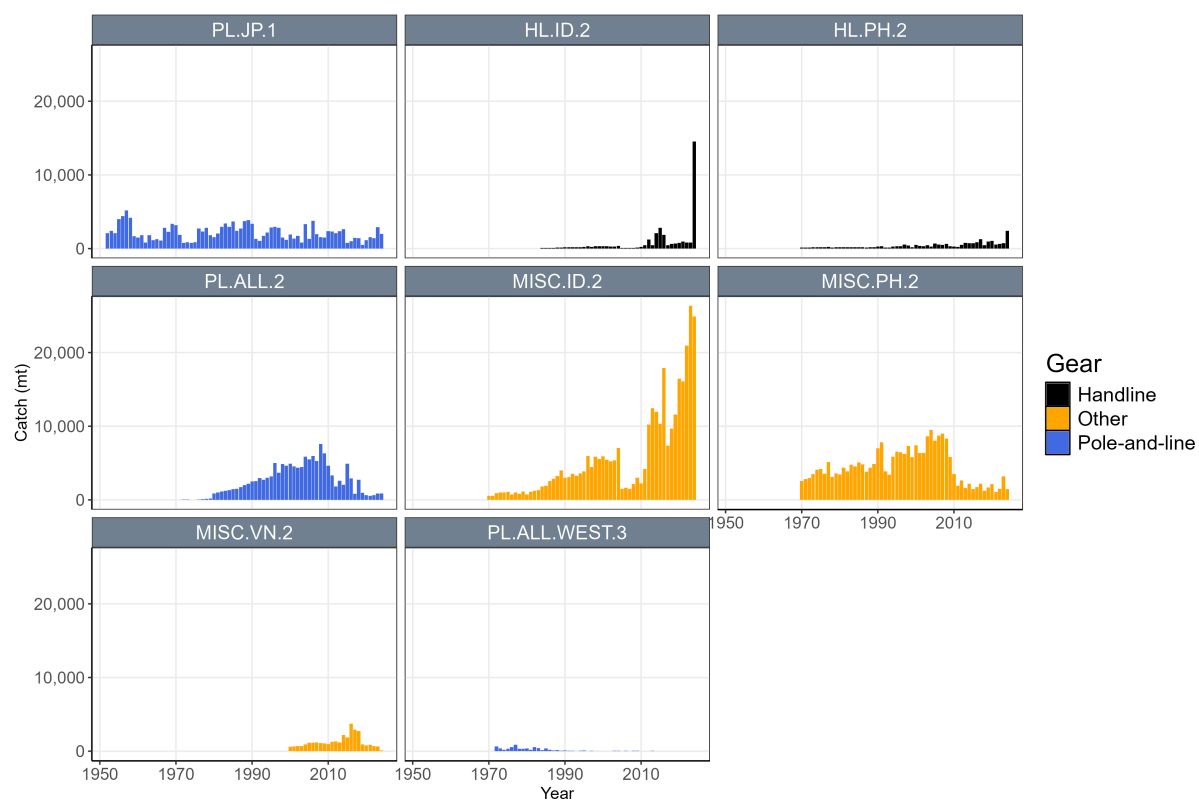


Figure 9: Time series of catches (mt) by fishery: other.

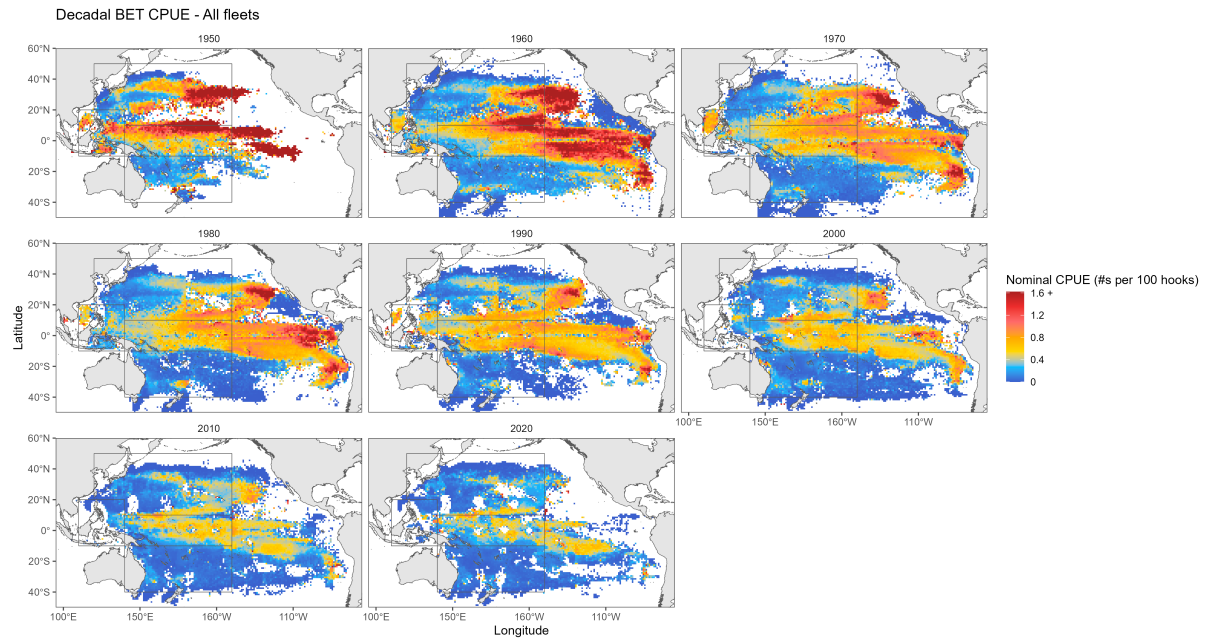


Figure 10: Spatial distribution of nominal longline CPUE (all fleets) for bigeye in the Pacific.

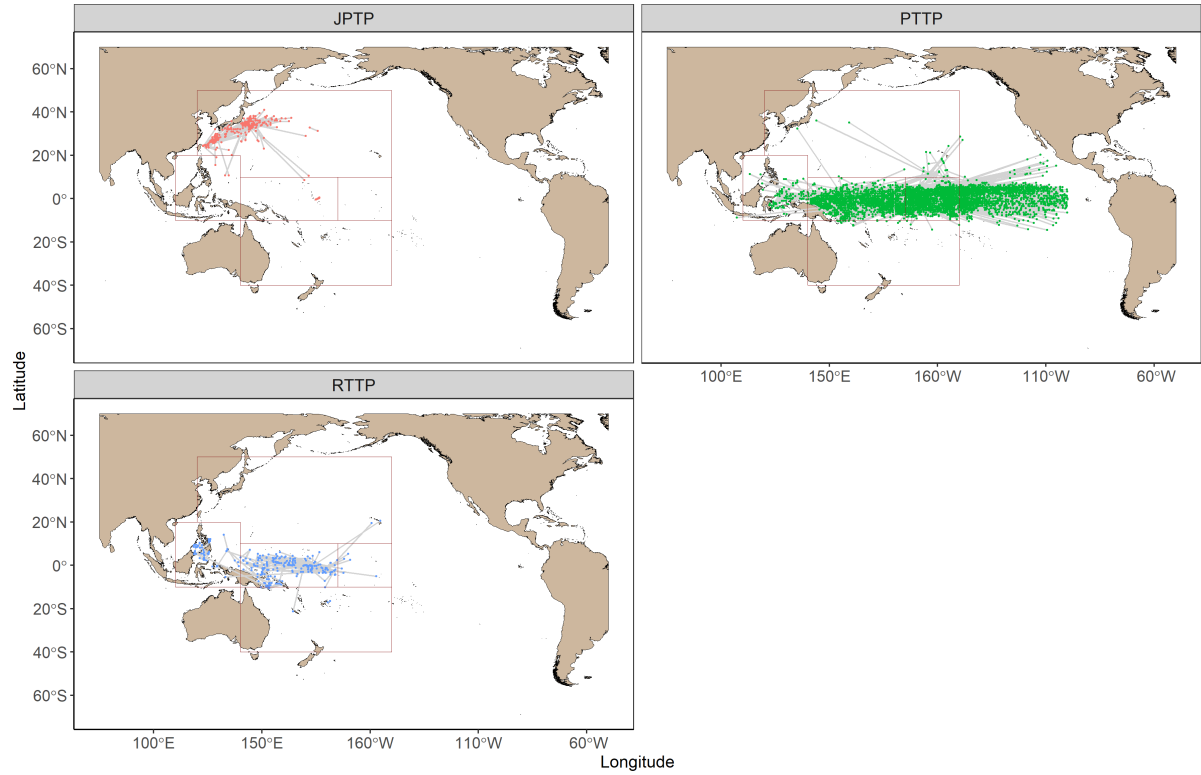


Figure 11: Map of tag recaptures. The panel shows the distributions of release and recapture displacements for the different tagging programs. Pacific Tuna Tagging Program (PTTP), Regional Tuna Tagging Program (excluding Coral Sea tags) (RTTP) and the Japanese Tagging Program (JPTP). Dots indicate recapture locations.

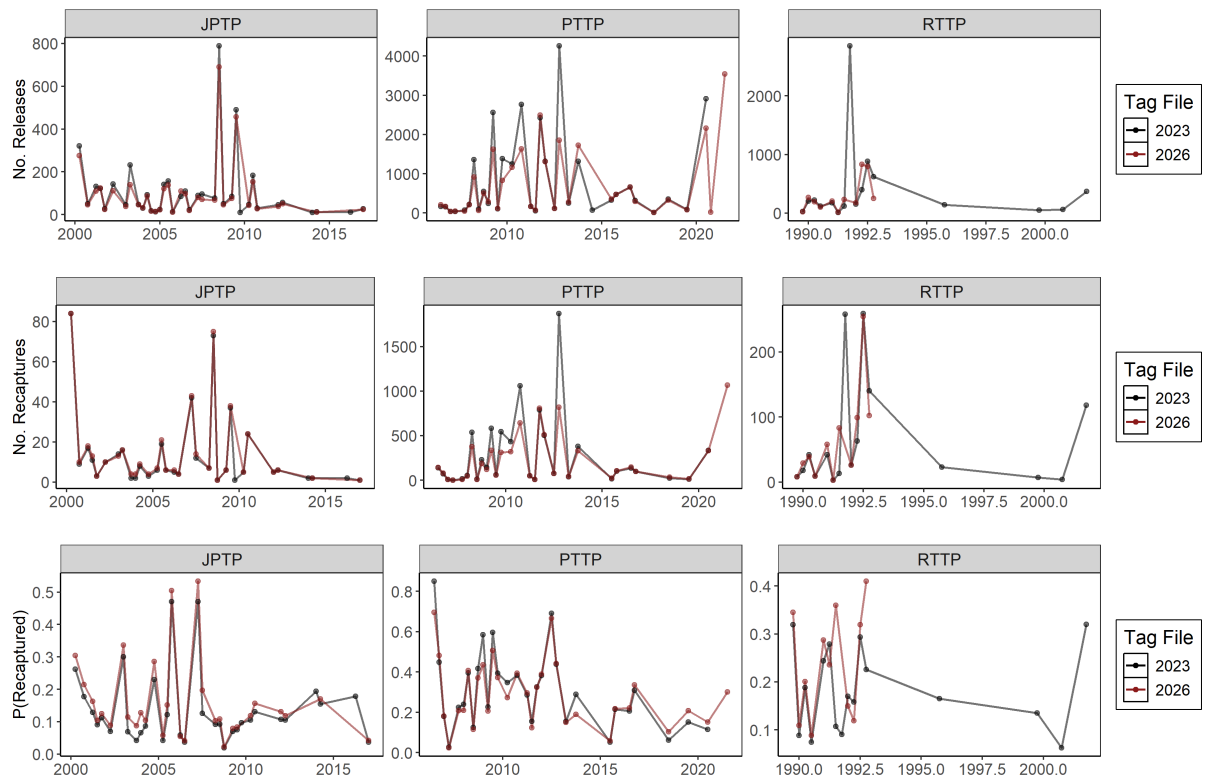


Figure 12: Summary plots of the number of releases, recaptures, and recapture rate of tags, by tagging program and region.

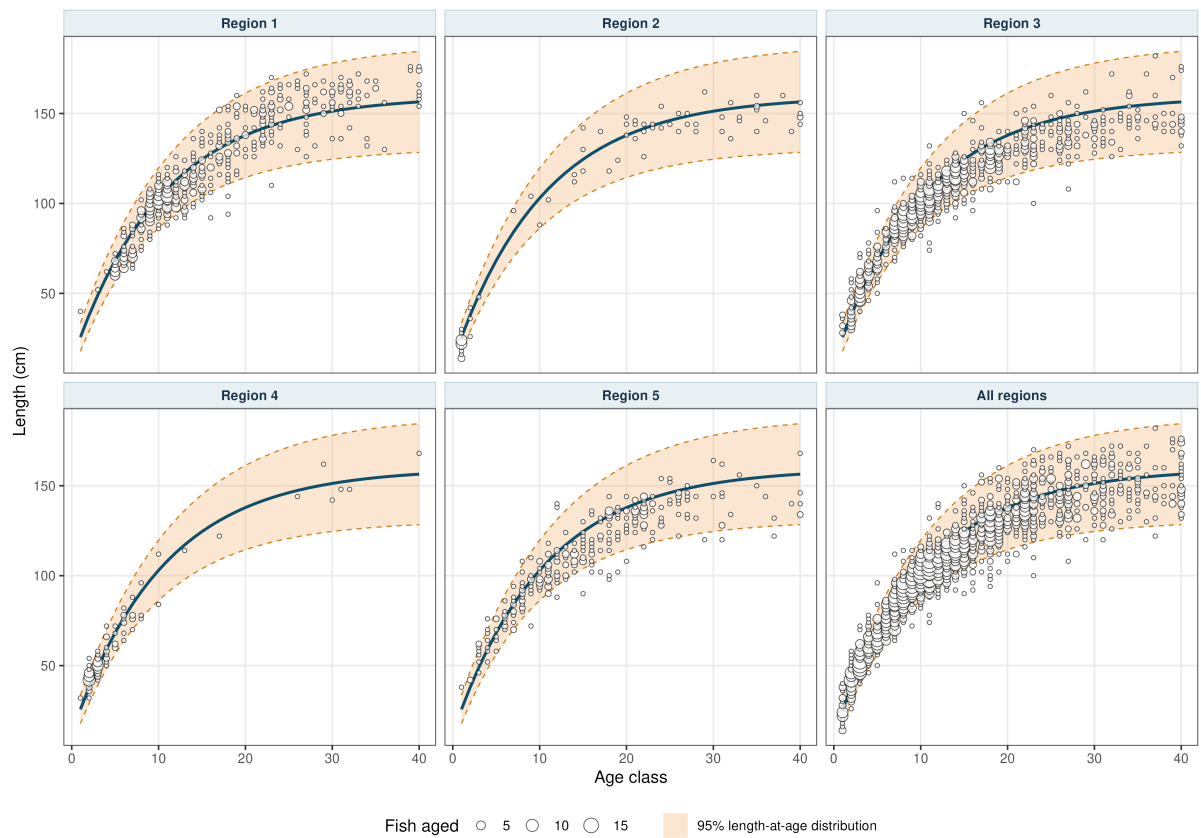


Figure 13: Observed age-length data and fitted mean length at age for the five model regions and all regions combined. Circle area is proportional to the number of aged fish. Shading is the fitted mean plus or minus 1.96 length-at-age standard deviations and represents fish-level length variability, not parameter-estimation uncertainty.

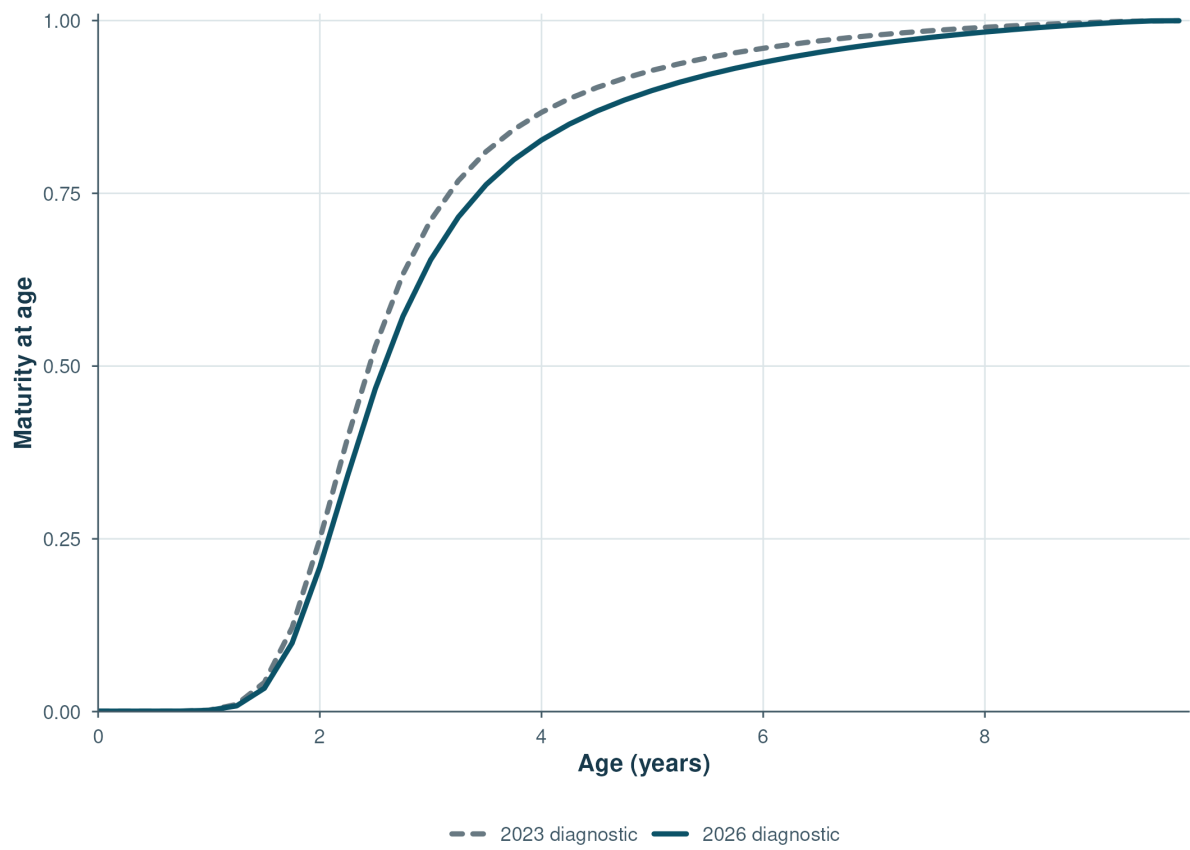


Figure 14: Maturity-at-age schedules in the 2023 and 2026 diagnostic models. The maturity-at-length relationship was unchanged; the small difference results from applying it to the updated 2026 growth curve.

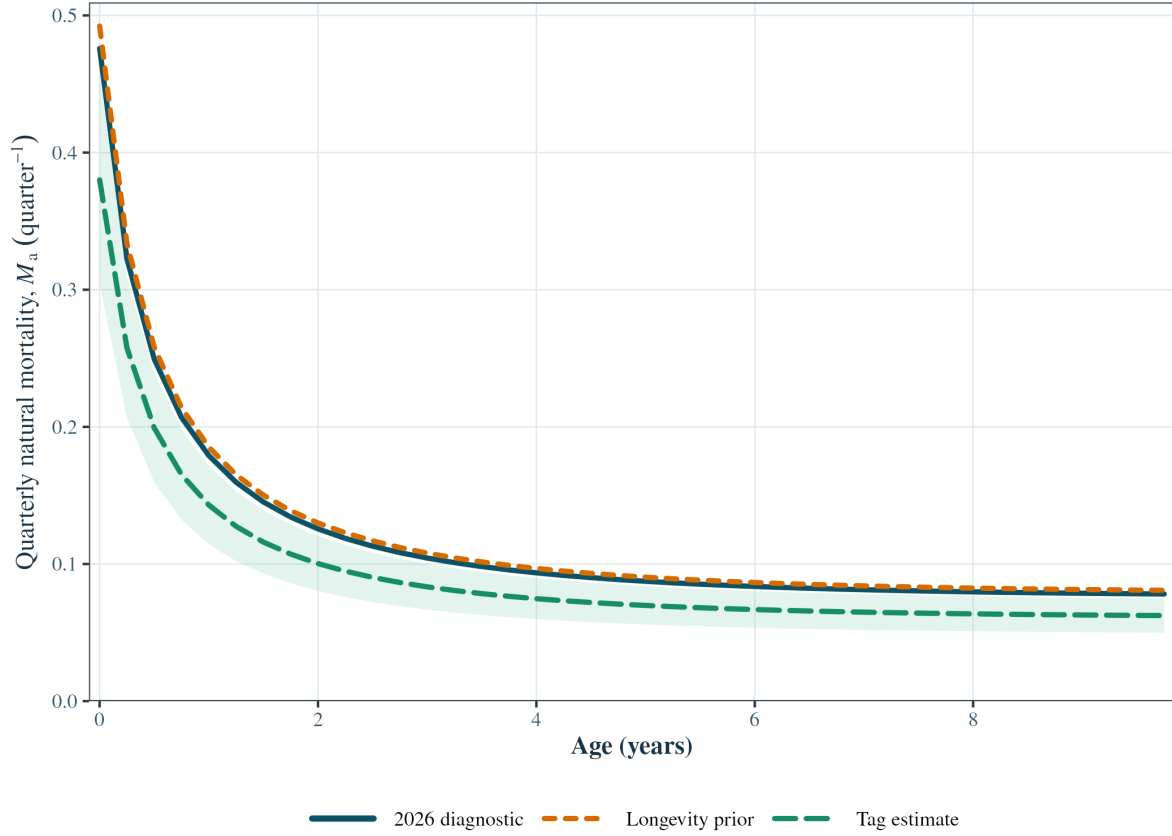


Figure 15: Age-specific quarterly natural mortality obtained by applying three assessment-relevant M_0 values to the same 2026 diagnostic growth schedule and fixed Lorenzen exponent (-1). Curves show the diagnostic value (0.07814), the model-aligned longevity-prior median (0.08082) derived following Hamel (2015) and Hamel & Cope (2022), and the tag estimate from Ducharme-Barth et al. (2026) (0.0624). Shading is the tag estimate's 90% confidence interval (0.0500–0.0749) propagated over age. Open the [ensemble report](#).

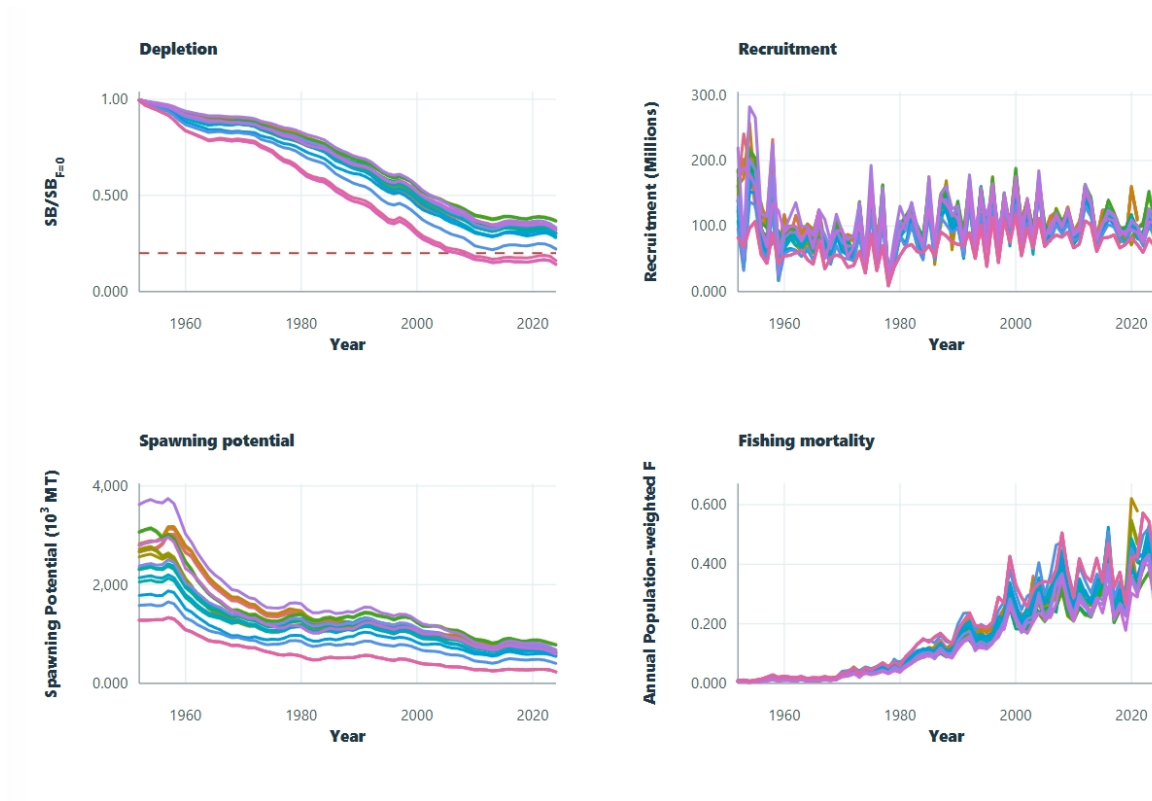


Figure 16: Stepwise model changes for depletion, recruitment, spawning potential and fishing mortality. The individual stepwise model changes can be explored in the [interactive HTML viewer](#).

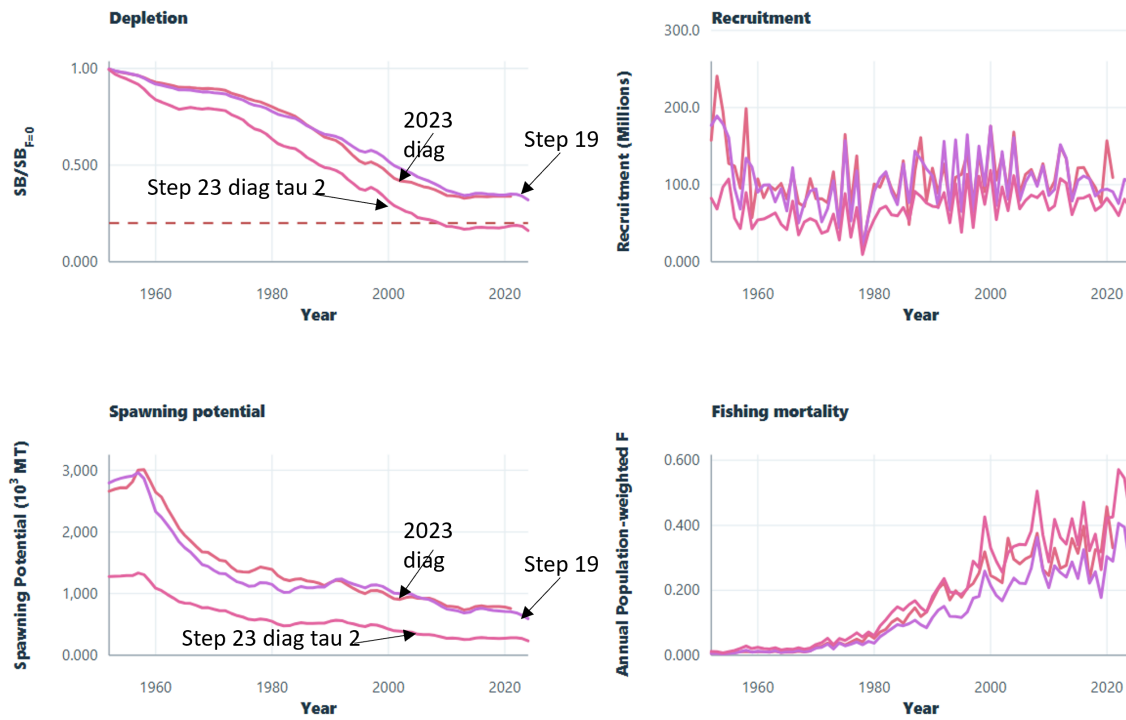
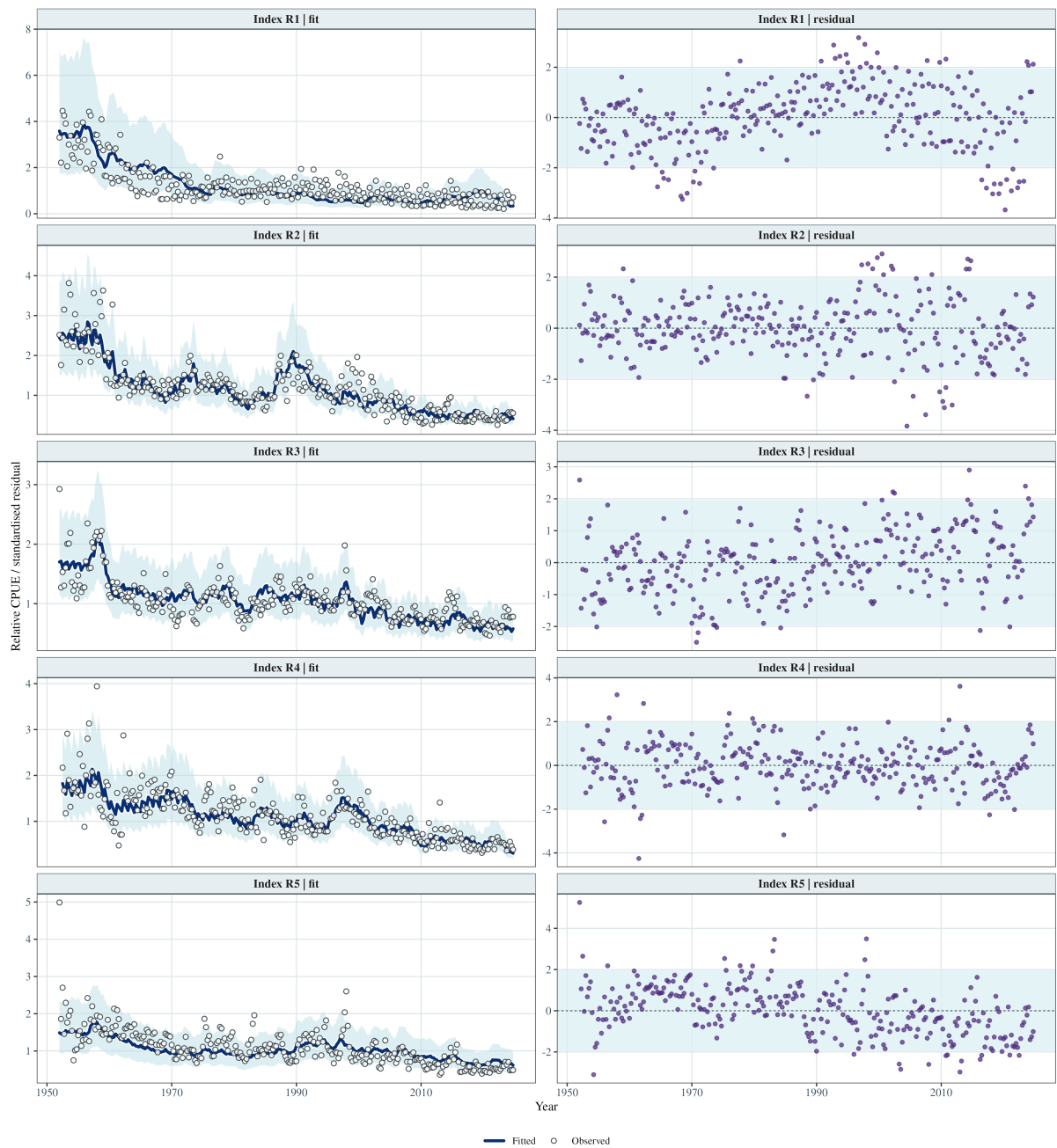


Figure 17: Selected stepwise trajectories for depletion ($SB/SB_{F=0}$), recruitment, spawning potential and fishing mortality. The figure contrasts the 2023 diagnostic-model refit, step 19 and the 2026 diagnostic model, highlighting the change after tag overdispersion τ was increased from 1 to 2. Open the [interactive stepwise viewer](#).



Residual panels: shaded band = ± 1.96 fixed observation-error SD.

Figure 18: Observed and fitted relative CPUE (left panels) with standardised log-scale residuals (right panels) for the five regional index fisheries. Fit shading gives the central 95% observation interval implied by the fixed regional log-scale error values; it is not parameter-estimation uncertainty. Residual shading marks the corresponding ± 1.96 observation-error SD range. Open the [interactive diagnostic report](#).

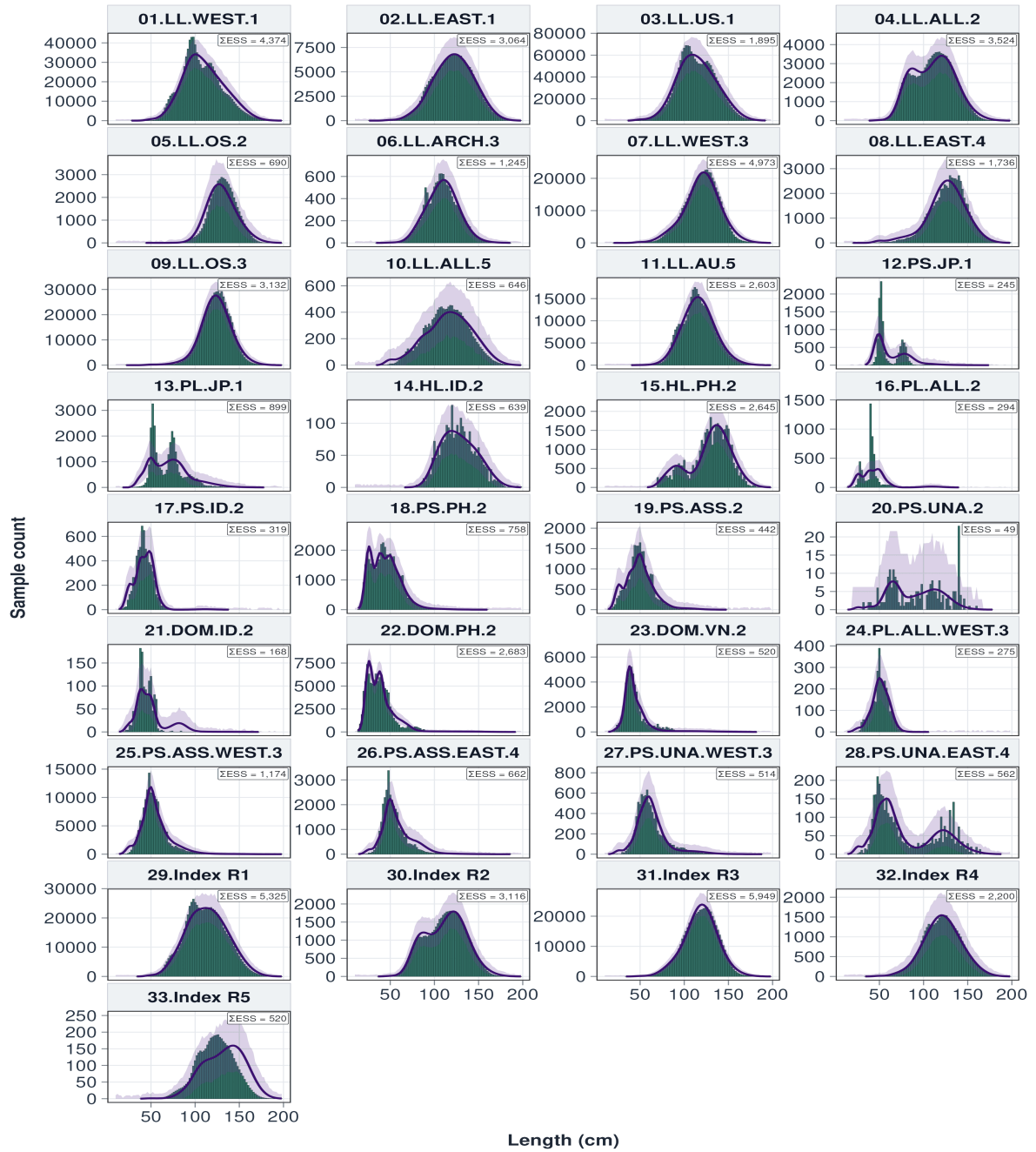


Figure 19: Observed (bars) and fitted (purple lines) length-frequency sample counts by fishery, pooled over years. Shading is the 95% pointwise conditional predictive interval for repeated observations under the fitted Dirichlet–multinomial model; parameter uncertainty is not included. The top-right ΣESS is the sum of model-implied incident effective sample sizes in each fishery. Open the [interactive diagnostic report](#) for fishery- and year-level selections.

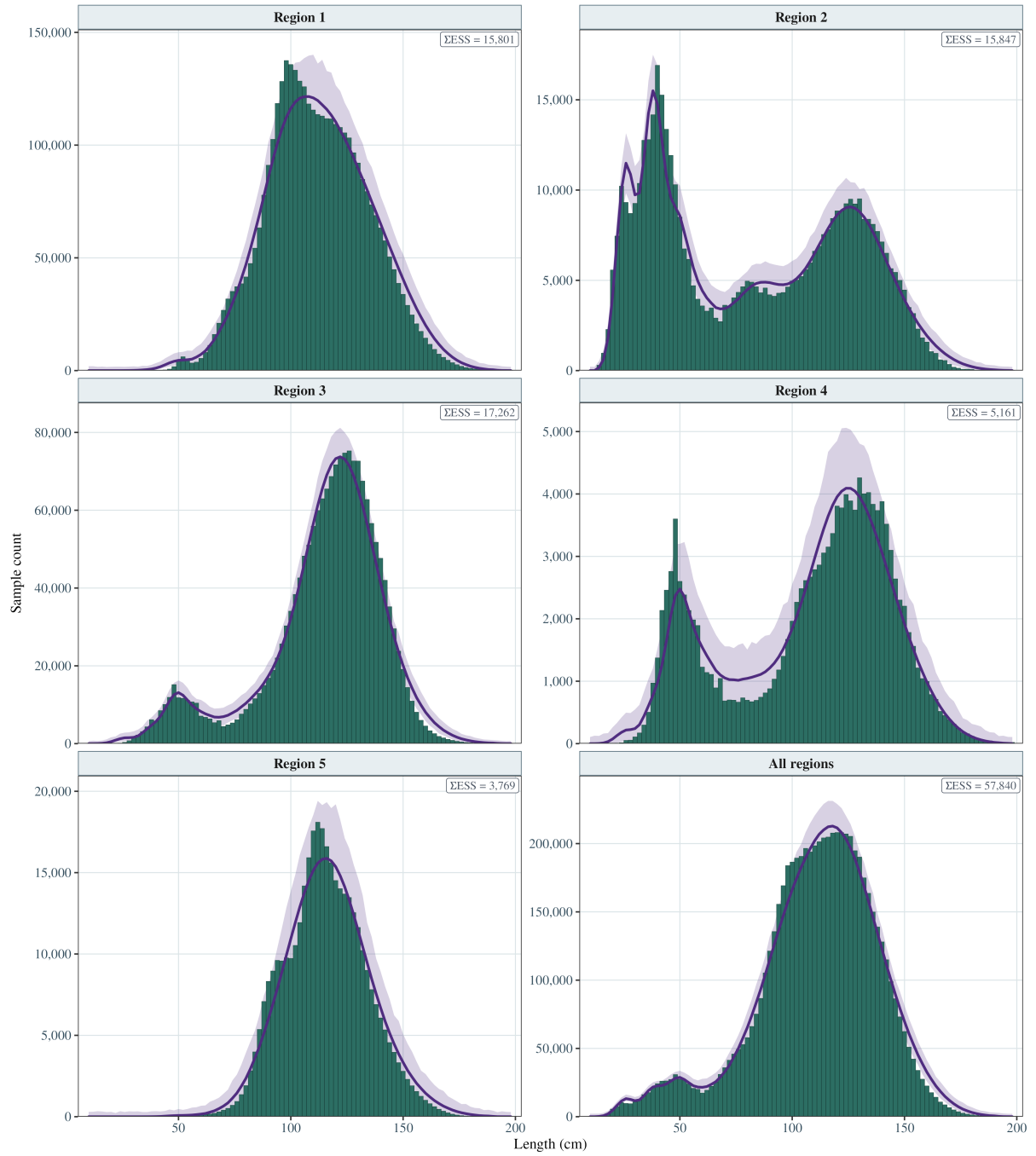


Figure 20: Observed (bars) and fitted (purple lines) length-frequency sample counts aggregated by model region and over all regions. Shading is the 95% pointwise conditional predictive interval for repeated observations under the fitted Dirichlet–multinomial model; parameter uncertainty is not included. The top-right ΣESS is the sum of model-implied incident effective sample sizes within each panel. Counts are pooled over years and fisheries for this complementary regional summary. Panel y axes are free and begin at zero. Open the [interactive diagnostic report](#) for fishery- and year-level selections.

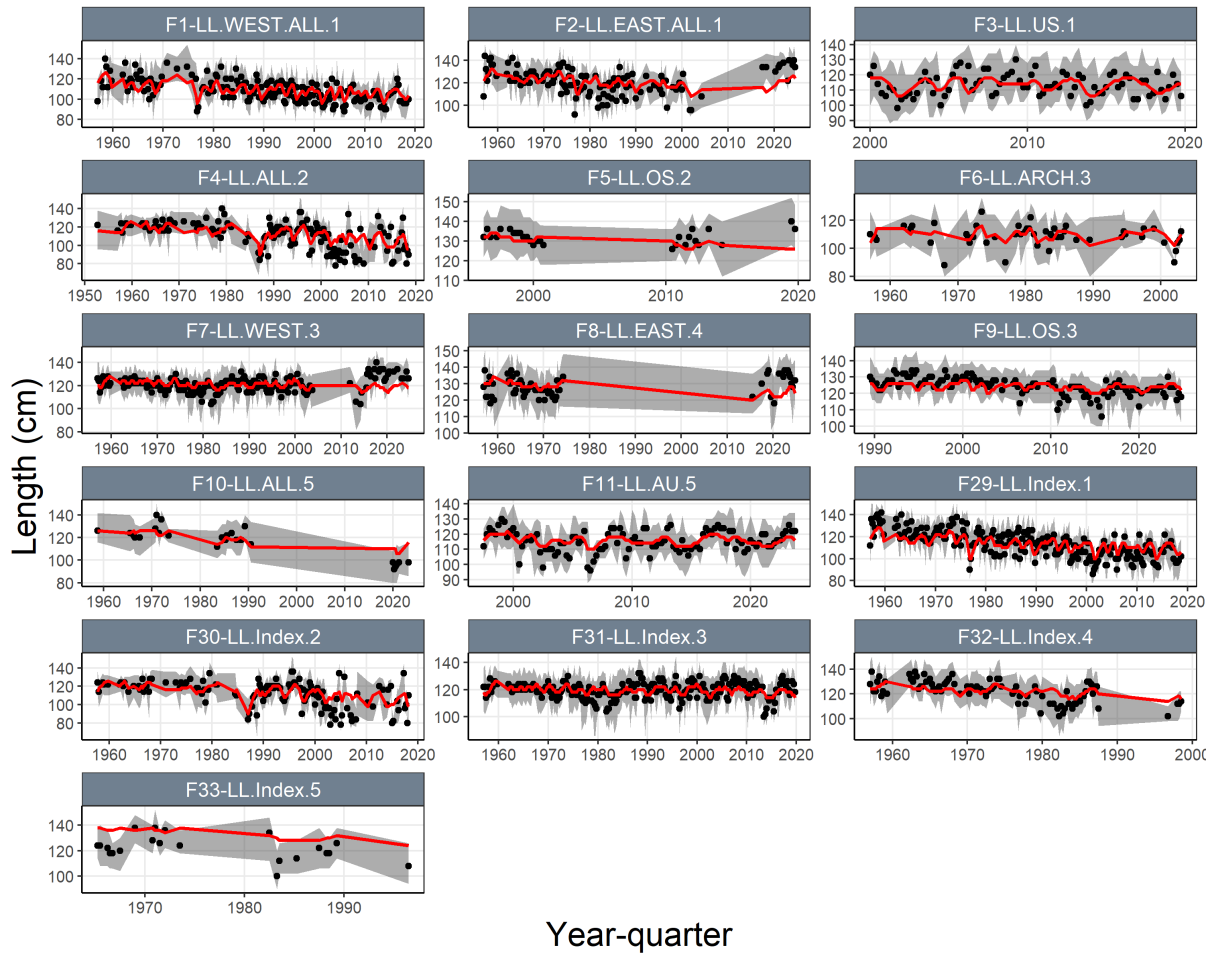


Figure 21: Observed (red points) and predicted (red line) median fish lengths (FL, cm) for the longline fisheries for the 2026 diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples with more than 250 fish per quarter/fishery strata are plotted.

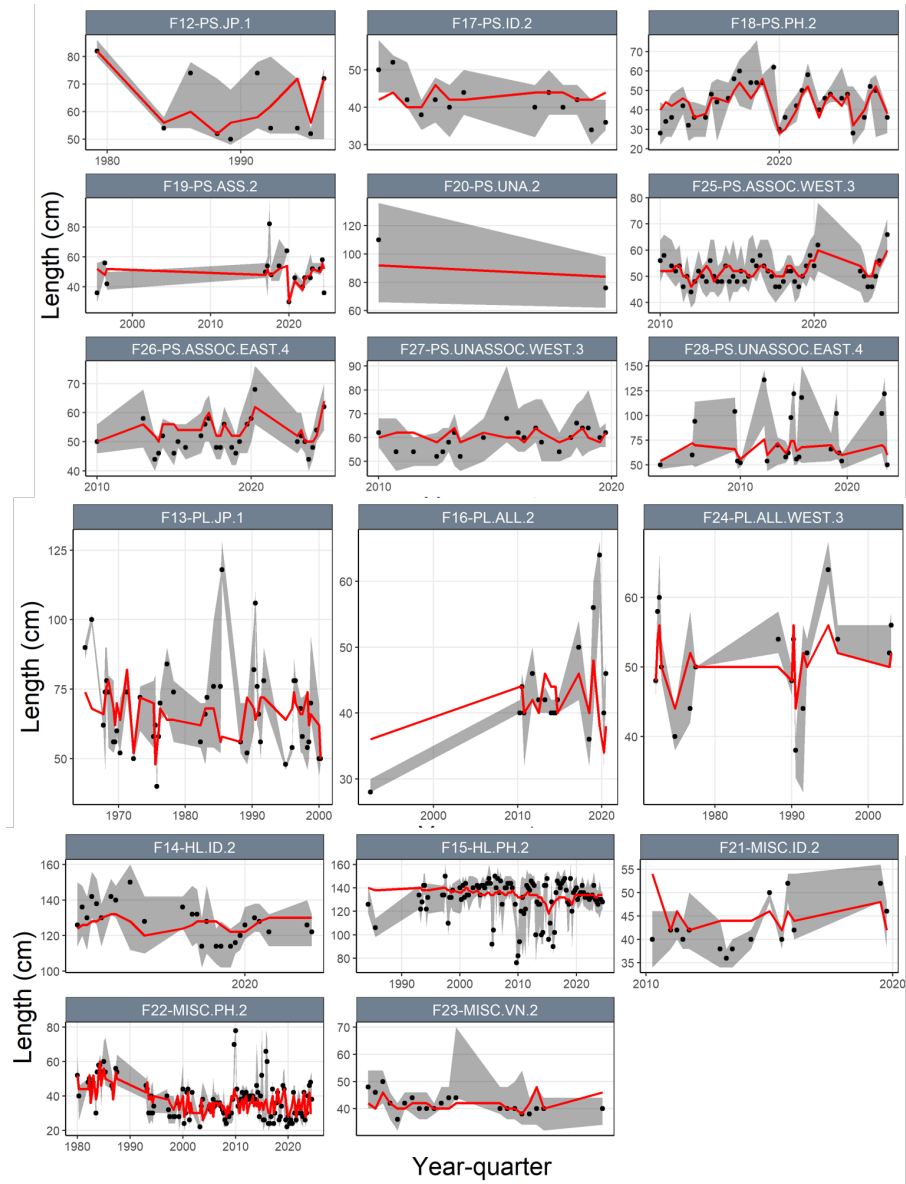


Figure 22: Observed (red points) and predicted (red line) median fish lengths (FL, cm) for the other fisheries for the 2026 diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples with more than 100 fish per quarter/fishery strata are plotted.

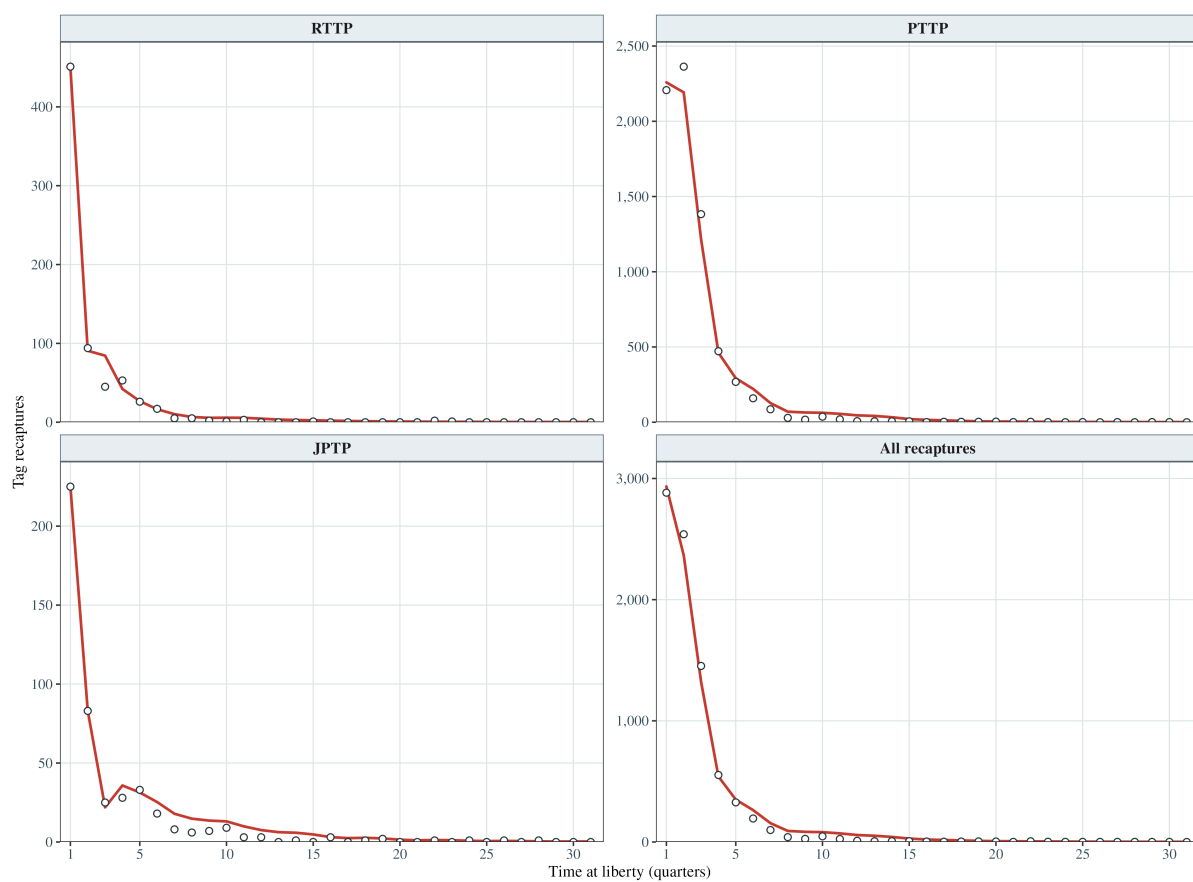


Figure 23: Observed (black points) and model-predicted (red line) tag recaptures by time at liberty in quarters for RTTP, PTPP, JPTP and all recaptures. Programme predictions apply MFCL's premixing reporting-rate rule and reconcile to the official all-recapture diagnostic within its printed precision.

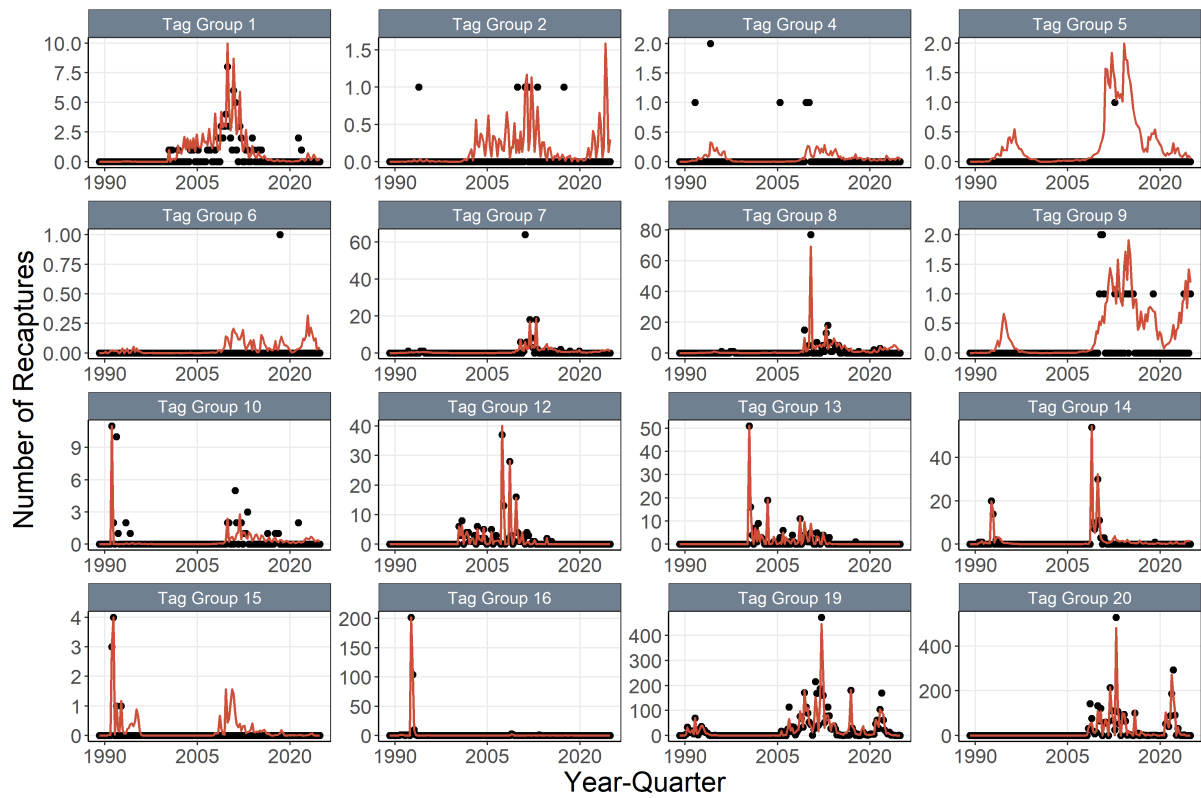


Figure 24: Observed (black points) and model-predicted (red lines) tag recaptures through time for the principal tag-recapture groups in the 2026 diagnostic model.

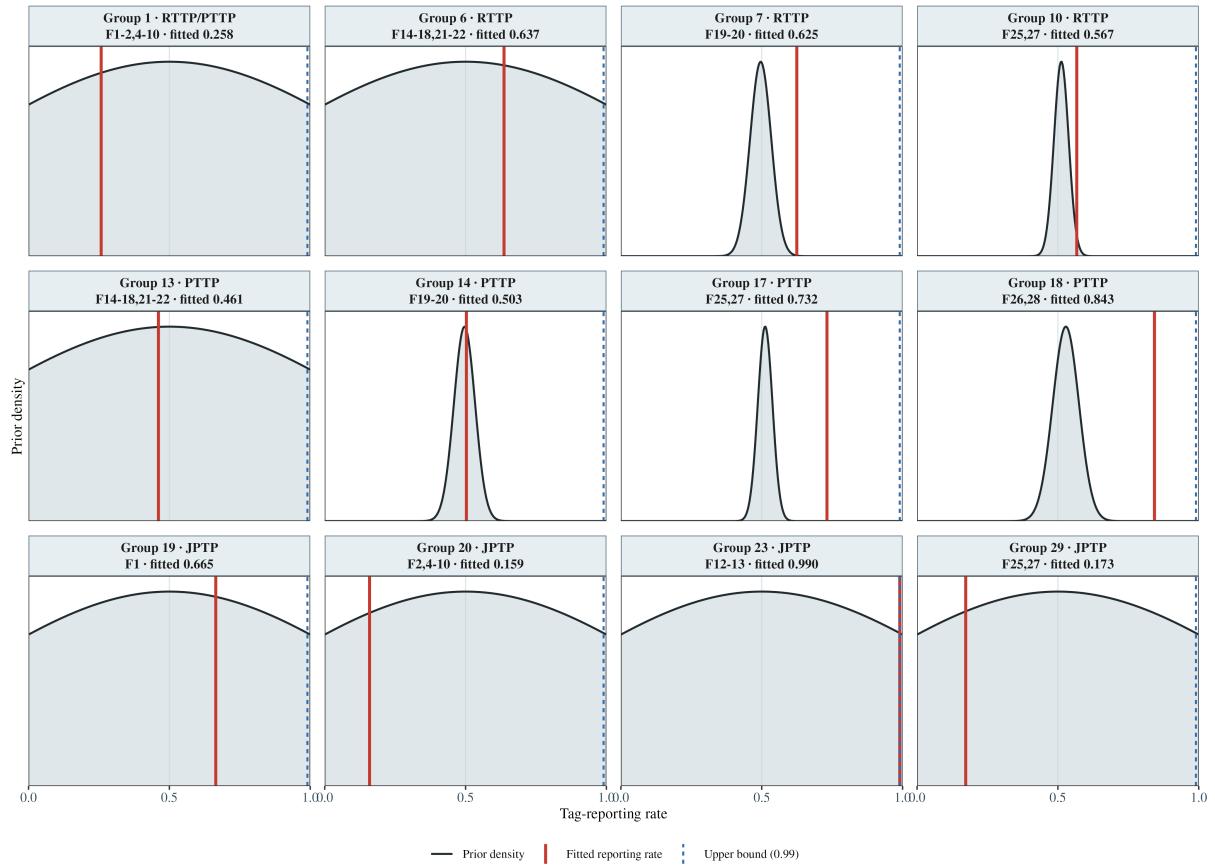


Figure 25: Fitted tag-reporting rates for the 12 groups estimated in the diagnostic model. Panels identify the MFCL reporting-rate group, tagging programme and recapture fisheries. Black curves show the prior density, red vertical lines show fitted reporting rates and blue dashed lines mark the 0.99 upper bound. Group 23 reached the upper bound. Fixed groups and the inactive Group 31 placeholder assigned to Index fisheries F29–F33 are not shown. Open the [interactive diagnostic report](#).

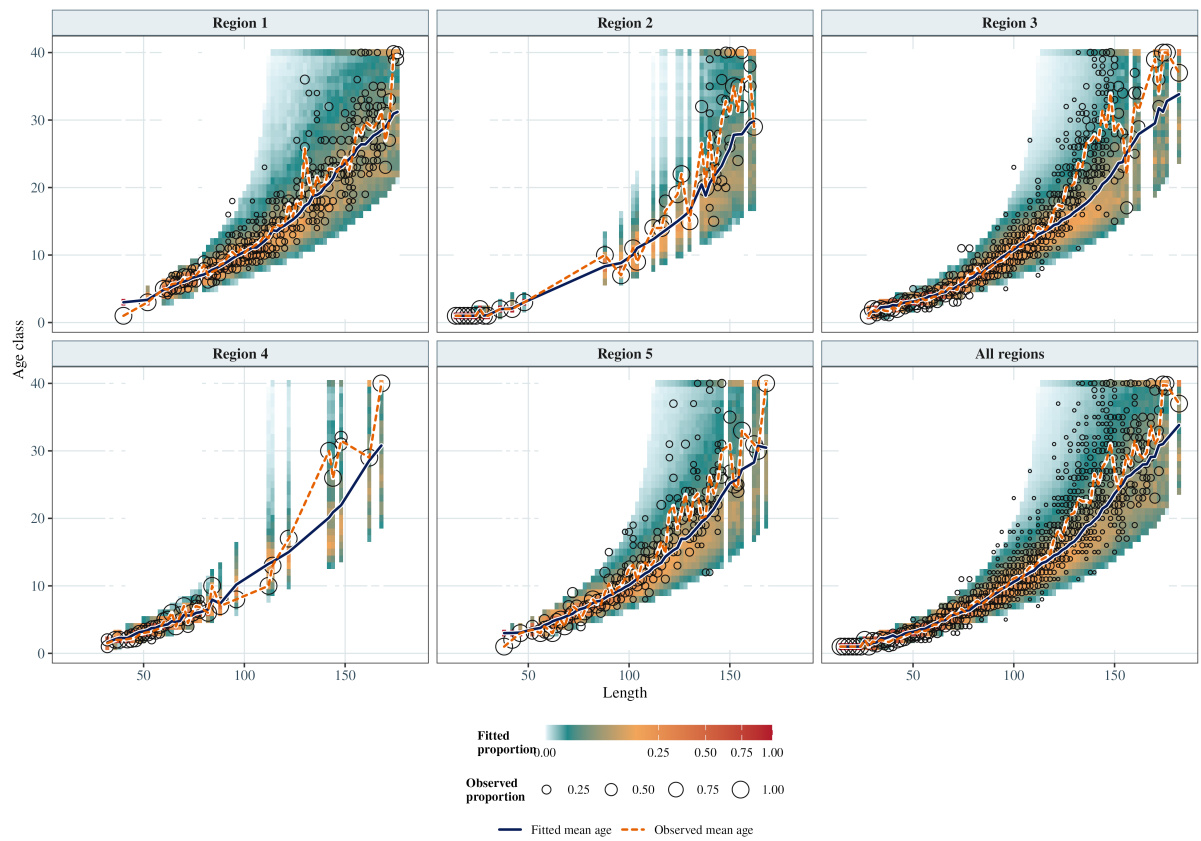


Figure 26: Observed and fitted conditional age-at-length distributions for the five model regions and the pooled All regions panel (last). Cells and point sizes show fitted and observed proportions; solid and dashed lines show their corresponding mean ages at length. Open the [interactive diagnostic report](#) for fishery-level detail.

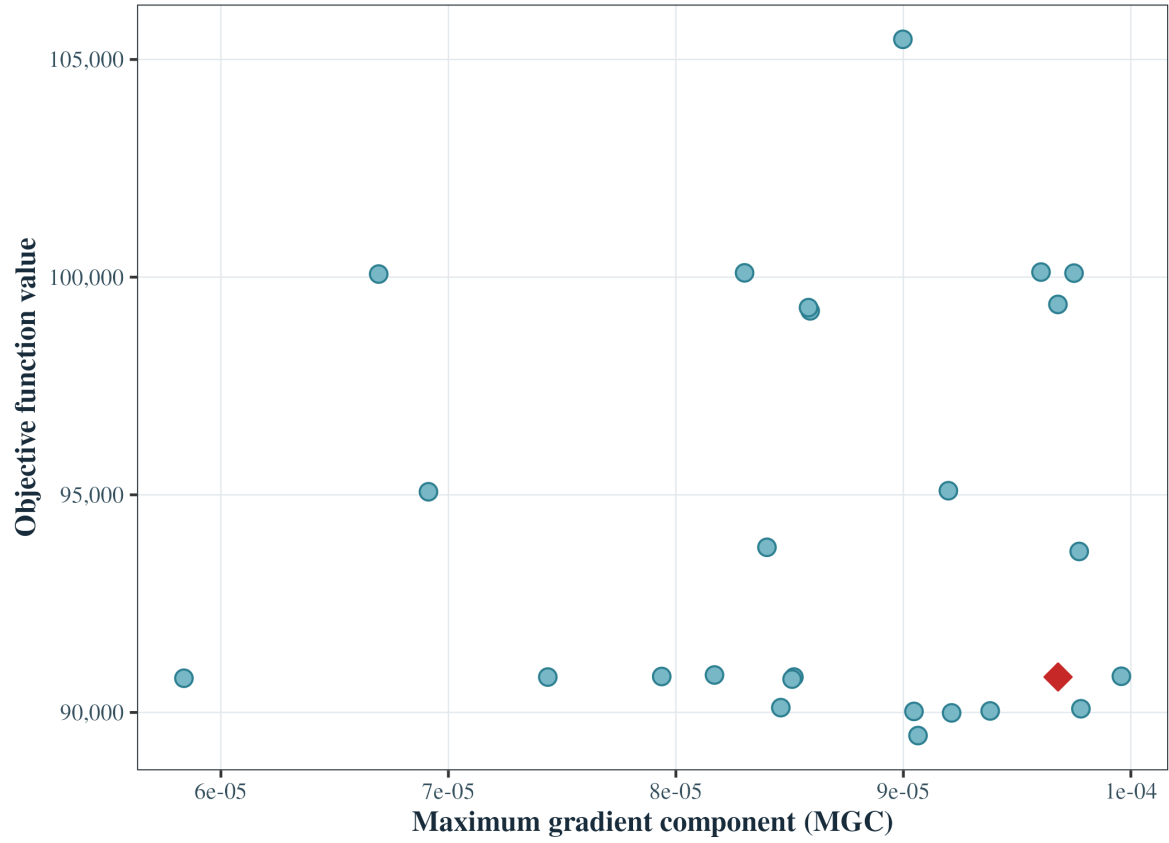


Figure 27: Jitter convergence diagnostics for the diagnostic model. Each circle represents one of the 25 retained jitter fits and the red diamond identifies the diagnostic model without jitter. Retained fits had a maximum gradient component no greater than 1.0×10^{-4} ; lower objective-function values indicate improved fit. Open the [interactive jitter report](#).

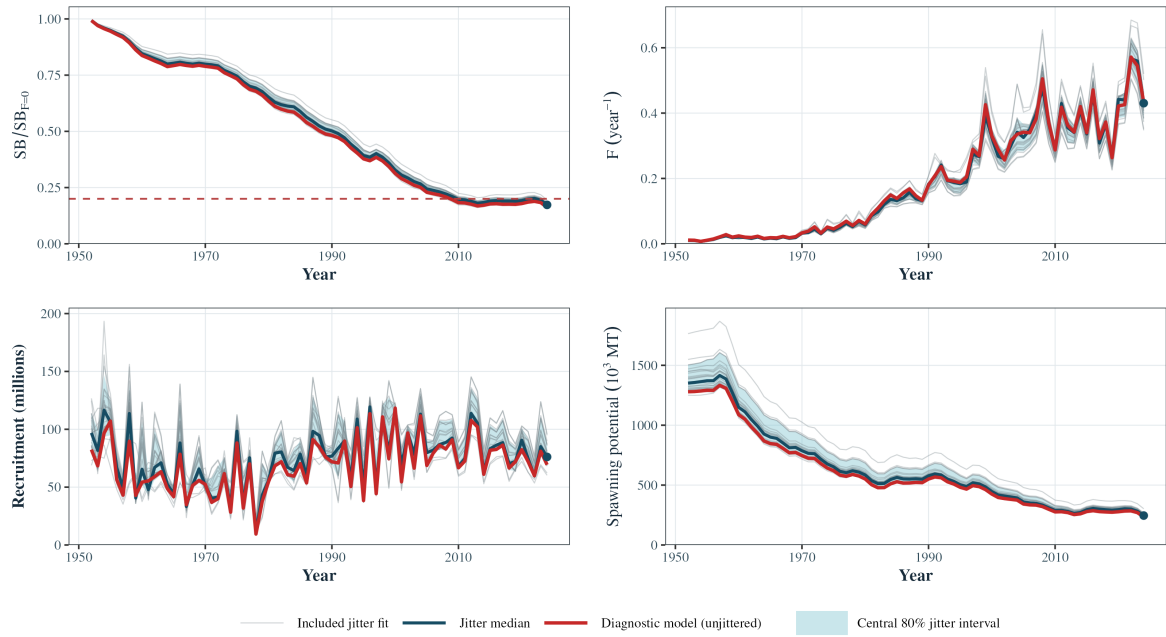


Figure 28: Annual derived quantities across the 25 retained jitter fits. Thin grey-blue lines are individual fits, the shaded band is the pointwise central 80% interval, the dark line is the jitter median with its terminal value marked by a circle, and the red line is the diagnostic model without jitter. Recruitment is in millions of fish and spawning potential is in thousand metric tonnes. Open the [interactive jitter report](#).

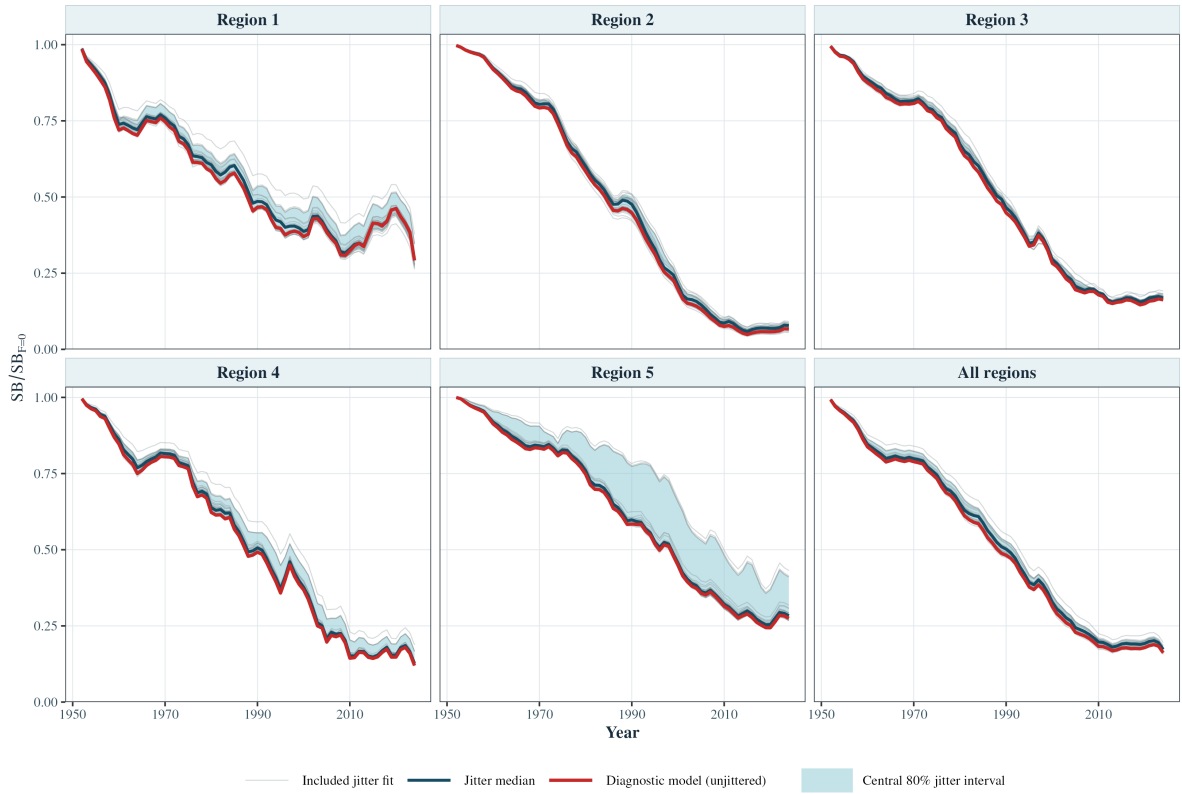


Figure 29: Regional and all-region depletion trajectories across the 25 retained jitter fits. Thin grey-blue lines are individual fits, the shaded band is the pointwise central 80% interval, the dark line is the jitter median and the red line is the diagnostic model without jitter.

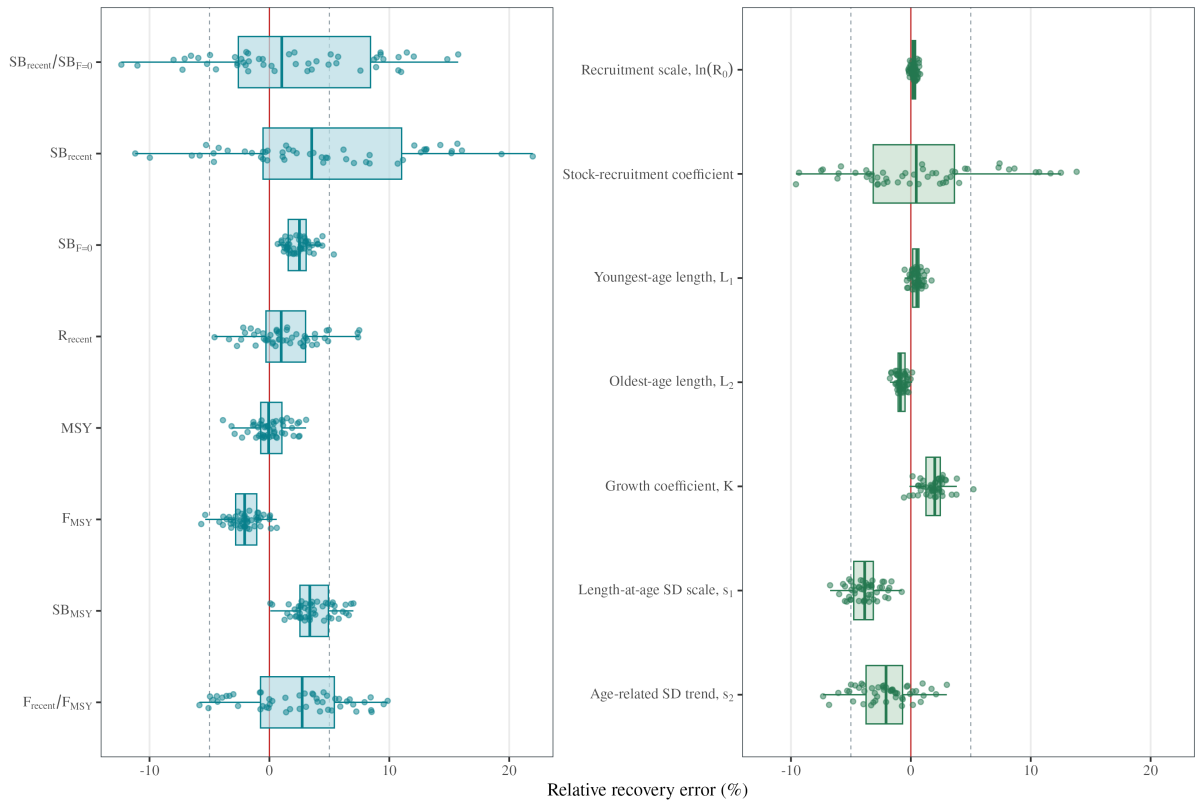


Figure 30: Relative recovery error for recent management and assessment quantities and selected estimated parameters in the diagnostic model. Points are the 50 pseudo-data refits, boxes show their empirical distributions, the red line marks exact recovery and grey dashed lines mark errors of -5% and +5%. Open the [interactive self-test report](#).

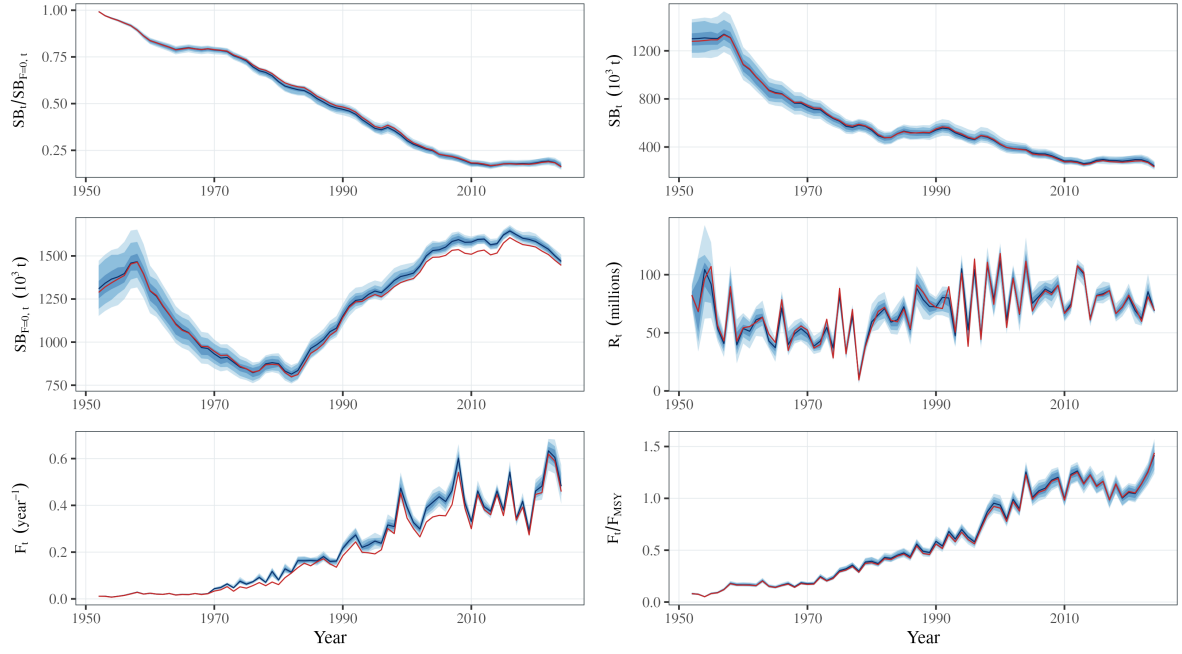


Figure 31: Self-test recovery of annual assessment quantities. The red line is the generating value, the dark-blue line is the median across 50 replicate refits and the nested blue bands are the central 50%, 80% and 95% empirical ranges. Open the [interactive self-test report](#).

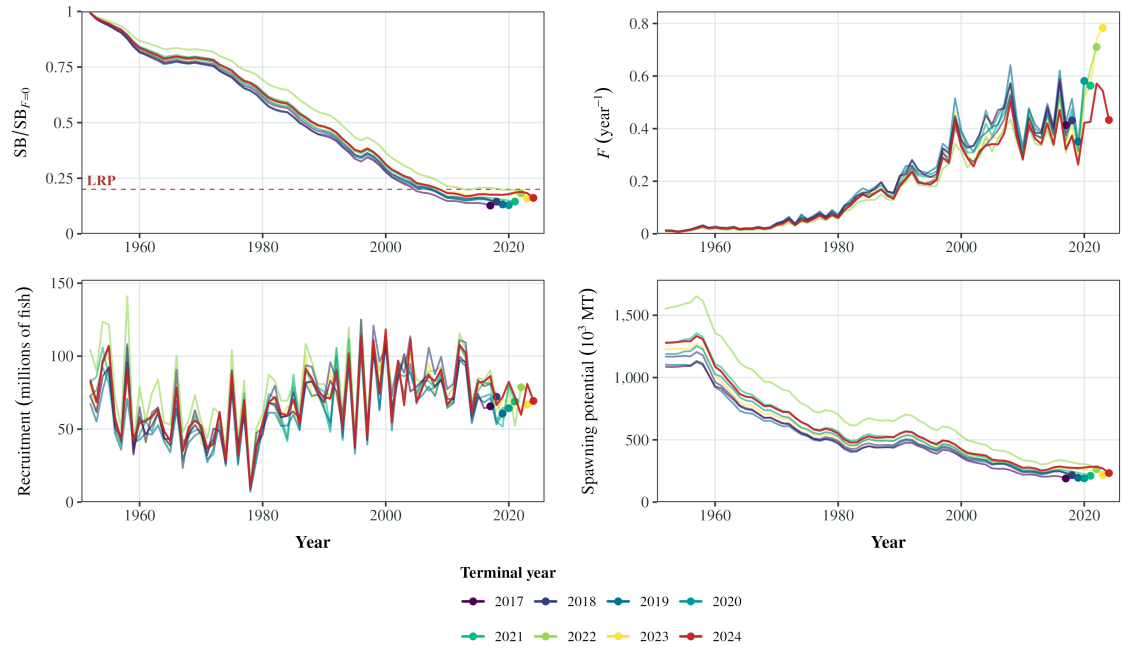


Figure 32: Retrospective trajectories for (a) dynamic spawning depletion, (b) spawning potential, (c) recruitment and (d) annual population-weighted fishing mortality. Colours distinguish the diagnostic fit and seven terminal-year peels; labels give Mohn's ρ . Open the [interactive retrospective report](#).

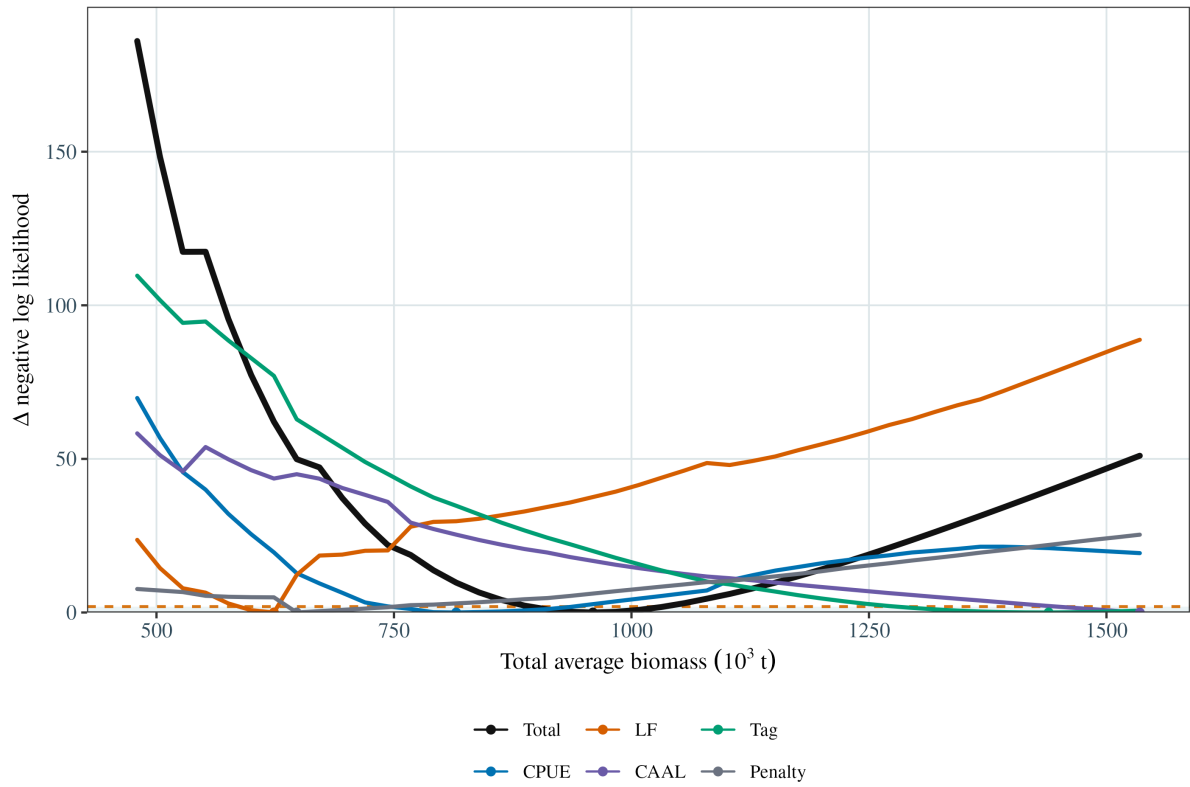


Figure 33: Overall-objective and component likelihood profiles overlaid on the absolute total-average-biomass scale. Each curve is expressed as a change from its own minimum; the horizontal line marks a change of 1.92. Open the [interactive profile viewer](#).

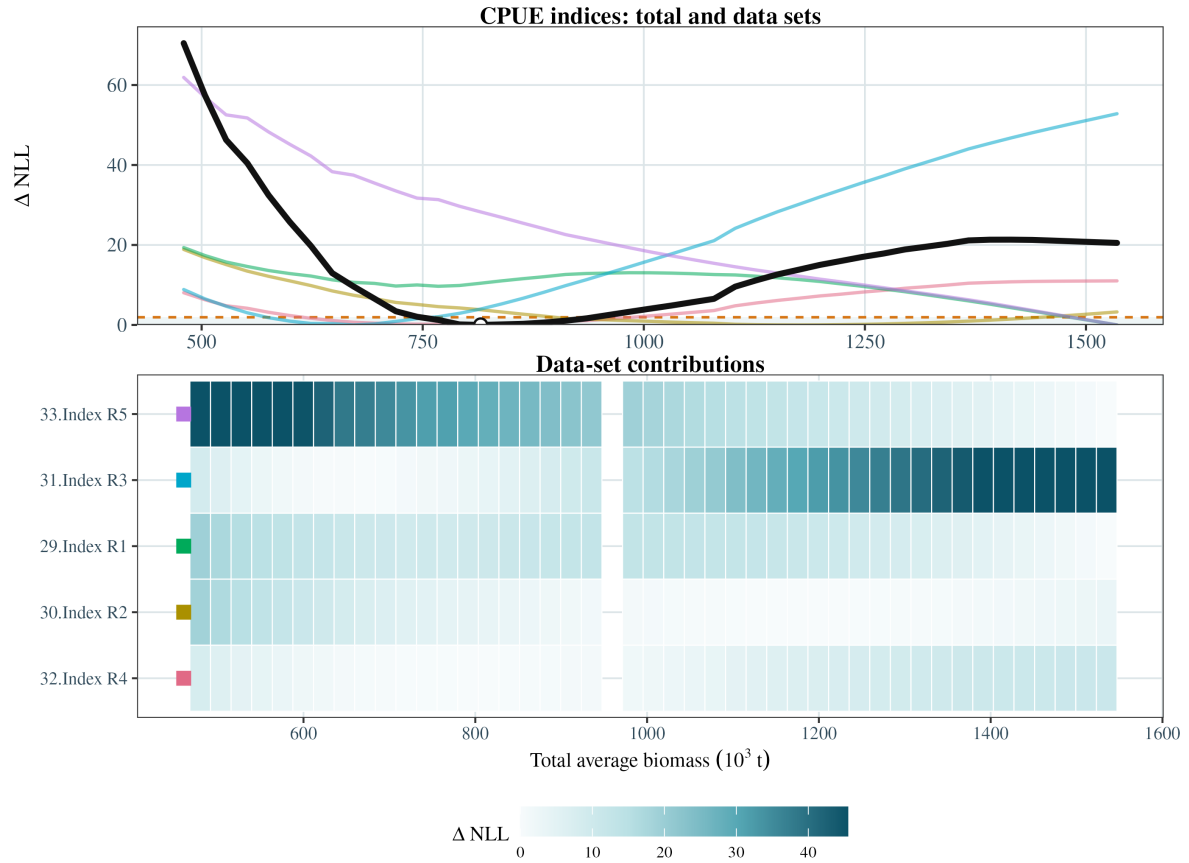


Figure 34: CPUE likelihood profile. The black line is the total CPUE contribution; the coloured lines and matching heatmap rows show the five regional indices. Larger ΔNLL indicates a poorer fit at that biomass, and the dashed line marks $\Delta NLL = 1.92$. Select individual indices in the [interactive profile viewer](#).

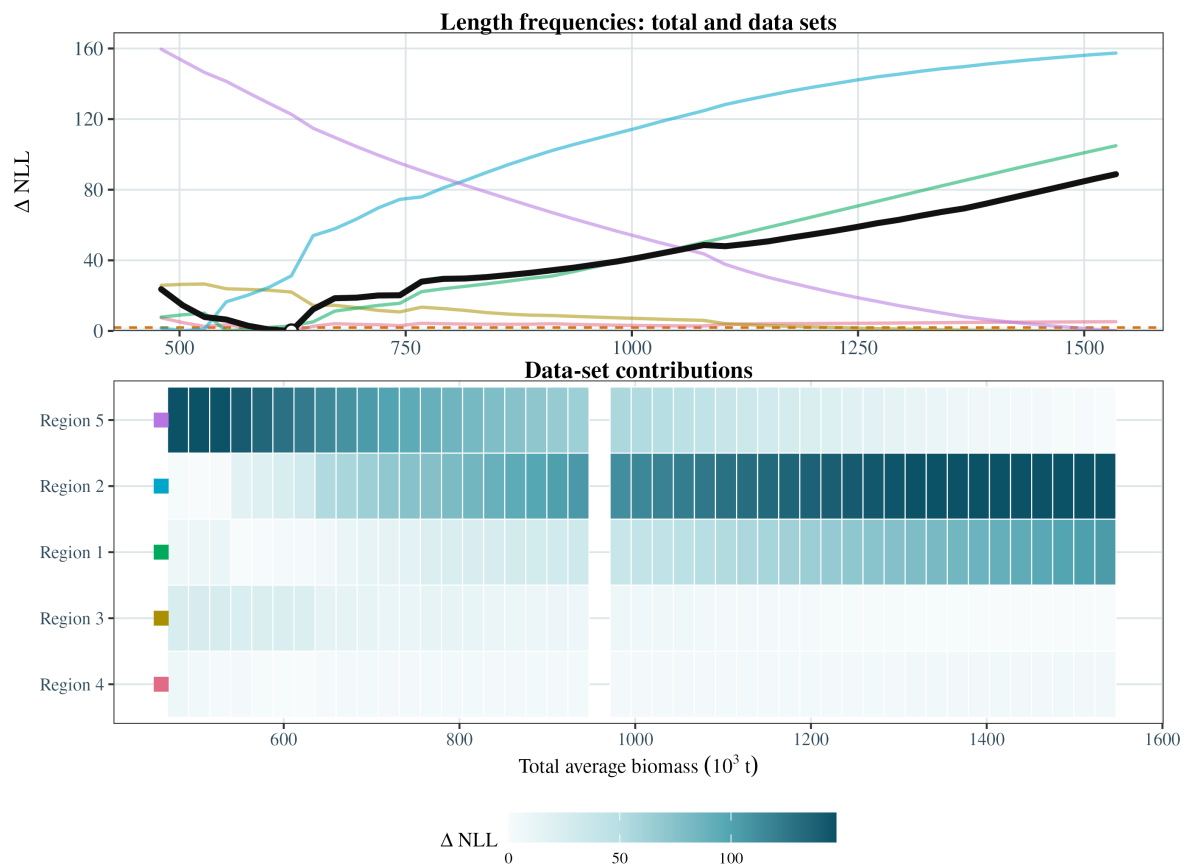


Figure 35: Length-frequency likelihood profile by model region. The black line is the total length-frequency contribution; the coloured lines and matching heatmap rows show the five regions. Larger ΔNLL indicates a poorer fit at that biomass, and the dashed line marks $\Delta NLL = 1.92$. Open a region in the [interactive profile viewer](#) for fishery-level detail.

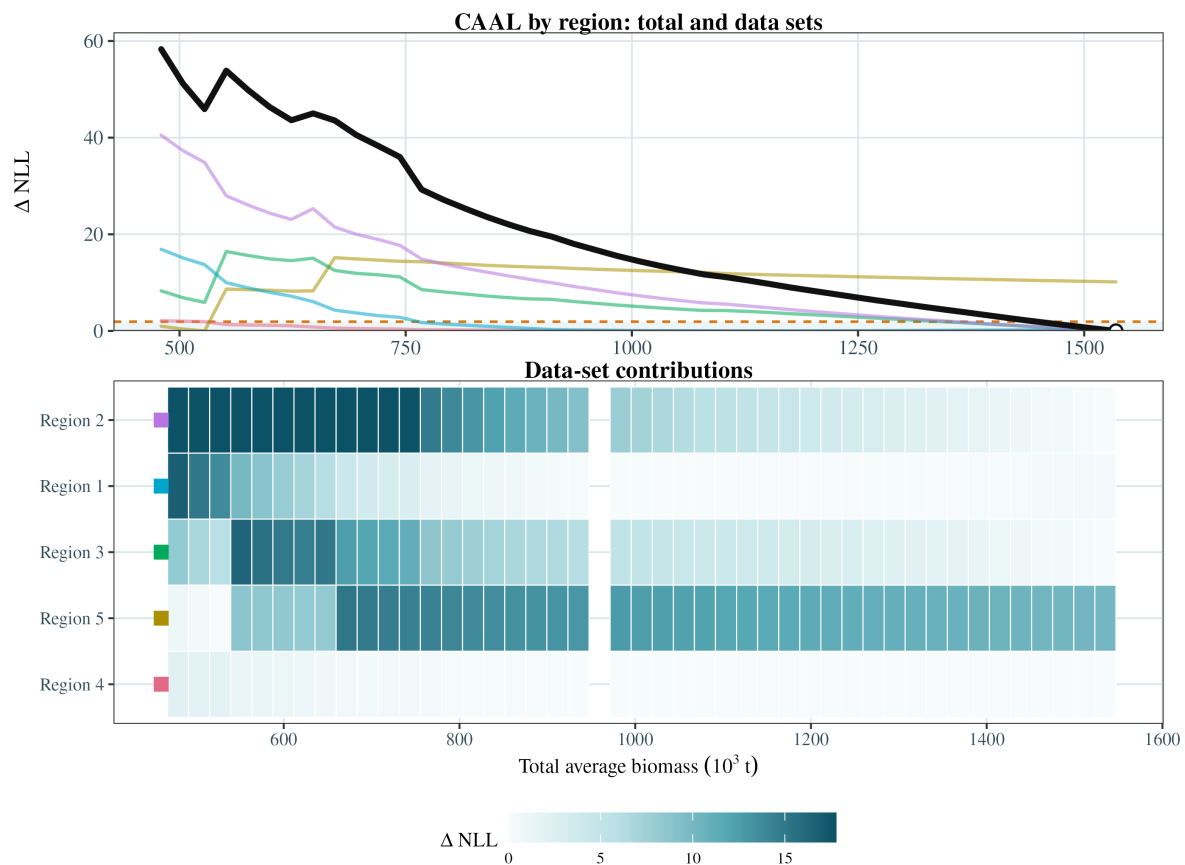


Figure 36: Conditional age-at-length likelihood profile by model region. The black line is the total CAAL contribution; the coloured lines and matching heatmap rows show the five regions. Larger ΔNLL indicates a poorer fit at that biomass, and the dashed line marks $\Delta NLL = 1.92$. Open a region in the [interactive profile viewer](#) for fishery-level detail.

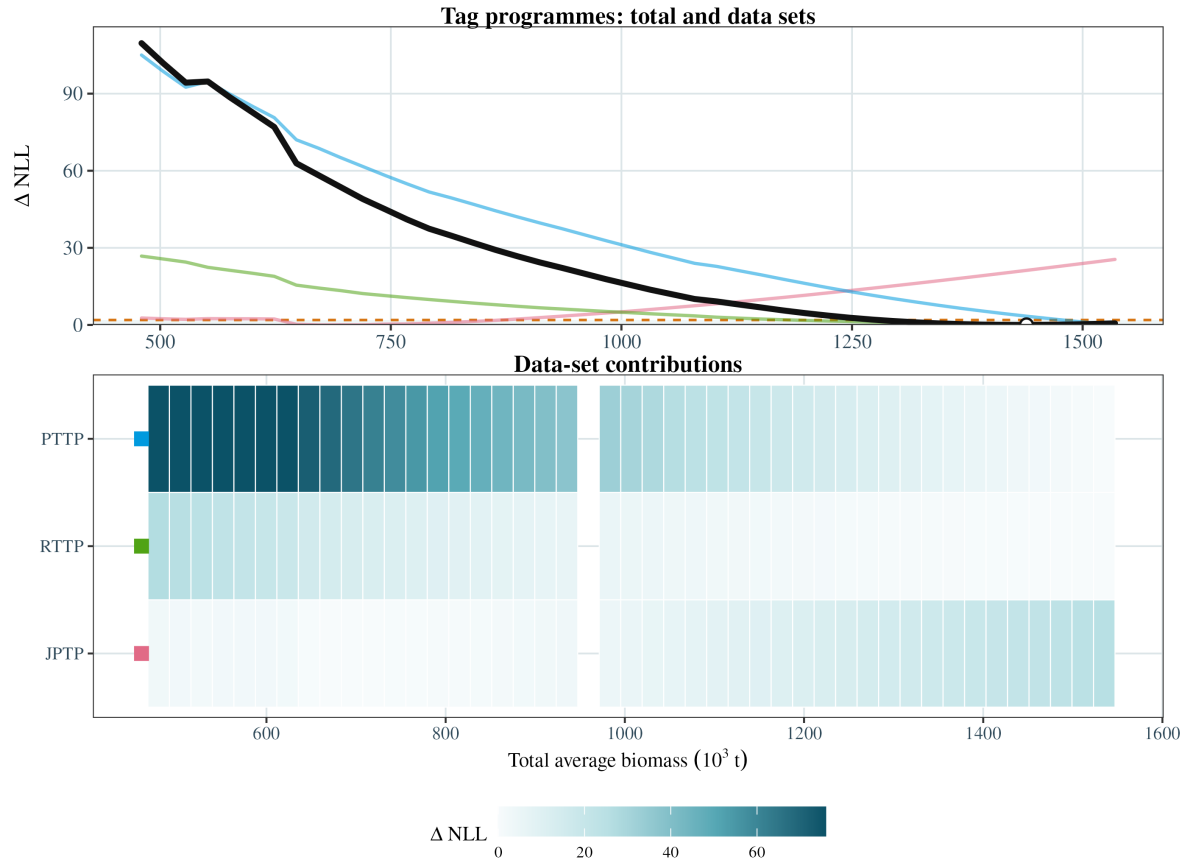


Figure 37: Tag likelihood profile by tagging programme. The black line is the total tag contribution; the coloured lines and matching heatmap rows show RTTP, PTTP and JPTP. Larger ΔNLL indicates a poorer fit at that biomass, and the dashed line marks $\Delta NLL = 1.92$. Open a programme in the [interactive profile viewer](#) for release-group detail.

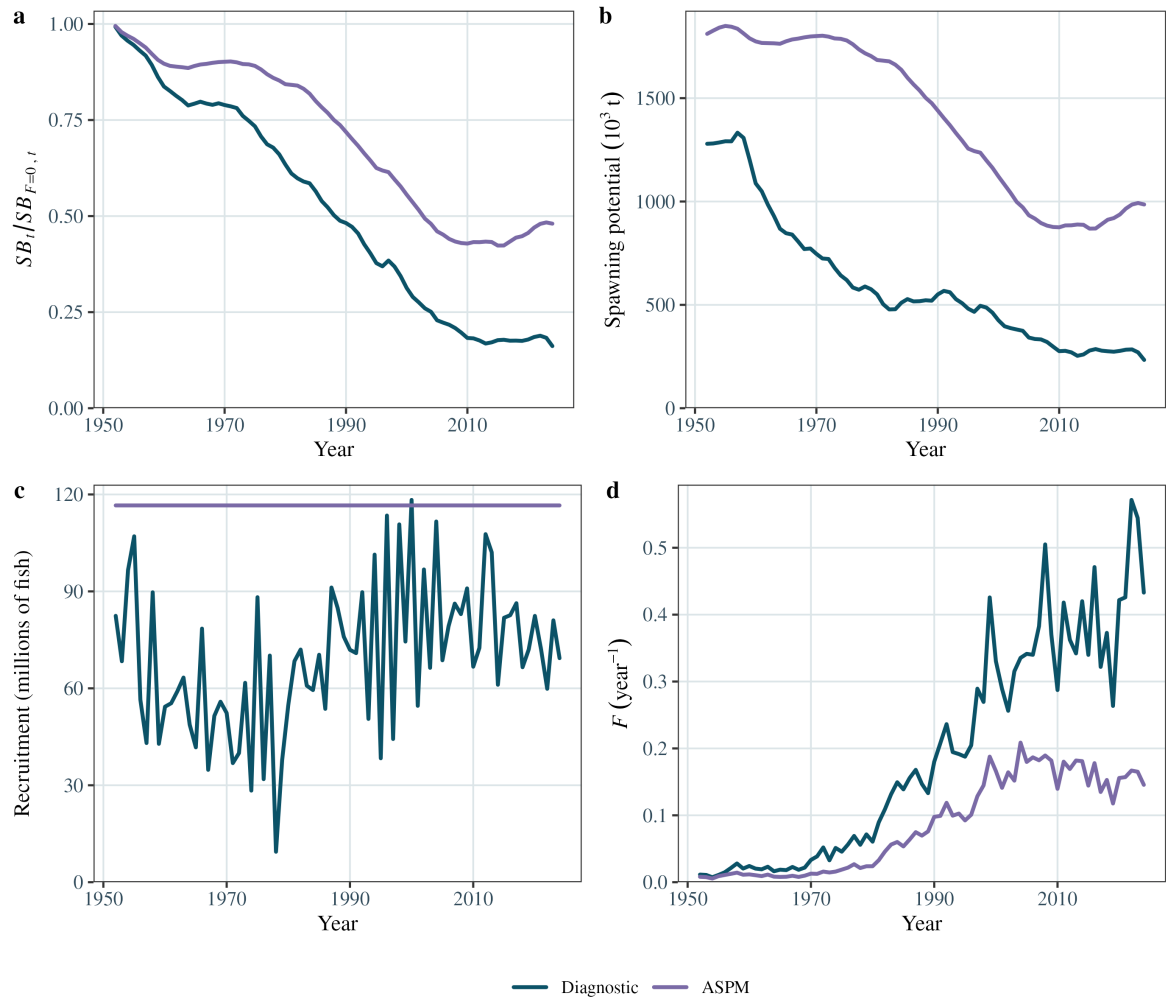


Figure 38: Comparison of the diagnostic model and constant-recruitment age-structured production model (ASPM): (a) dynamic spawning depletion, (b) spawning potential, (c) recruitment and (d) annual population-weighted fishing mortality.

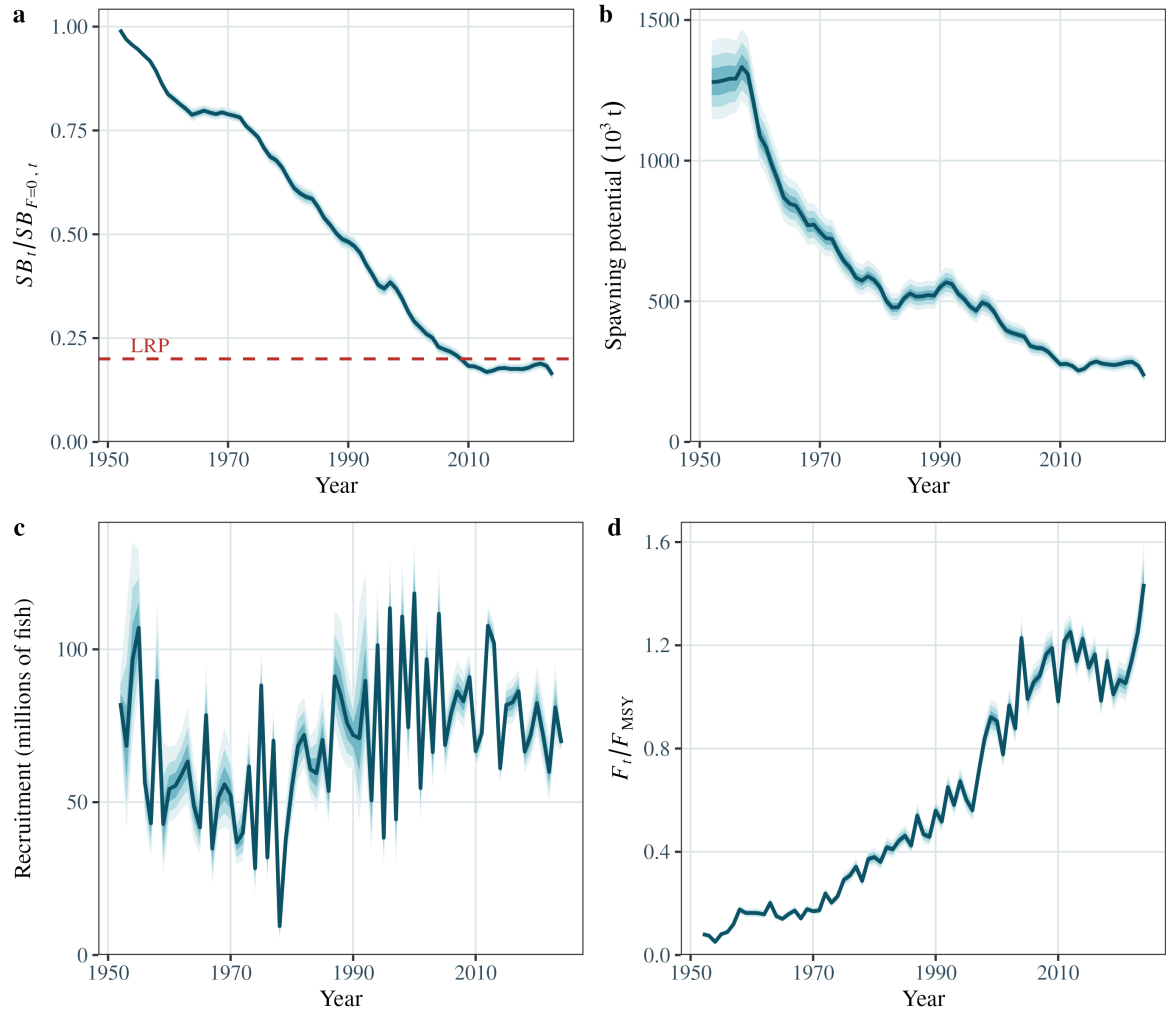


Figure 39: Diagnostic-model estimates of (a) dynamic spawning depletion, (b) spawning potential, (c) recruitment and (d) annual fishing mortality relative to F_{MSY} . Shading gives central 50%, 80% and 95% Hessian delta-method intervals.

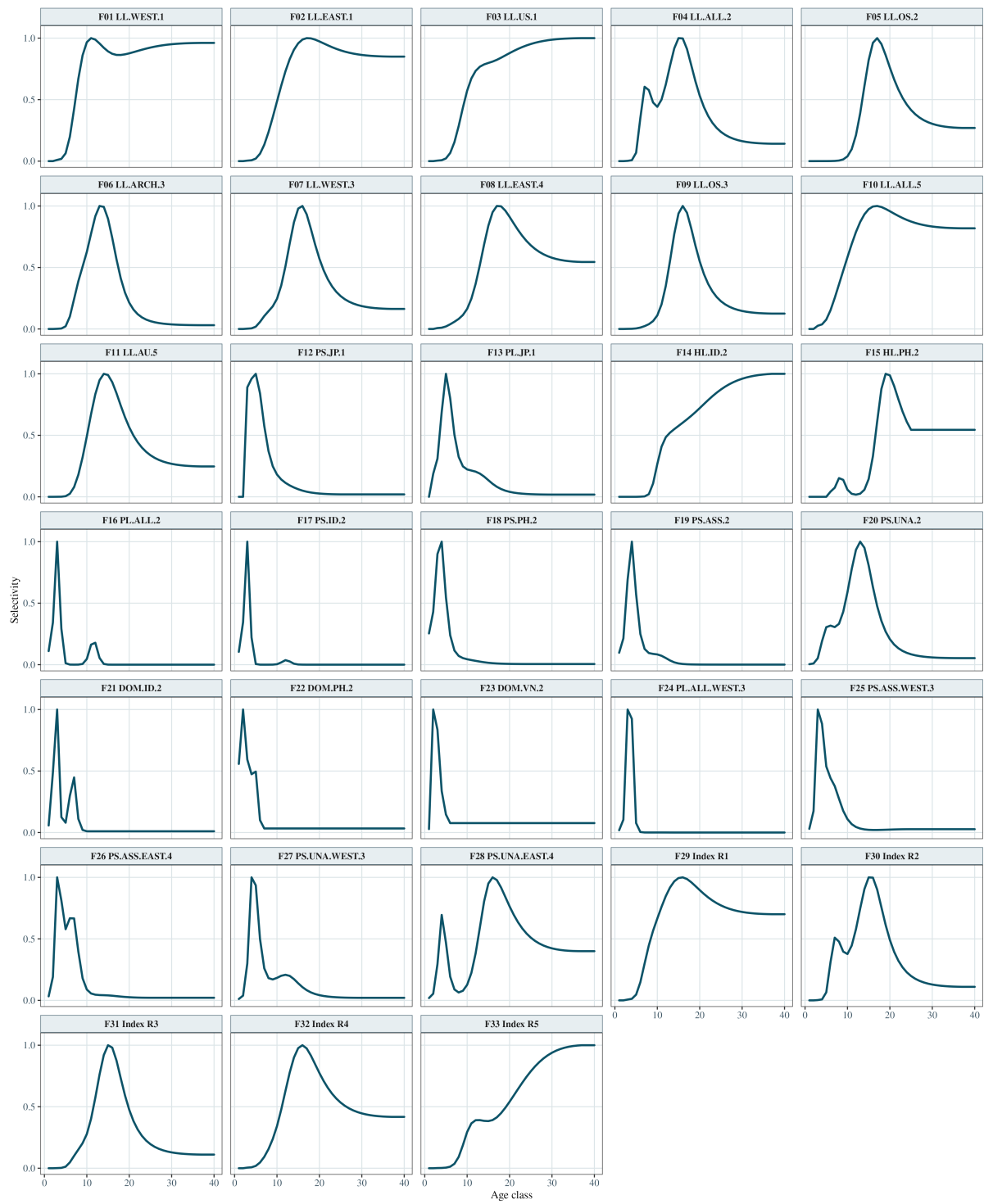


Figure 40: Estimated fishery selectivity at age for all 33 fisheries. Each panel shows one fitted fishery-specific cubic-spline curve; panel labels give the fishery ID and definition.

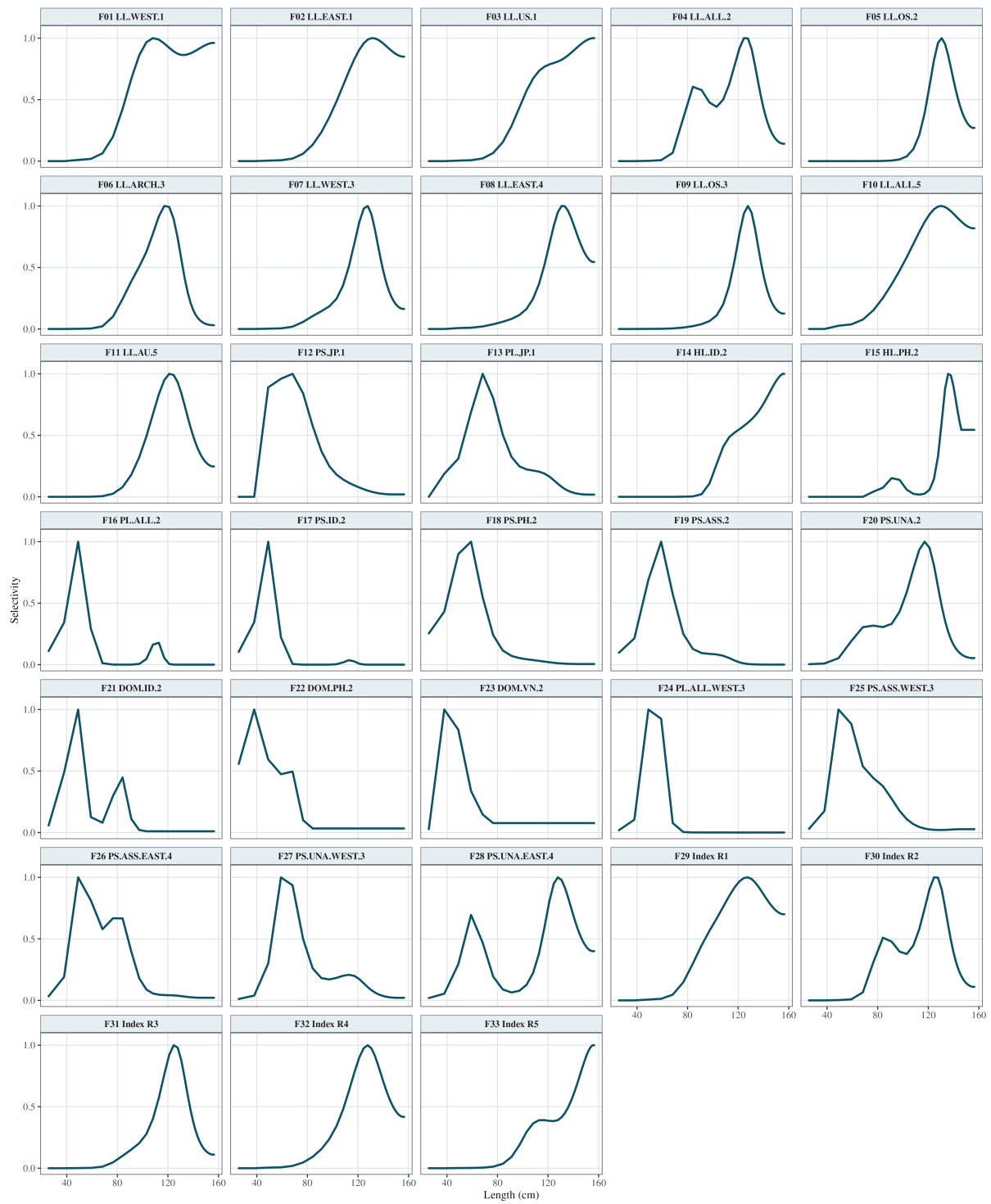


Figure 41: Estimated fishery selectivity at length for all 33 fisheries. Each panel shows one fishery, with age-specific selectivity mapped to the fitted mean length at age.

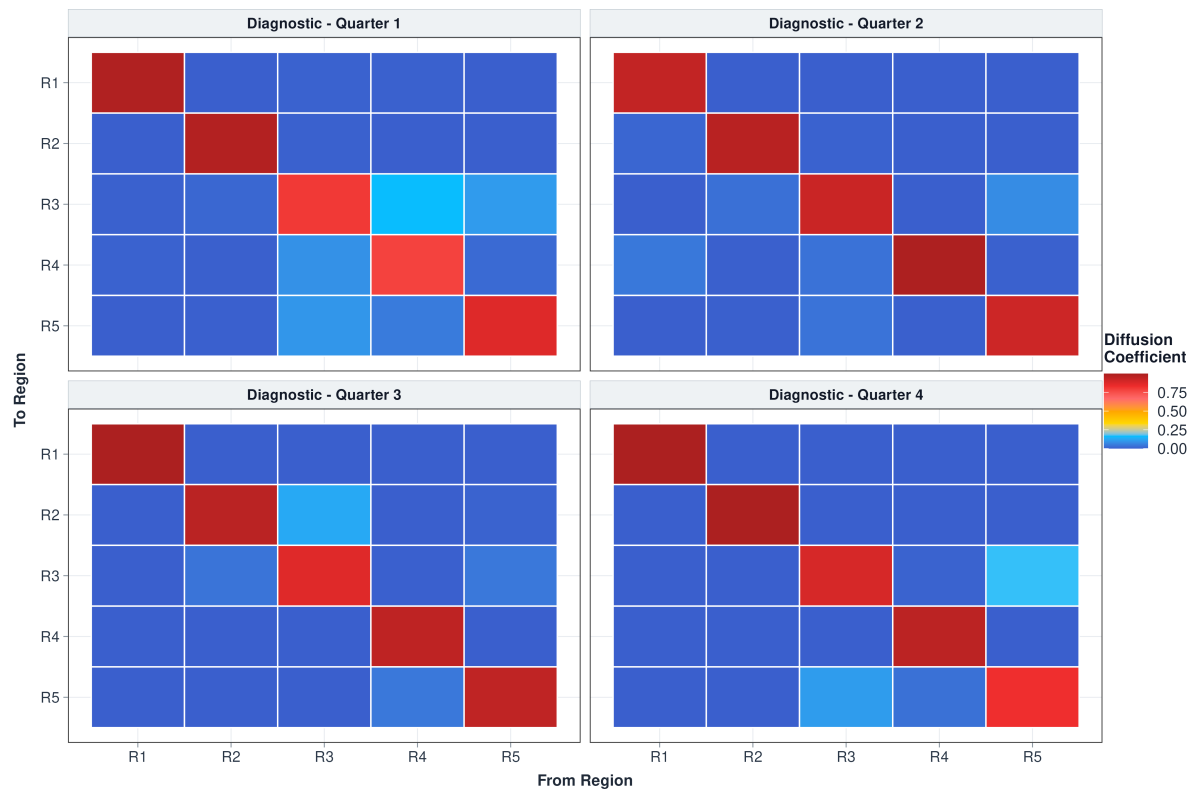


Figure 42: Estimated quarterly movement probabilities among the five model regions.



Figure 43: Estimated annual recruitment by model region and the pooled All regions series (last). The two-column layout provides additional panel height; panel y axes are free and begin at zero.

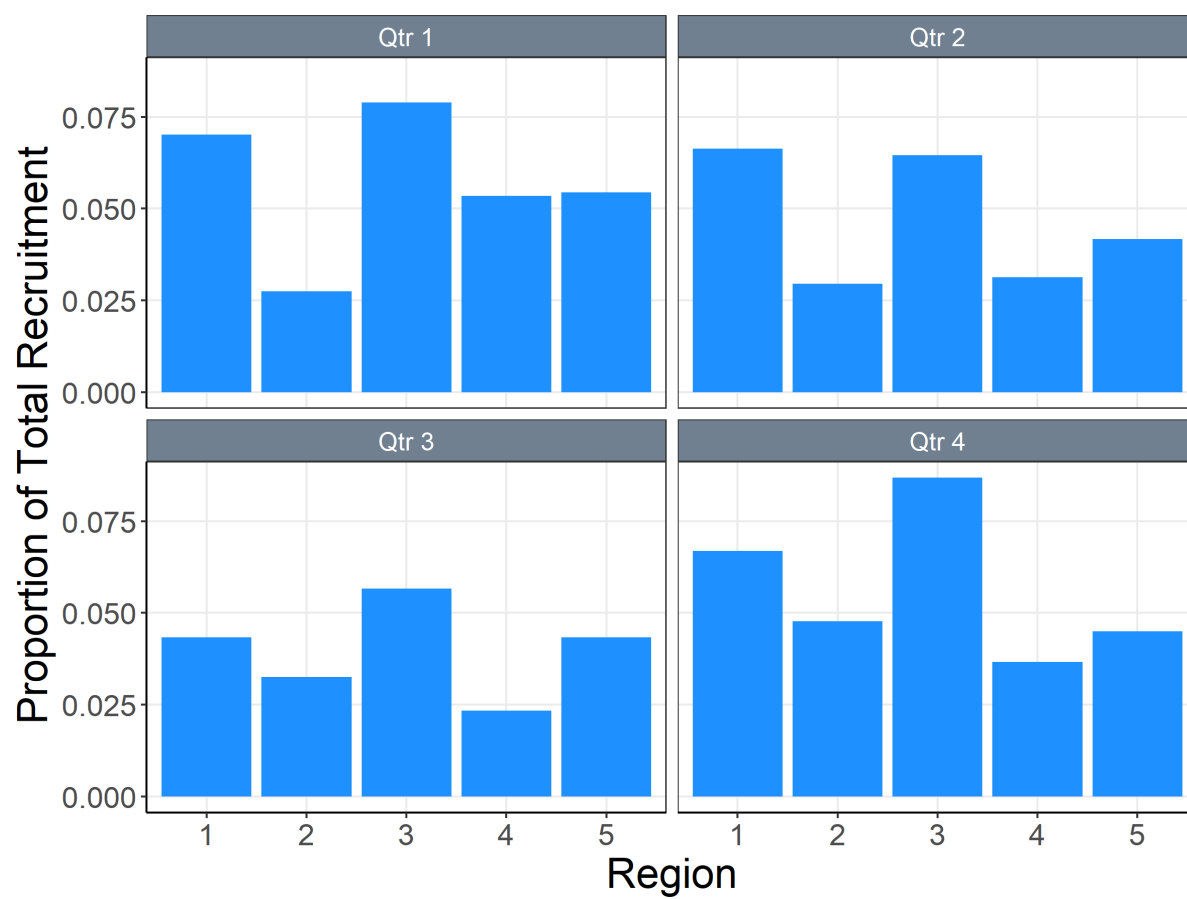


Figure 44: Estimated recruitment distribution by region and quarter.

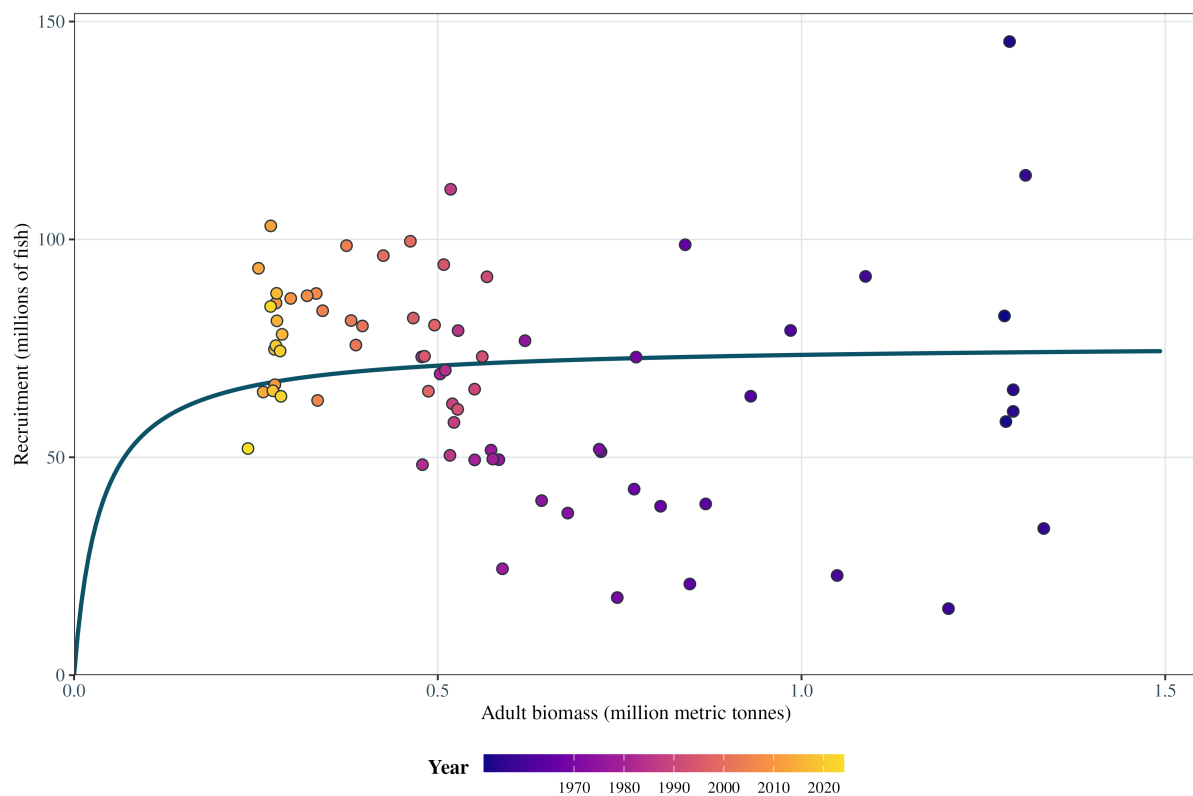


Figure 45: Fitted Beverton–Holt stock–recruitment relationship for the 2026 diagnostic model. Points show annual recruitment and adult-biomass estimates and are coloured by year. Open the [diagnostic report](#).

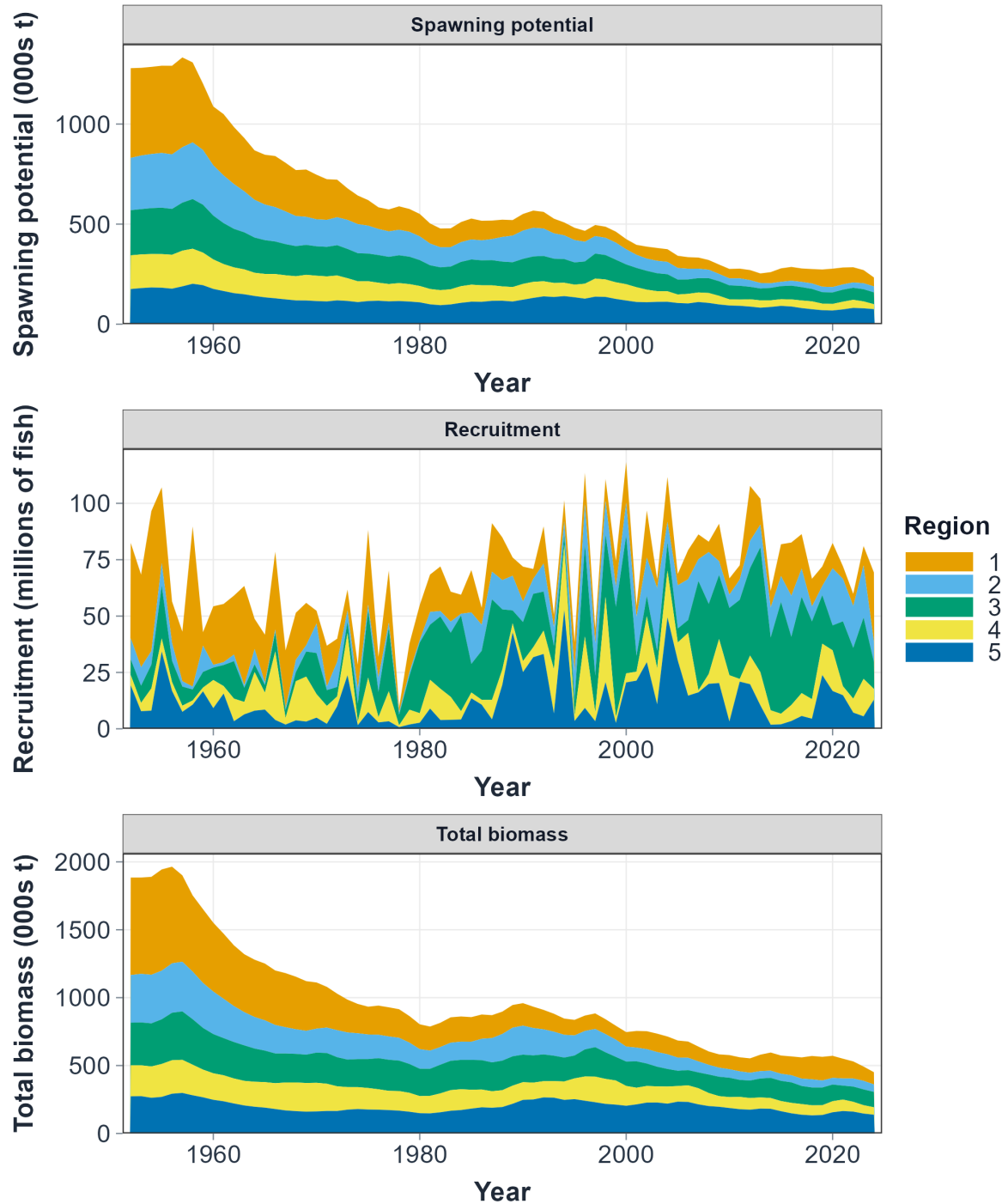


Figure 46: Time series of estimated annual spawning potential, recruitment and total biomass by model region for the diagnostic model, showing the relative proportions among regions. Note the data represent the averages of the quarterly model time steps for each year for spawning potential and total biomass and the sum of the quarterly recruitment estimates for annual recruitment.

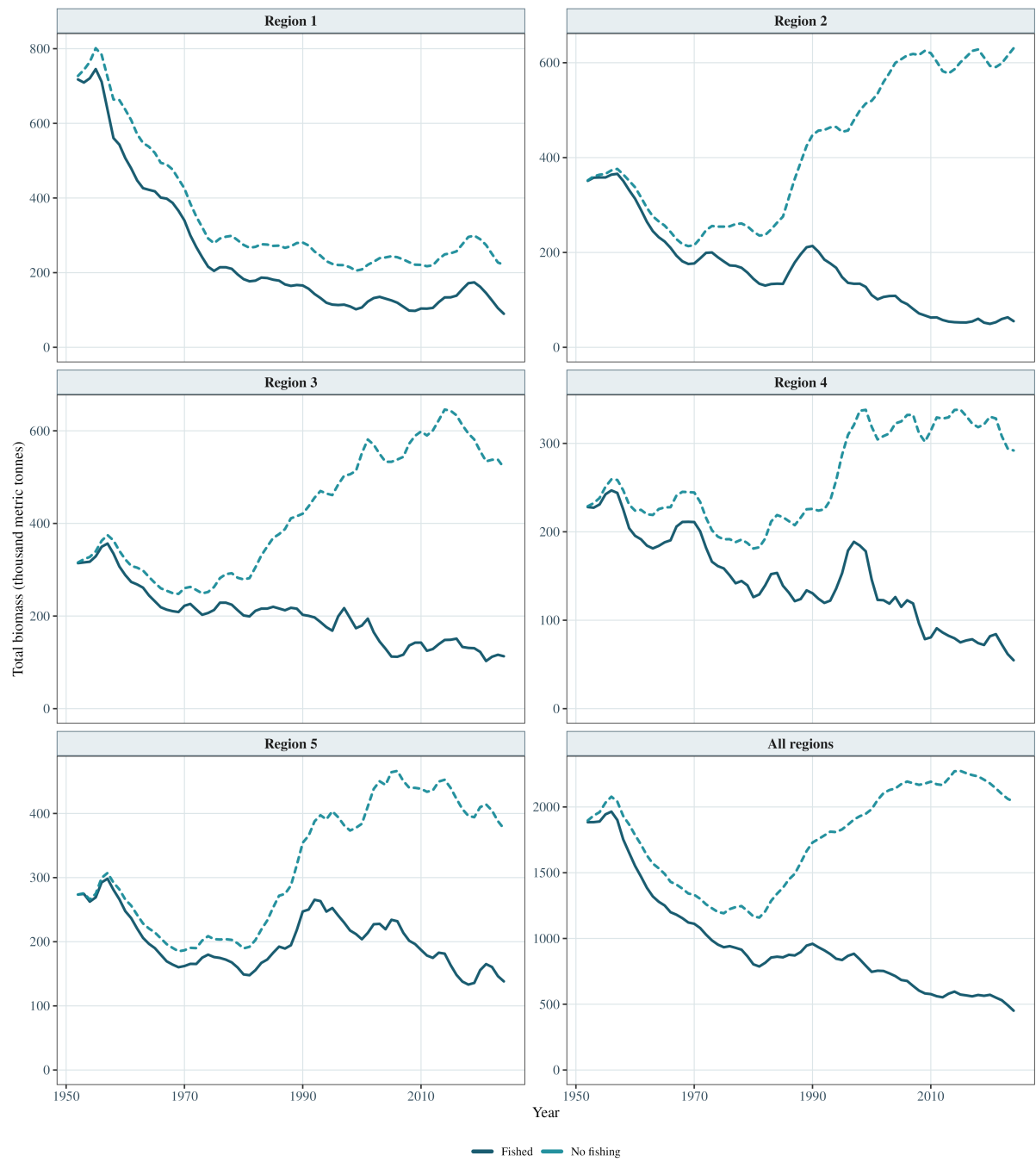


Figure 47: Estimated annual total biomass by model region and over all regions under the fitted fishing history (solid dark line) and the dynamic no-fishing counterfactual (dashed teal line). Annual values average the four quarterly estimates. The two-column layout provides additional panel height; panel y axes are free and begin at zero.

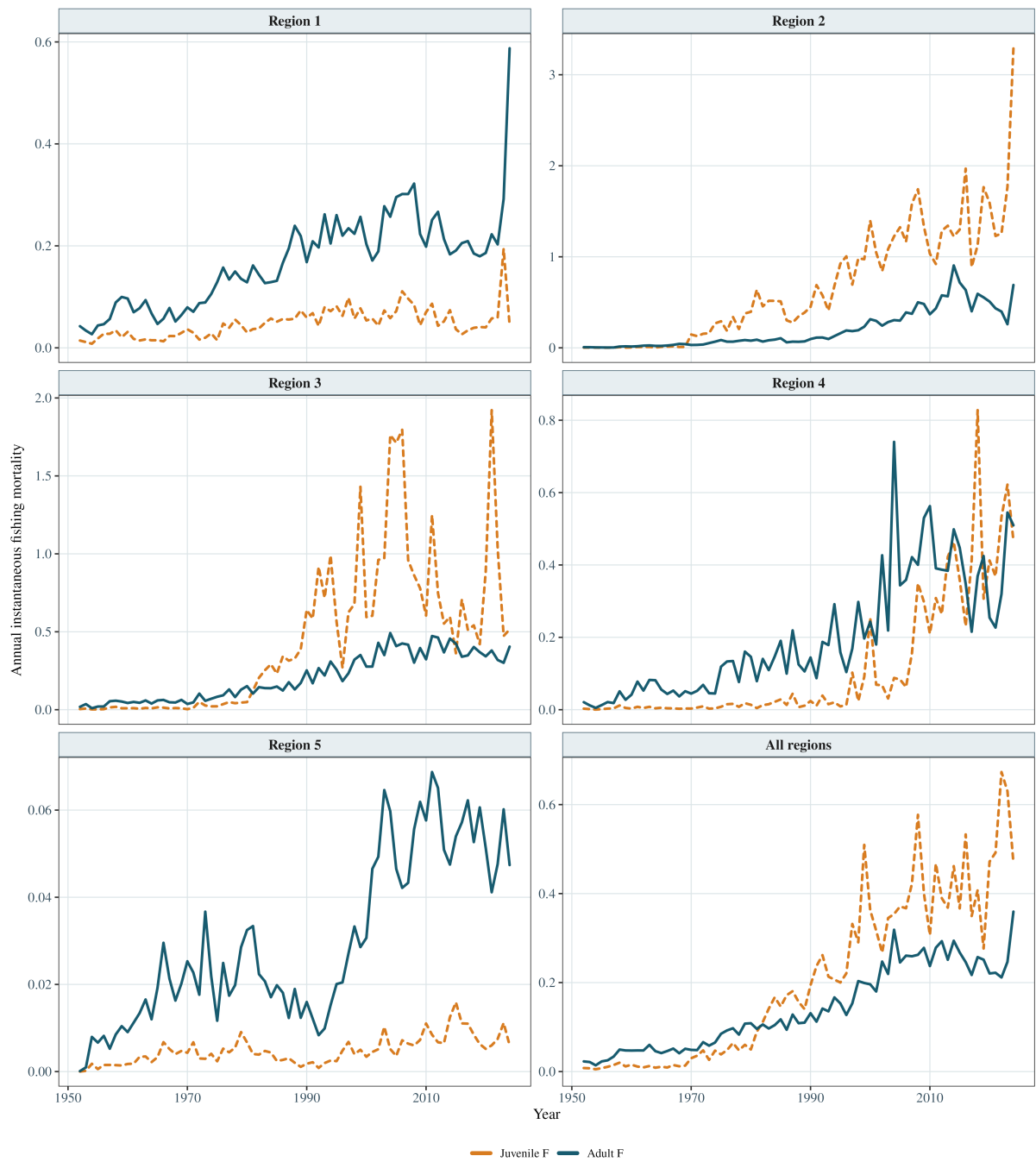


Figure 48: Annual juvenile (orange dashed) and adult (dark solid) instantaneous fishing mortality by model region and the pooled All regions series (last). The two-column layout provides additional panel height; panel y axes are free and begin at zero.

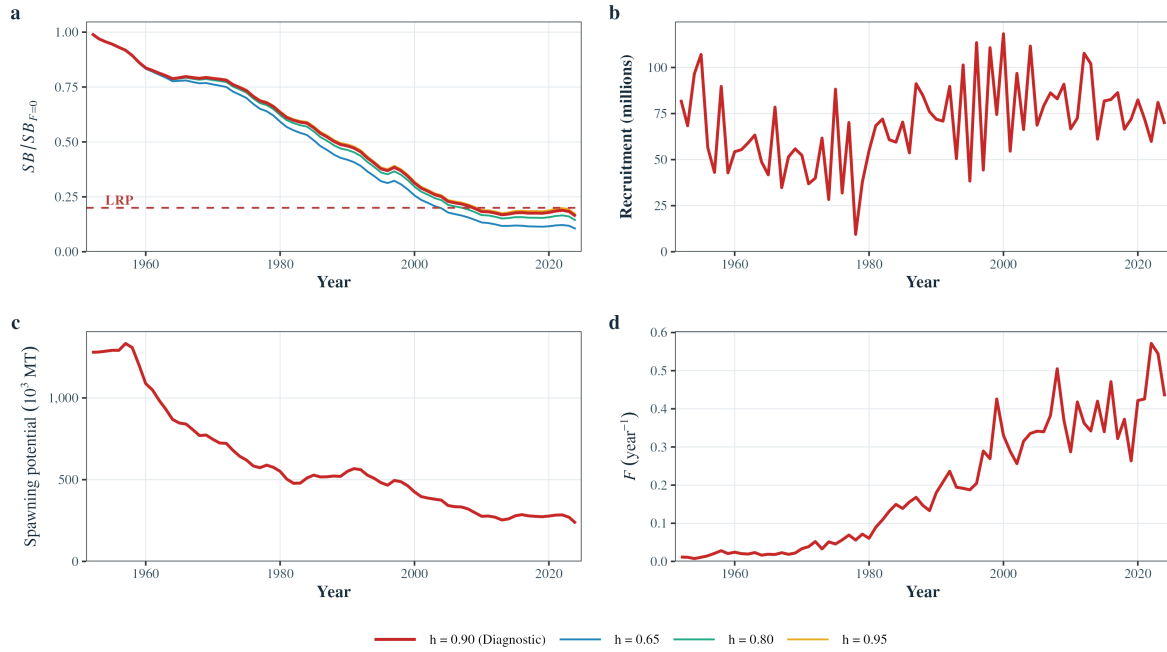


Figure 49: Estimated dynamic spawning depletion (Top left) and spawning potential (Bottom left), recruitment (Top right) and fishing mortality (Bottom right) for the one-off sensitivities from the 2026 diagnostic model for steepness of the stock recruitment relationship. Open the [interactive sensitivity viewer](#).

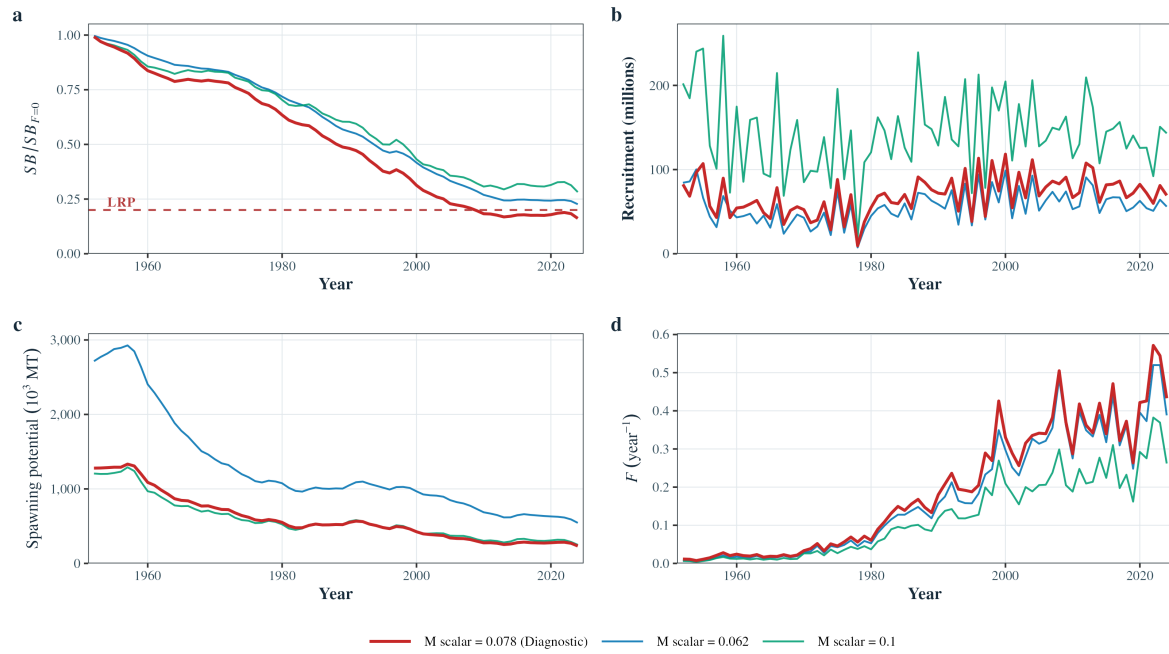


Figure 50: Estimated dynamic spawning depletion (Top left) and spawning potential (Bottom left), recruitment (Top right) and fishing mortality (Bottom right) for the one-off sensitivities from the 2026 diagnostic model for natural mortality scale (M at age 40 quarters). Open the [interactive sensitivity viewer](#).

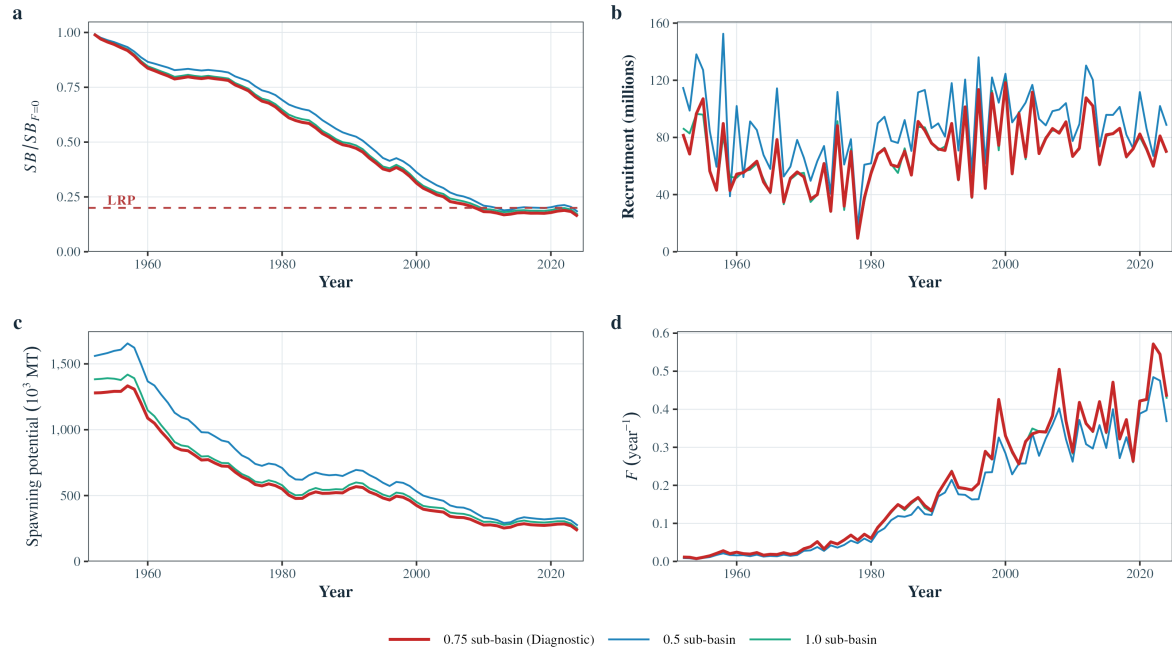


Figure 51: Estimated dynamic spawning depletion (Top left) and spawning potential (Bottom left), recruitment (Top right) and fishing mortality (Bottom right) for the one-off sensitivities from the 2026 diagnostic model for conditional age at length (CAAL) data weighting. Open the [interactive sensitivity viewer](#).

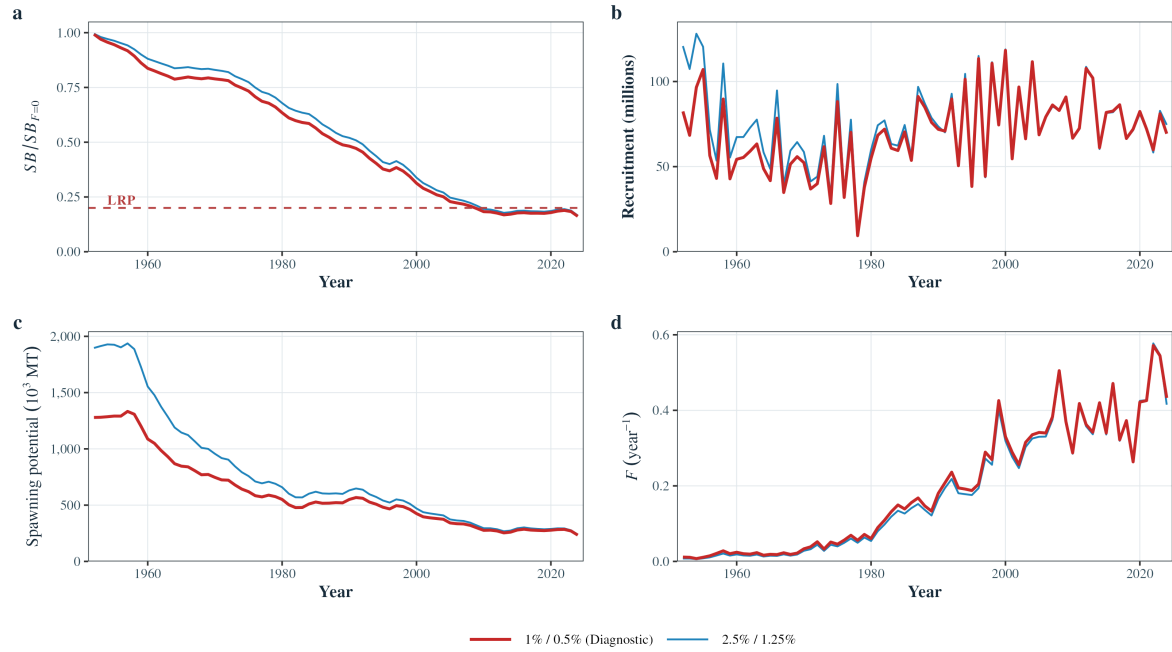


Figure 52: Estimated dynamic spawning depletion (Top left) and spawning potential (Bottom left), recruitment (Top right) and fishing mortality (Bottom right) for the one-off sensitivities from the 2026 diagnostic model for effort creep adjustment to CPUE indices. Open the [interactive sensitivity viewer](#).

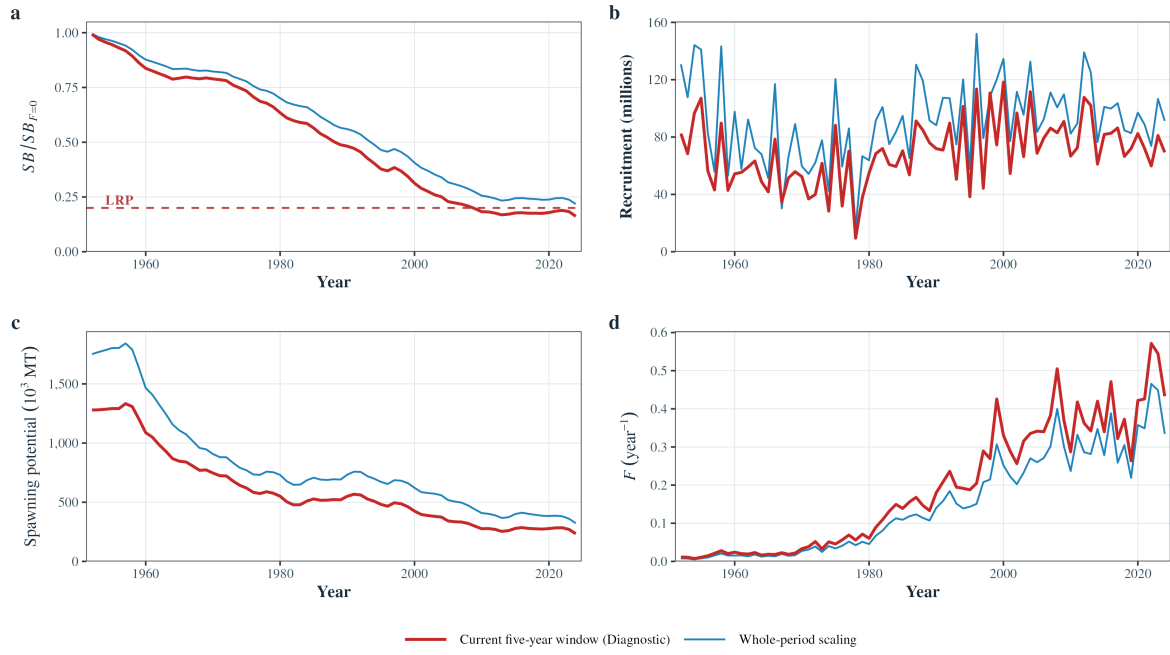


Figure 53: Dynamic spawning depletion (top left), spawning potential (bottom left), recruitment (top right) and fishing mortality (bottom right) for one-off sensitivities to the global-CPUE period used in the regional-scaling penalty. The diagnostic model uses the 1965–1969 period of broad spatial coverage recommended by PAW; the alternative uses the full 1952–2024 series. Open the [interactive sensitivity viewer](#).

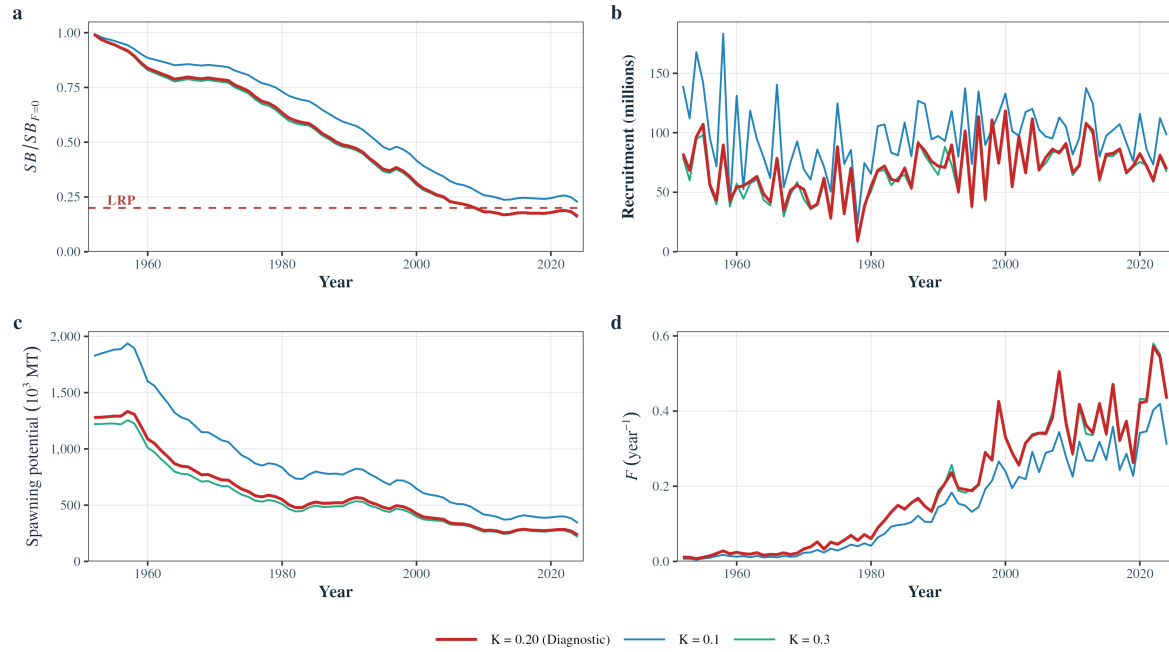


Figure 54: Estimated dynamic spawning depletion (Top left) and spawning potential (Bottom left), recruitment (Top right) and fishing mortality (Bottom right) for the one-off sensitivities from the 2026 diagnostic model on the K-statistic for assigning release group specific mixing periods - higher value assume less conservative shorter mixing periods, lower value assume more conservative longer mixing periods. Open the [interactive sensitivity viewer](#).

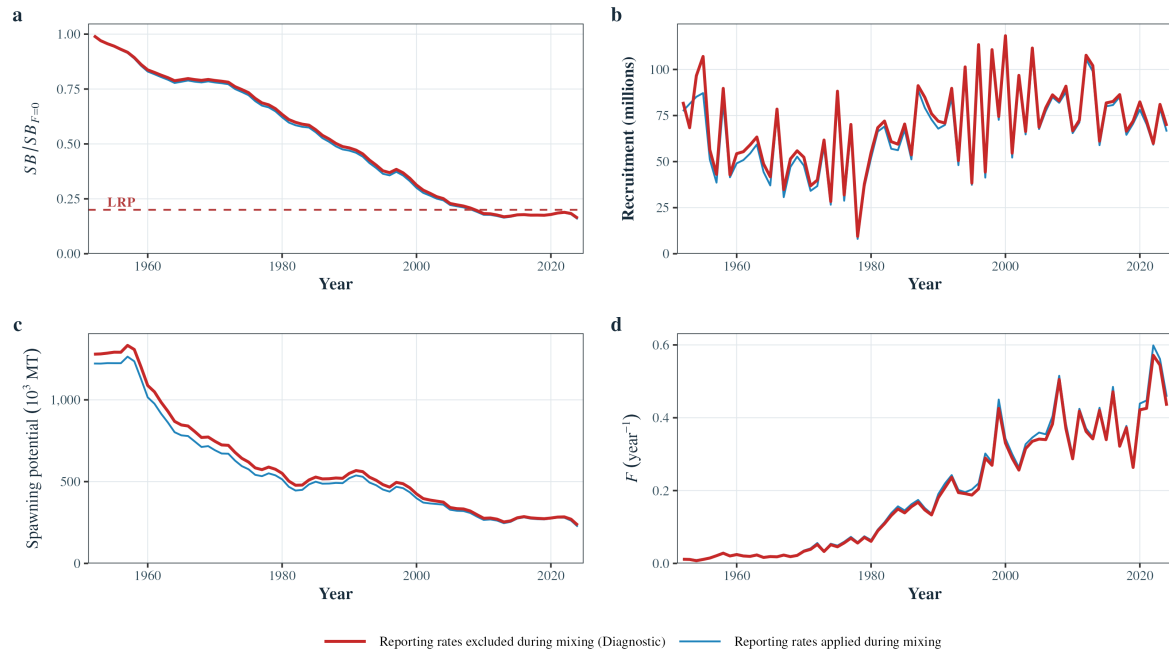


Figure 55: Estimated dynamic spawning depletion (Top left) and spawning potential (Bottom left), recruitment (Top right) and fishing mortality (Bottom right) for the one-off sensitivities from the 2026 diagnostic model on the option to activate or deactivate reporting rate adjustments during tag mixing periods. Open the [interactive sensitivity viewer](#).

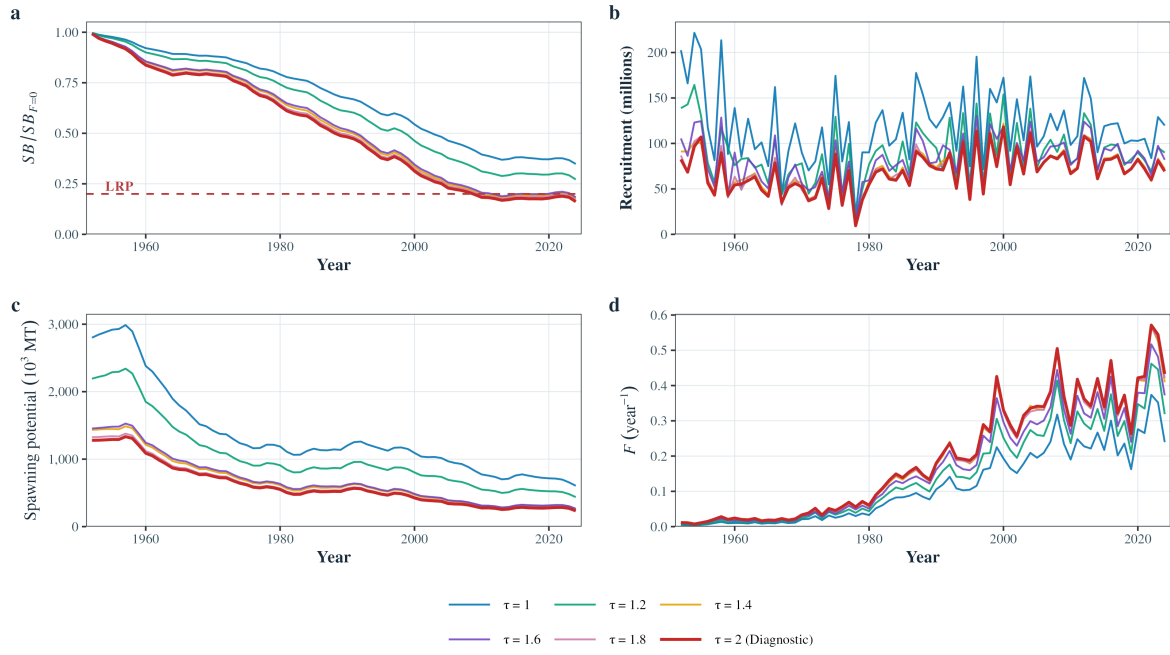


Figure 56: Dynamic spawning depletion (top left), spawning potential (bottom left), recruitment (top right) and fishing mortality (bottom right) for one-off sensitivities to tag overdispersion τ . Lower values assign greater likelihood weight to tag data; higher values allow greater extra-Poisson variation. Open the [interactive sensitivity viewer](#).

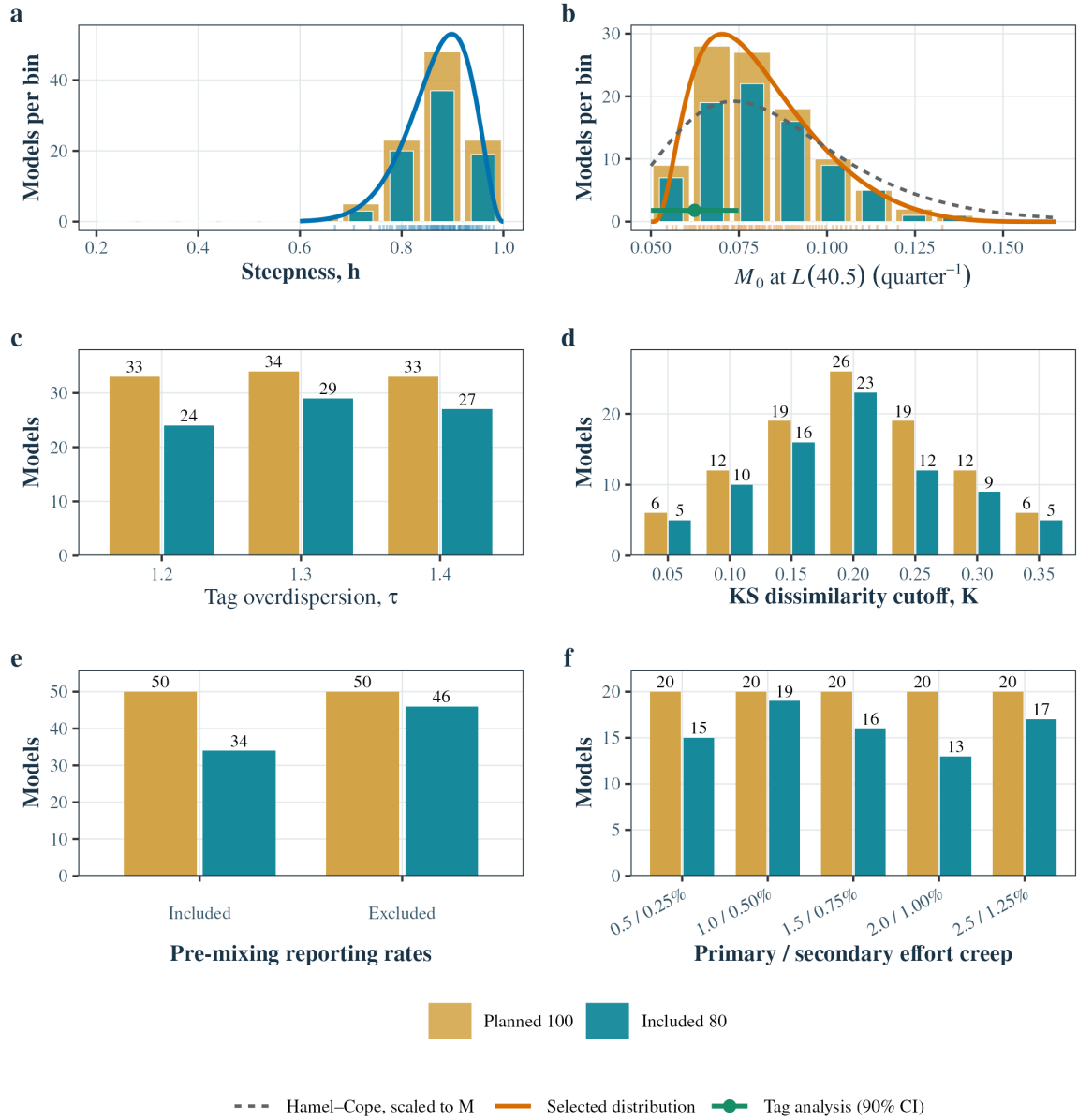


Figure 57: Planned and retained ensemble model inputs: steepness (a), natural mortality at the reference length (age 40 quarters) (b), tag overdispersion (c), tag-mixing cutoff (d), pre-mixing reporting treatment (e), and effort creep (f). Gold shows the 100 planned configurations and teal shows the 80 retained models. Curves in panels a–b show the specified continuous input distributions; the green point and line in panel b are the tag-based estimate and 90% confidence interval from Ducharme-Barth et al. (2026). Open the [interactive ensemble viewer](#).

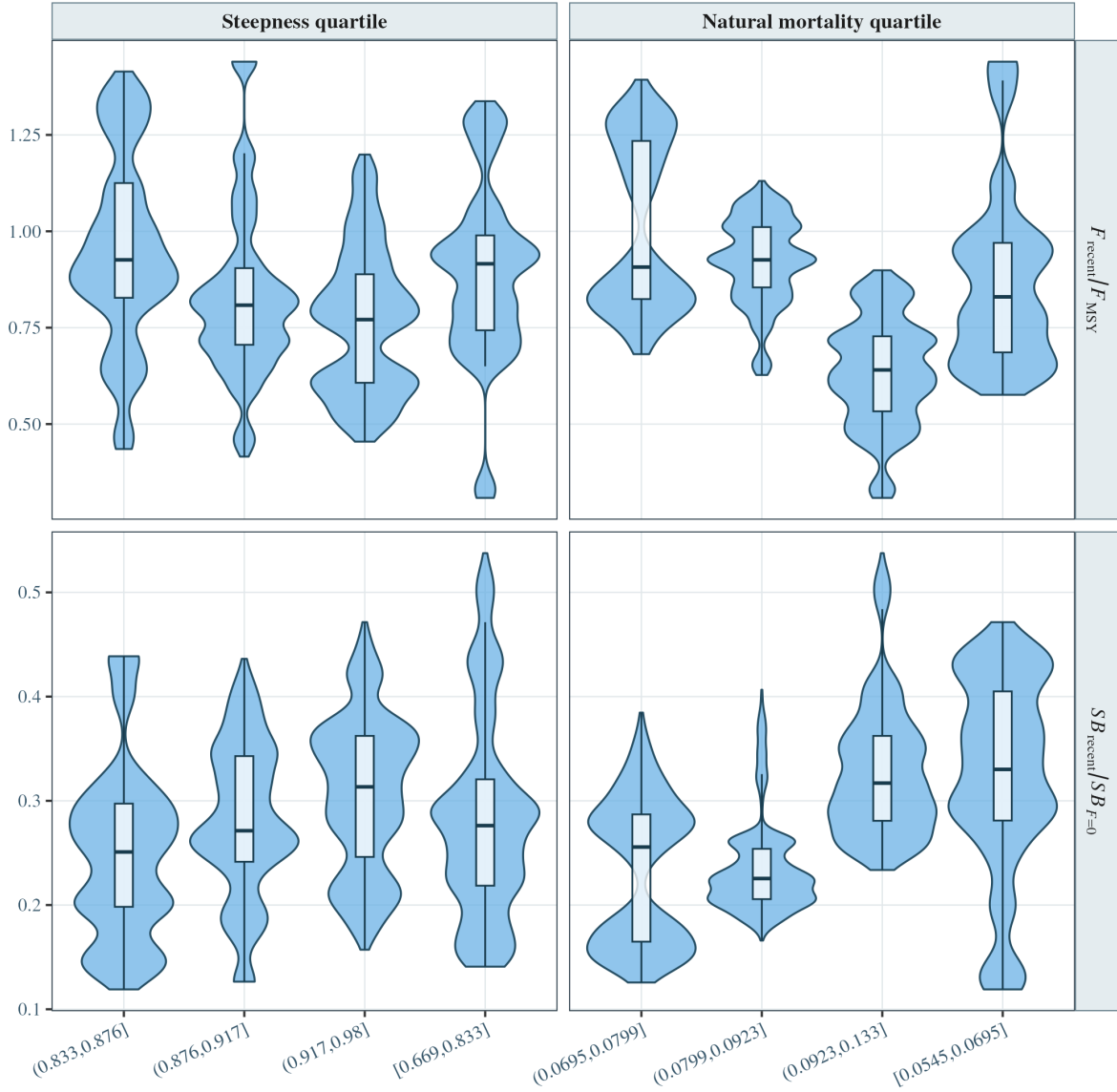


Figure 58: All-model uncertainty distributions of F_{recent}/F_{MSY} (top) and $SB_{recent}/SB_{F=0}$ (bottom), grouped by quartiles of steepness and natural mortality at the reference length. Violin width represents density and the inset box shows the interquartile range and median. Each model has equal total weight; PDH models contribute joint Hessian draws and Near-PDH models contribute their central estimates. Because all ensemble axes vary jointly, differences among groups are descriptive and not one-factor causal effects. Open the [interactive ensemble viewer](#).

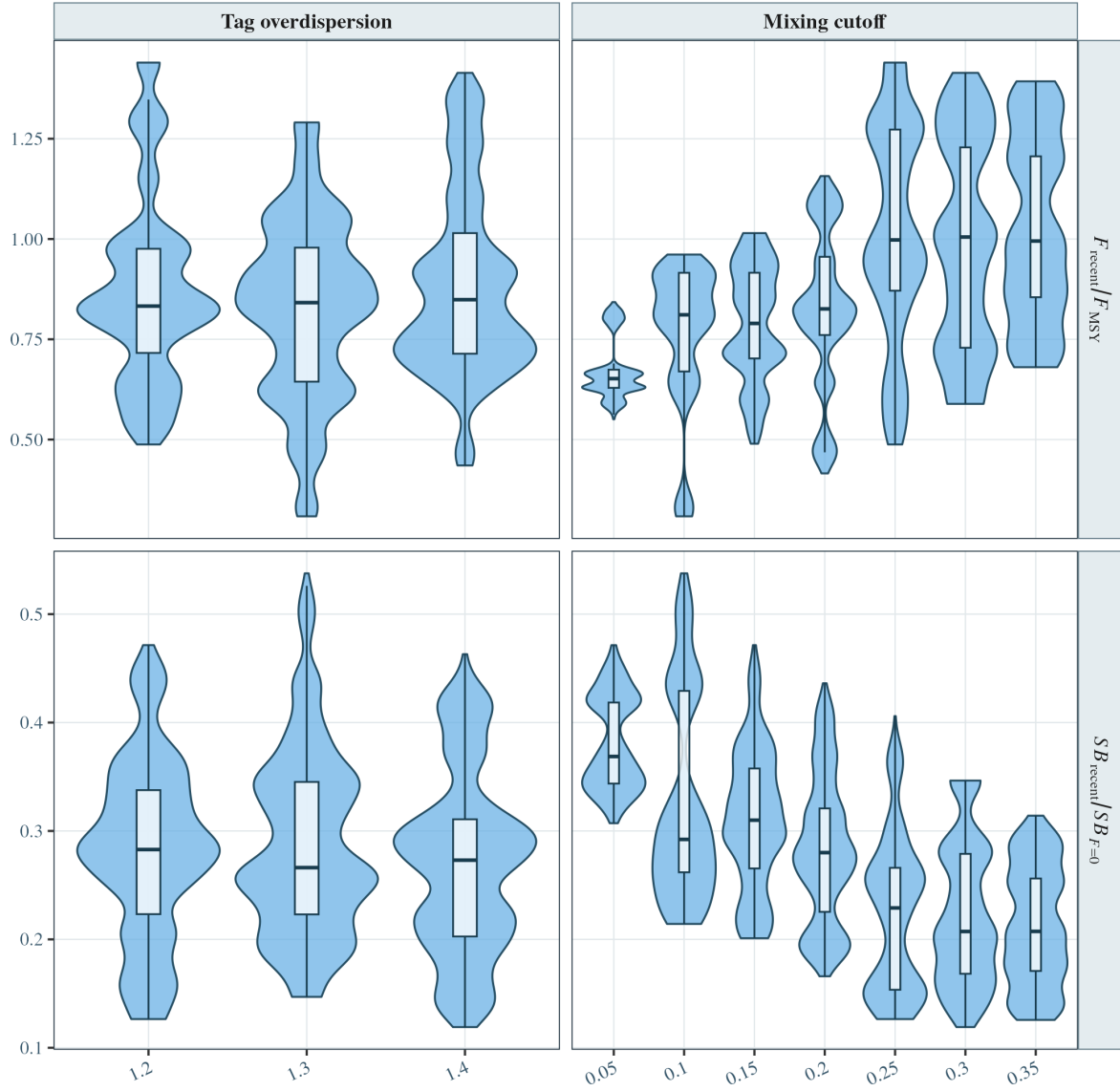


Figure 59: All-model uncertainty distributions of F_{recent}/F_{MSY} (top) and $SB_{recent}/SB_{F=0}$ (bottom), grouped by fixed tag overdispersion and tag-mixing cutoff. Each assessment model has equal total weight. Violins show density and boxes show the median and interquartile range; comparisons are marginal descriptions of a jointly varying ensemble. Open the [interactive ensemble viewer](#).

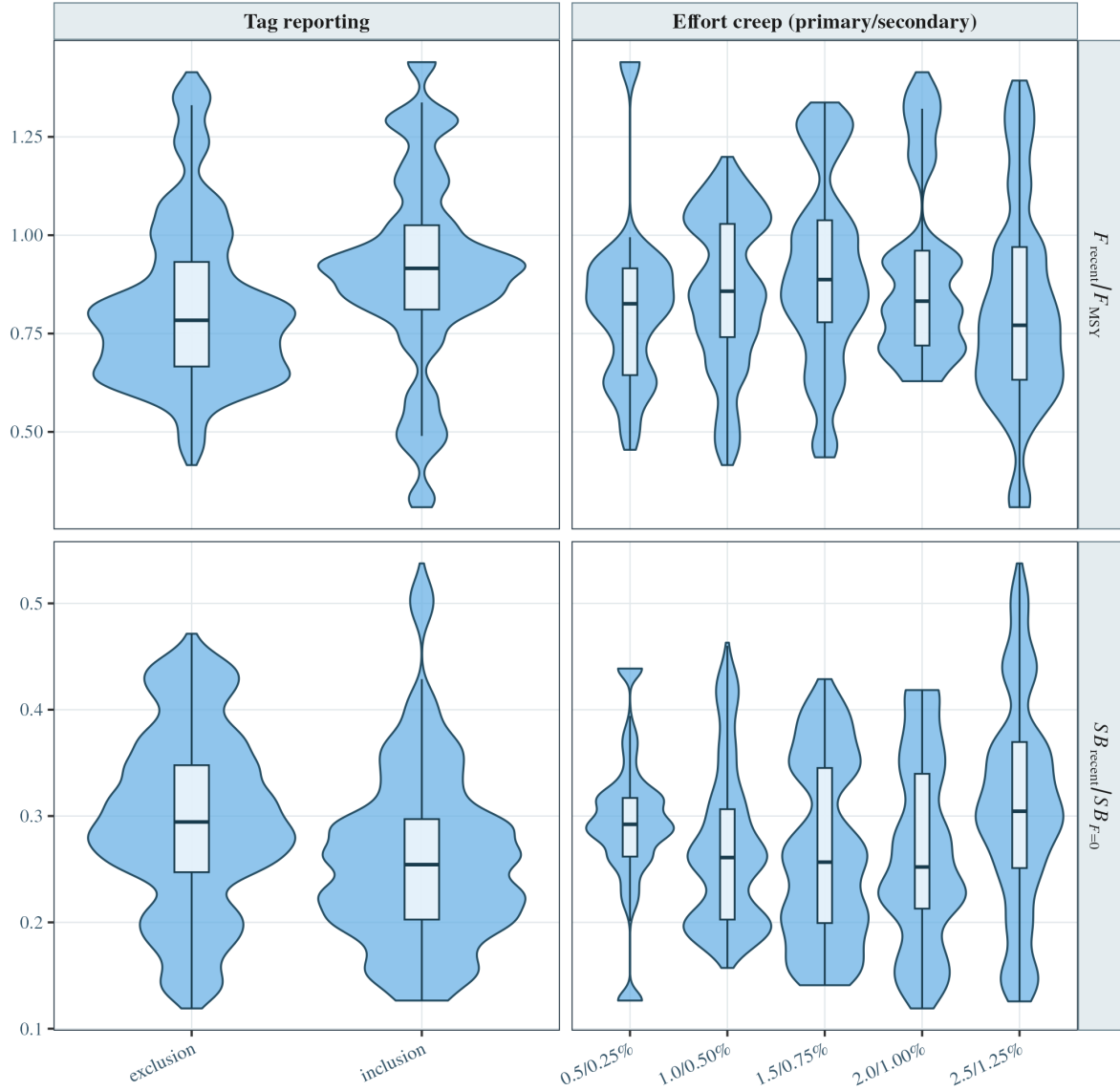


Figure 60: All-model uncertainty distributions of F_{recent}/F_{MSY} (top) and $SB_{recent}/SB_{F=0}$ (bottom), grouped by tag-reporting treatment and paired annual effort-creep rates. Each assessment model has equal total weight. Violins show density and boxes show the median and interquartile range; comparisons are marginal descriptions rather than controlled one-factor sensitivities. Open the [interactive ensemble viewer](#).

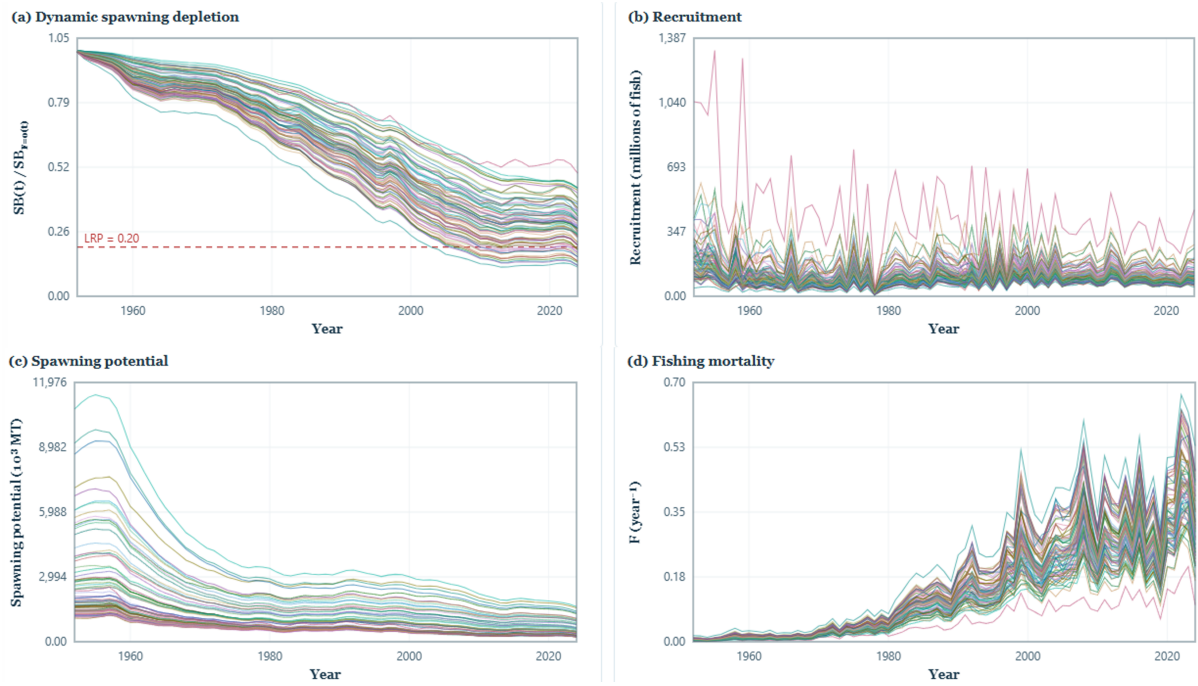


Figure 61: Estimated dynamic spawning depletion (top left), spawning potential (bottom left), recruitment (top right) and fishing mortality (bottom right) for the 80 individual models retained for the uncertainty ensemble. Open the [interactive ensemble viewer](#).

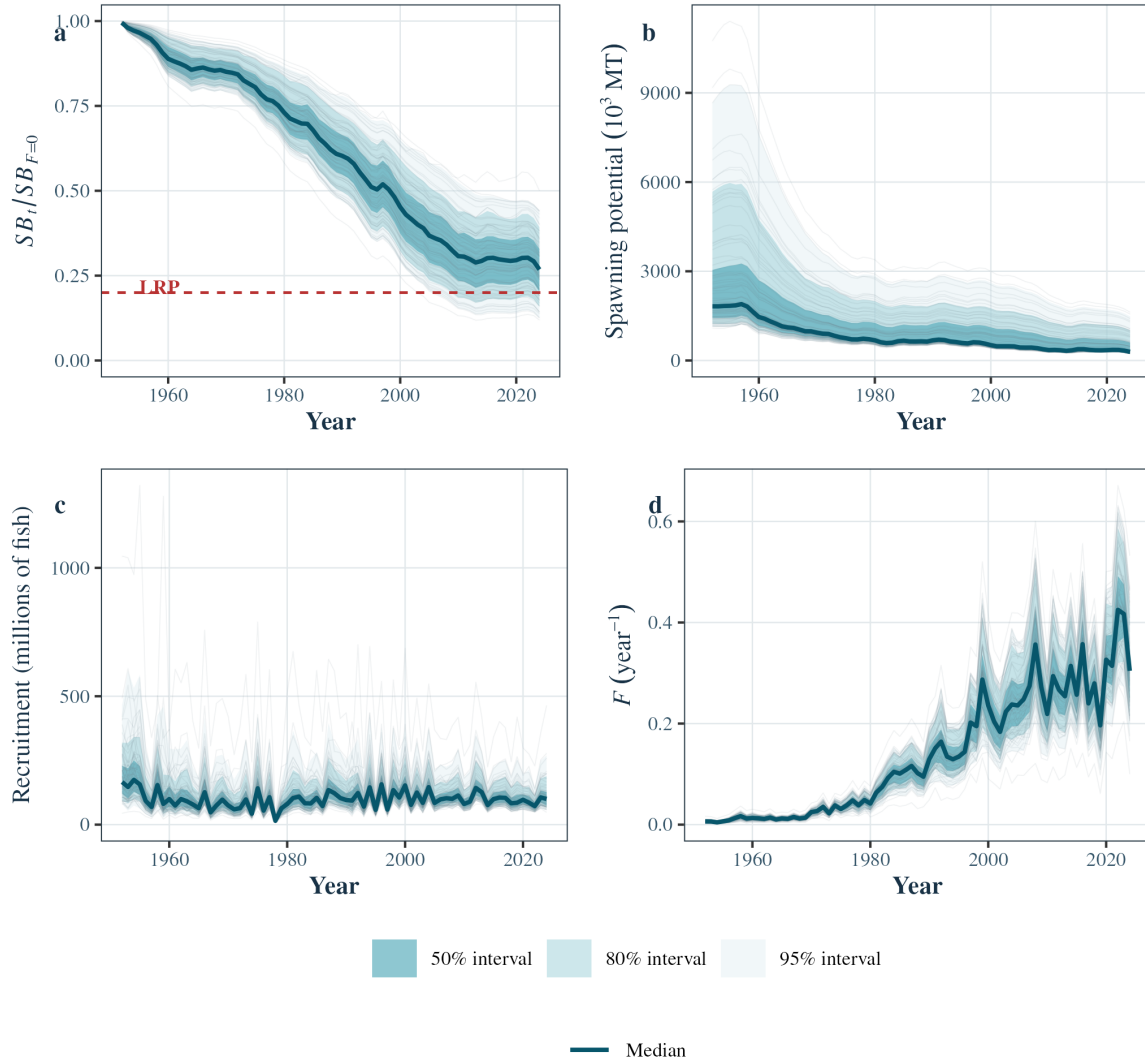


Figure 62: Ribbon plots of estimated dynamic spawning depletion (Top left) and spawning potential (Top right), recruitment (Bottom left) and fishing mortality (Bottom right) for the 80 models retained for the uncertainty ensemble including the estimation uncertainty. The median and pointwise central 50%, 80% and 95% intervals use the uncertainty samples available for each quantity. Grey lines show central model trajectories. Open the [interactive ensemble viewer](#).

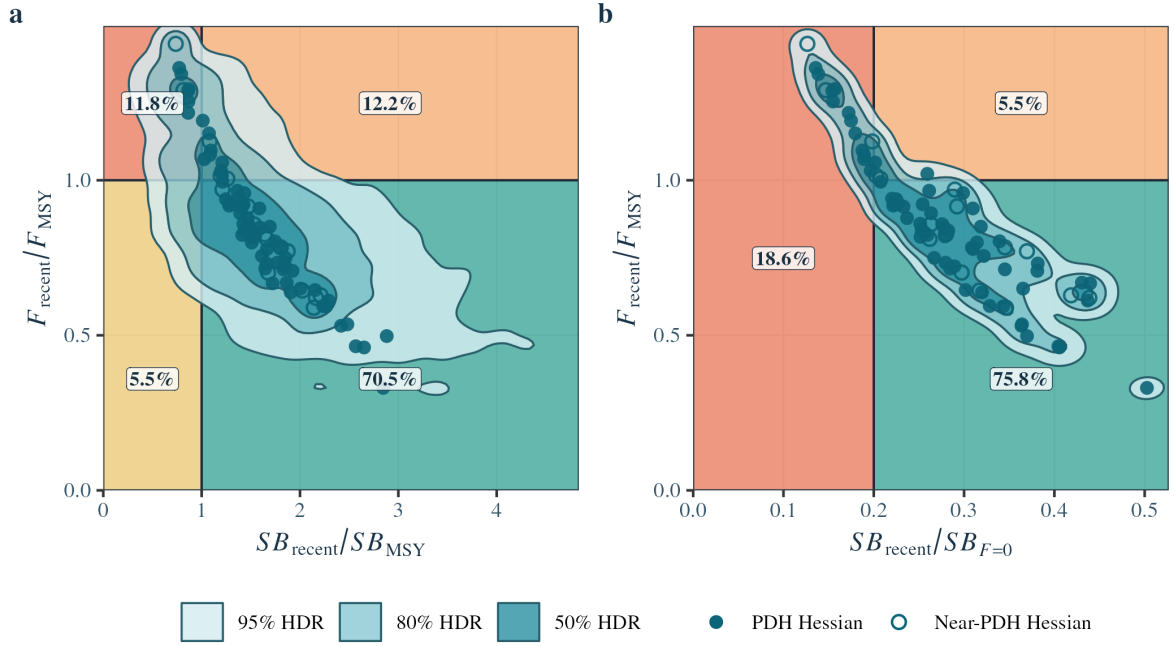


Figure 63: Stock status based on $SB_{recent}/SB_{F=0}$ (SBrecent 2021–2024) and F_{recent}/F_{MSY} (Frecent 2020–2023). Panel (a) Kobe: is relative to SB_{recent}/SB_{MSY} and F_{recent}/F_{MSY} panel (b) Majuro: $SB_{recent}/SB_{F=0}$ and F_{recent}/F_{MSY} with the LRP of $SB_{recent}/SB_{F=0} = 0.20$. Points are the 80 model estimates; filled and open symbols distinguish PDH and Near-PDH fits. Nested 50%, 80% and 95% HDRs combine structural uncertainty with available joint Hessian uncertainty. Background colours follow the Kobe four-category and Majuro three-category definitions. Labels give the probability in each category from the same equal-model-weight distribution used for the combined HDRs. Open the [interactive ensemble viewer](#).

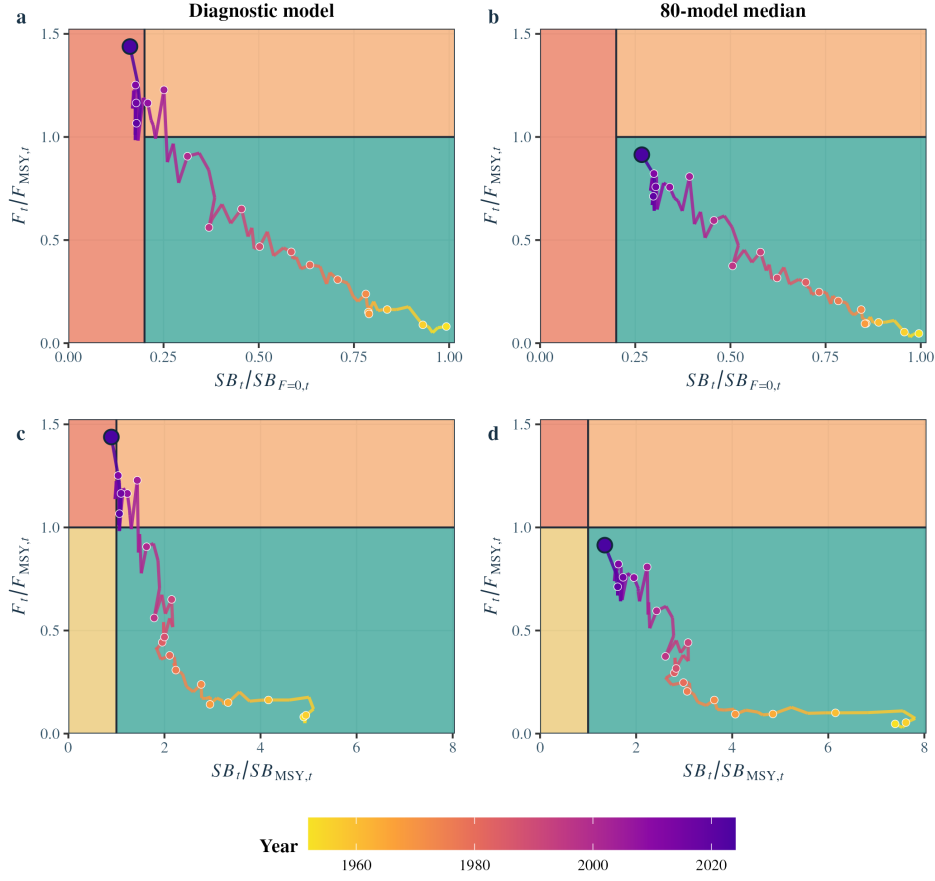


Figure 64: Time-dynamic Majuro and Kobe trajectories for the 2026 diagnostic model and the median of the 80-model ensemble. Panels (a) and (b) are Majuro trajectories for the diagnostic model and the annual ensemble median, respectively; panels (c) and (d) are the corresponding Kobe trajectories. The axes are $SB_t/SB_{F=0,t}$ and $F_t/F_{MSY,t}$ for Majuro, and $SB_t/SB_{MSY,t}$ and $F_t/F_{MSY,t}$ for Kobe. Colour shows year and the outlined point marks 2024. The ensemble trajectory is the median of the 80 retained models and therefore represents structural medians rather than Hessian estimation uncertainty. Open the [interactive ensemble viewer](#).

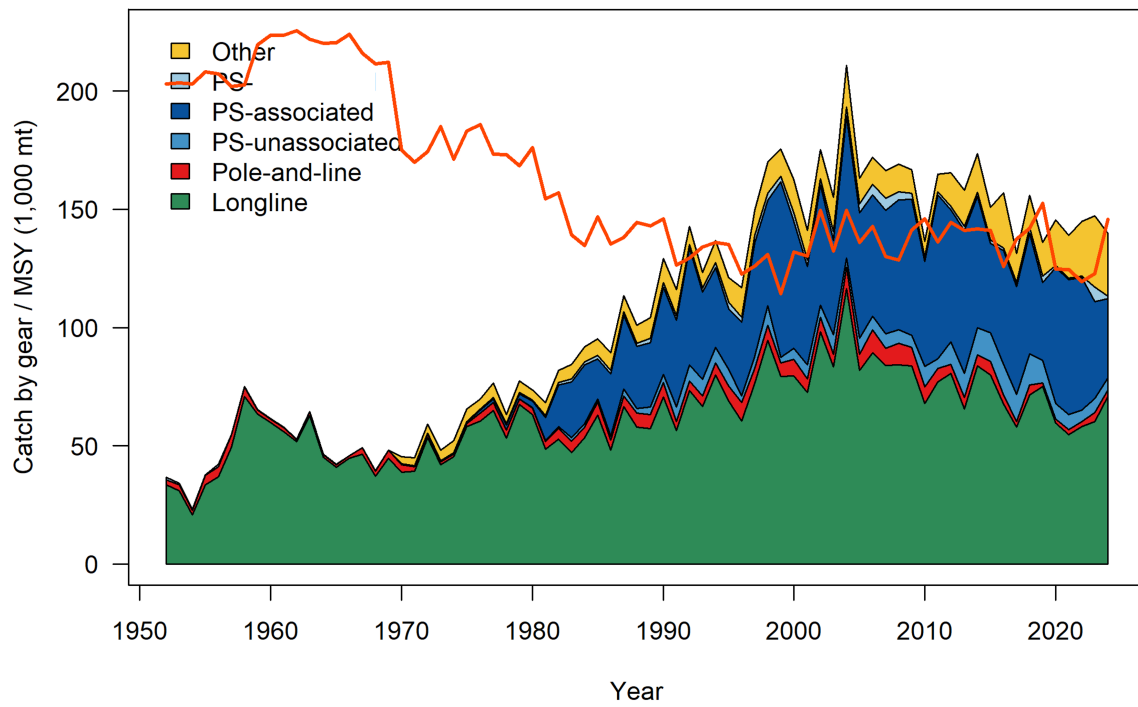


Figure 65: History of the annual estimates of MSY (red line) for the diagnostic model compared with annual catch by the main gear types. Note that this is a ‘dynamic’ MSY.



Figure 66: Estimates of fishery impact (%) by region and over all regions (bottom right), attributed to fishery groups for the diagnostic model. Total impact (%) is $100(1 - SB_t/SB_{t,F=0})$.

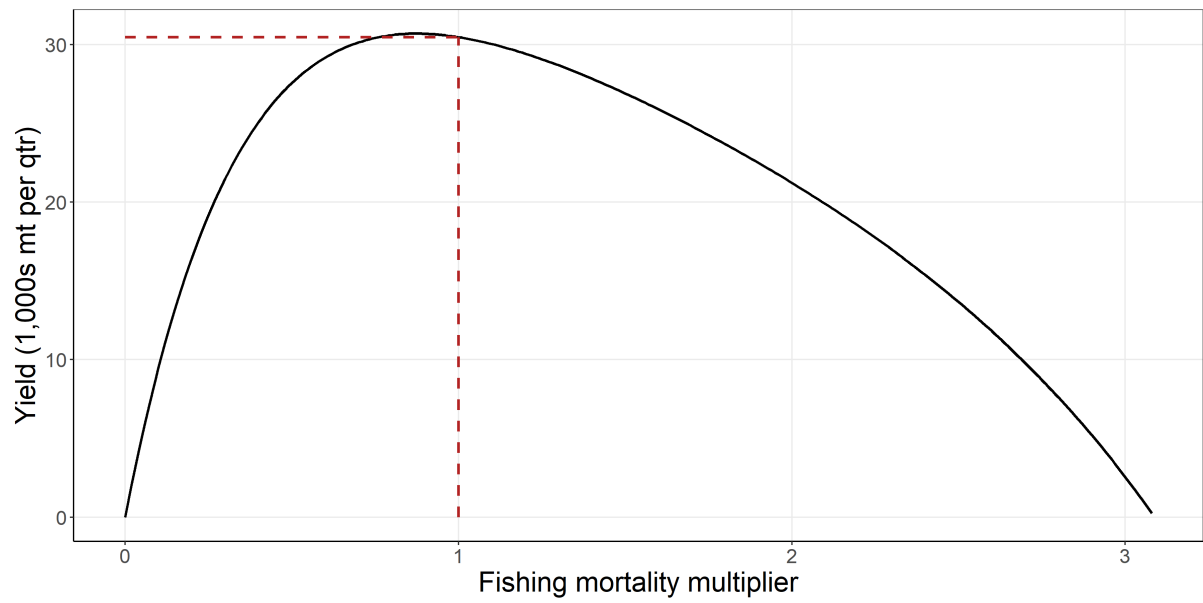


Figure 67: Estimated yield as a function of fishing mortality multiplier for the diagnostic model and the alternative mixing scenario model. The black dashed line indicates the equilibrium yield at current fishing mortality.

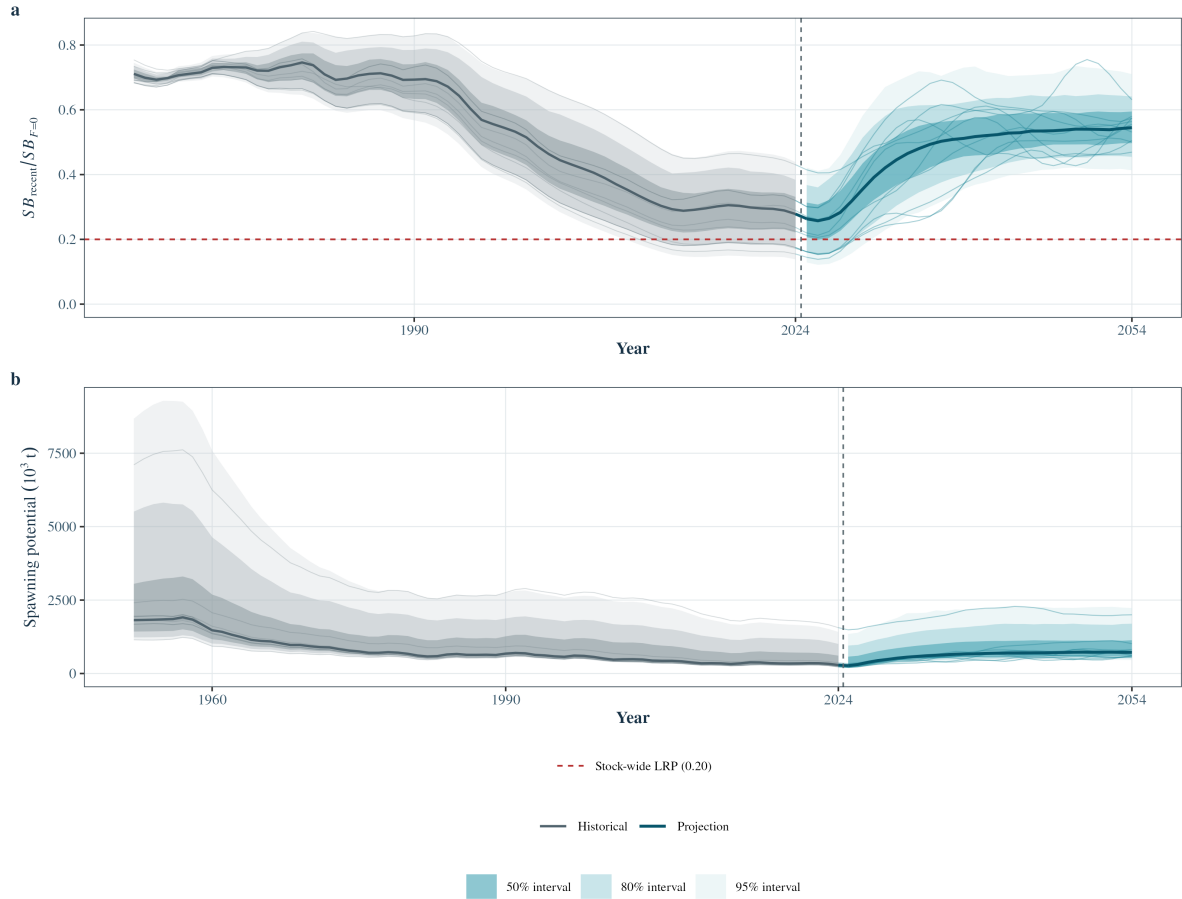


Figure 68: All-region LRP depletion metric (a) and spawning potential in thousand metric tonnes (b), joining estimates through 2024 to projections for 2025–2054. The LRP metric is the four-year mean spawning biomass divided by the preceding ten-year mean no-fishing biomass. Medians and central 50%, 80% and 95% intervals use all 800 model–recruitment realisations; ten thin lines show individual projection realisations. Historical intervals show structural uncertainty; projections add recruitment variability without Hessian parameter draws. The 0.20 line in panel a is the stock-wide LRP. Open the [interactive ensemble viewer](#).

A Appendices

A.1 No-tagging and alternative-movement model comparisons

The current bigeye tuna diagnostic model includes tagging data categorised into 98 tag-release groups modelled using a negative-binomial likelihood with assumed overdispersion of $\tau = 2$ (variance of tag recaptures equals twice the expected value). Tagging data inform the model about movement among the five model regions, fishing mortality and population size. However, several critical assumptions are required, including: (1) tagged fish are fully mixed with the untagged population in the release region following a number of pre-specified initial post-release periods designated as “mixing periods”; and (2) recaptured tagged fish are reported, and hence available for inclusion in the model, with estimated probabilities conditioned on reporting rates estimated for some fisheries from tag-seeding experiments (Peatman et al., 2026a). Other treatments correct the tagging data for variable initial survival of tagged fish and movement outside the assessment area (the Pacific Ocean west of 150°W) (Peatman et al., 2026c).

To determine the impact of tagging data on the assessment results, three alternative models were developed and compared with the diagnostic model:

- **5R-MFCL:** the diagnostic model with tagging data removed and the movement parameters estimated in the MULTIFAN-CL diagnostic model retained as fixed parameters. This removes the direct influence of tagging data on fishing mortality and population size while retaining their estimated movement information.
- **5R-SEAP:** a model similar to 5R-MFCL, but with movement parameters supplied by the current reference version of the bigeye tuna SEAPODYM model (Senina et al., 2026).
- **1R:** a single-region “fleets-as-areas” bigeye tuna model without tagging data (Hampton et al., 2026b).

The fixed movement estimates used by 5R-MFCL and 5R-SEAP are shown in Figure A1 and Figure A2. The patterns are qualitatively similar, although seasonal variation is evident in some MFCL-estimated movements but is limited in the SEAPODYM estimates. The SEAPODYM estimates vary with age, whereas the MFCL estimates are age-invariant. For example, SEAPODYM estimates that movement from regions 4 and 3 to region 1 increases with age.

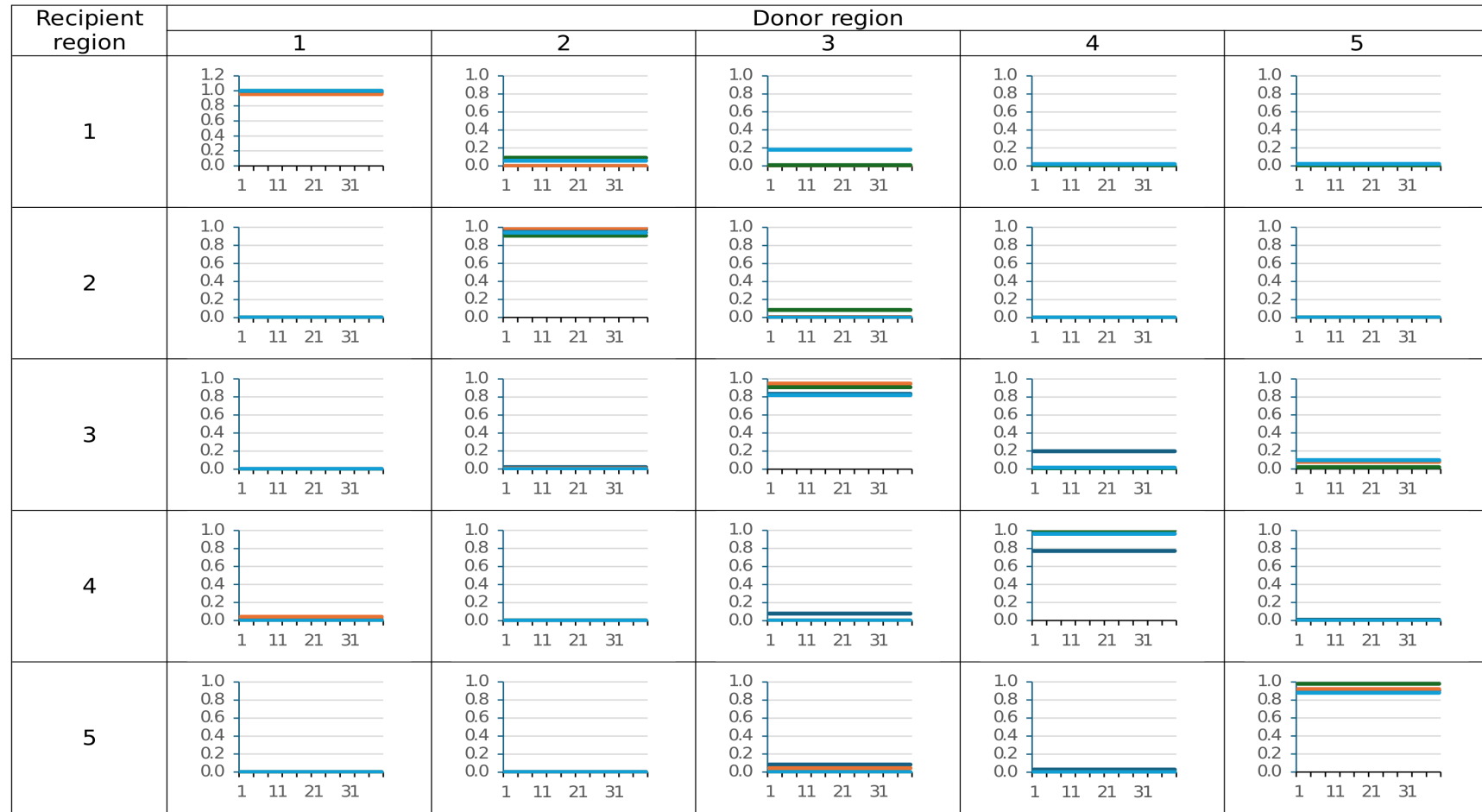


Figure A1: Bigeye tuna movement probabilities estimated in the MFCL diagnostic model. Columns identify donor regions and rows identify recipient regions; the horizontal axes are quarterly age classes. Q1 is dark blue, Q2 orange, Q3 green and Q4 light blue.

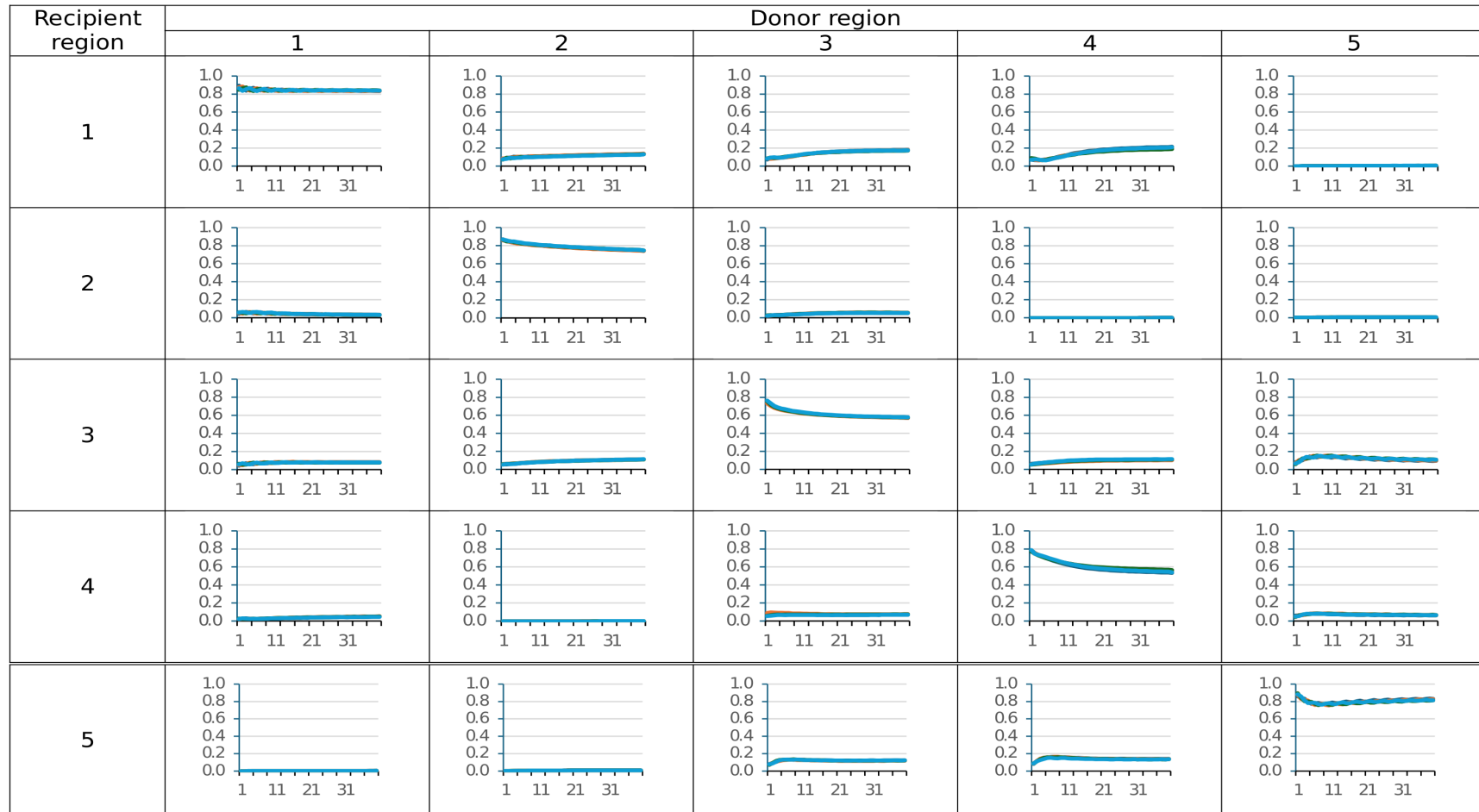


Figure A2: Bigeye tuna movement probabilities from the SEAPODYM reference model. Columns identify donor regions and rows identify recipient regions; the horizontal axes are quarterly age classes. Q1 is dark blue, Q2 orange, Q3 green and Q4 light blue.

Estimates of recruitment, spawning biomass, spawning-biomass depletion and regional spawning-biomass distribution are shown for the diagnostic model (5R-Diag) and the three alternatives in Figure A3. Regional-distribution estimates are not available from the single-region model, so the corresponding panel includes the native SEAPODYM distribution for comparison.

Overall, the headline estimates were reasonably consistent across models:

- **Recruitment (1952–2024):** the 5R-Diag average was 17.3 million fish per quarter. The 5R-MFCL average was 14% higher, 5R-SEAP was almost identical and 1R was 8% lower.
- **Spawning biomass (1952–2024):** the 5R-Diag average was 597 thousand t. The 5R-MFCL average was 9% higher, 5R-SEAP was 22% lower and 1R was 24% higher.
- **Spawning-biomass depletion (2020–2024):** the 5R-Diag average was 0.18. The 5R-MFCL average was 9% higher, 5R-SEAP was 15% lower and 1R was 5% higher.
- **Regional spawning-biomass distribution (1952–2024):** 5R-MFCL was most similar to 5R-Diag. The 5R-SEAP distribution was very similar to native SEAPODYM and placed considerably more spawning biomass in region 1 than did 5R-Diag and 5R-MFCL.

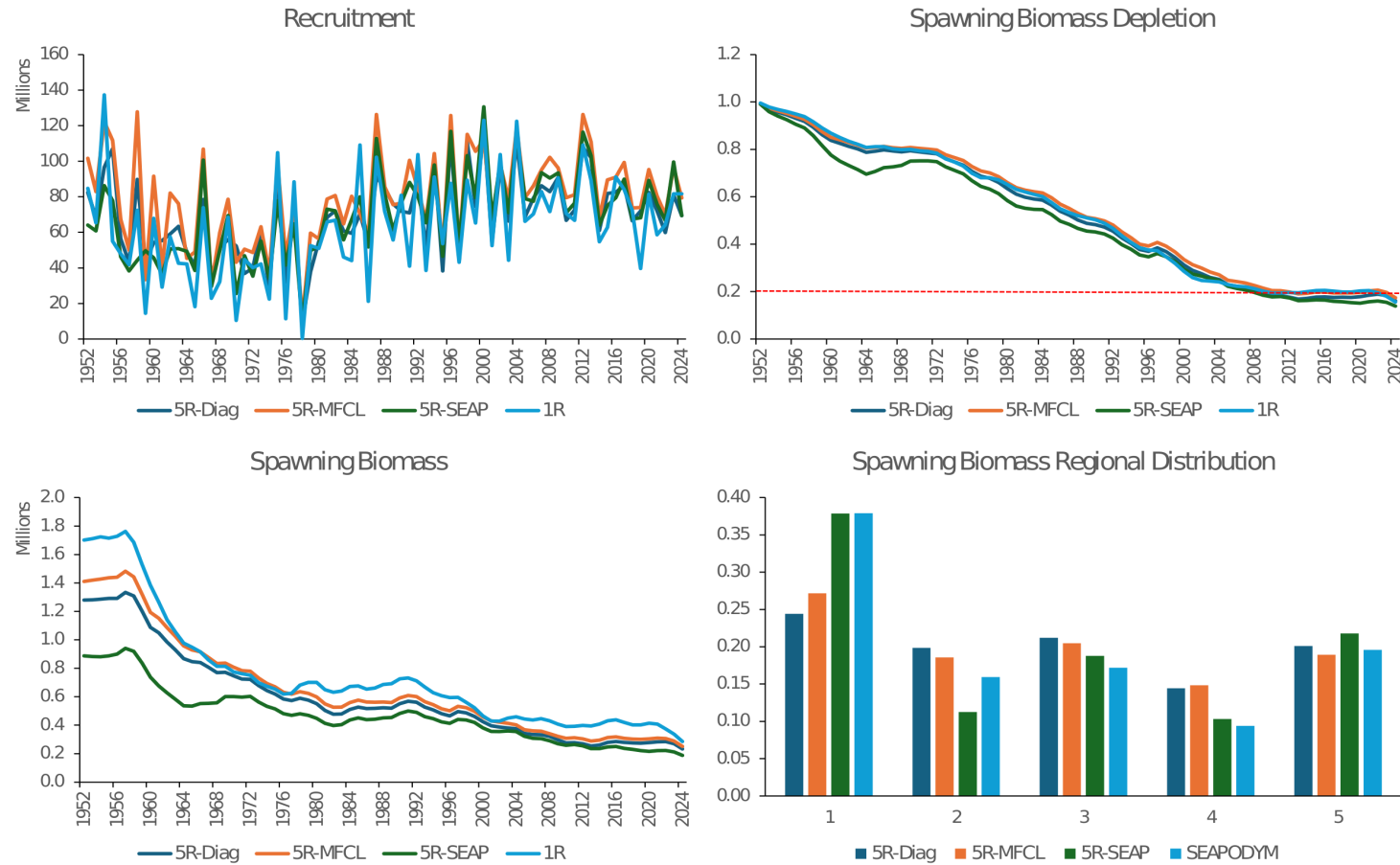


Figure A3: Comparisons of recruitment (upper left), spawning-biomass depletion (upper right), spawning biomass (lower left) and regional spawning-biomass distribution (lower right). The models are the current five-region diagnostic model (5R-Diag), the five-region model without tagging data and with fixed MFCL movement (5R-MFCL), the five-region model without tagging data and with fixed SEAPODYM movement (5R-SEAP), and the single-region model without tagging data (1R). Native SEAPODYM is included only in the regional-distribution panel.

A.1.1 Conclusions

The key estimates from 5R-Diag and the no-tagging model with the same fixed movement parameters (5R-MFCL) are very similar. This indicates either that tagging data have only a small direct effect on population scaling in the diagnostic model, or that their effect is expressed primarily through movement. Because the tagging-data likelihood profile favours greater population scaling, the second explanation is credible. A likelihood profile of the diagnostic model with movement fixed may help distinguish these explanations.

SEAPODYM estimates somewhat greater movement and mixing than the diagnostic model. In MFCL, tagging observations are aggregated into large regions, making assumptions about within-region mixing sensitive to the spatial distributions of tag releases and fishing effort. SEAPODYM operates at a finer $2^\circ \times 2^\circ$ spatial resolution and analyses tagging data in recapture-conditioned mode, making it less dependent on reporting-rate assumptions. These features provide a case that SEAPODYM movement may be less affected by large-region aggregation and more realistic because environmental drivers are represented explicitly.

Despite the different bases for estimating movement in MFCL and SEAPODYM, and some differences in the movement estimates themselves, the consistency of the headline assessment quantities is reassuring.

References

- Aires-da-Silva, A. M., Maunder, M. N., Schaefer, K. M., & Fuller, D. W. (2015). Improved growth estimates from integrated analysis of direct aging and tag-recapture data: An illustration with bigeye tuna (*thunnus obesus*) of the eastern pacific ocean with implications for management. *Fisheries Research*, 163, 119–126.
- Anderson, S. C., Ward, E. J., English, P. A., & Barnett, L. A. K. (2022). *sdmTMB: An r package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields* [Preprint]. Ecology. <https://doi.org/10.1101/2022.03.24.485545>
- Berger, A. M., McKechnie, S., Abascal, F., Kumasi, B., Usu, T., & Nichol, S. J. (2014). *Analysis of tagging data for the 2014 tropical tuna assessments: Data quality rules, tagger effects, and reporting rates* (No. WCPFC-SC10-2014/SA-IP-06). WCPFC.
- Cadigan, N. G., & 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. <https://doi.org/10.1016/j.icesjms.2004.11.015>
- Carvalho, F., Punt, A. E., Chang, Y.-J., Maunder, M. N., & Piner, K. R. (2017). Can diagnostic tests help identify model misspecification in integrated stock assessments? *Fisheries Research*, 192, 28–40. <https://doi.org/10.1016/j.fishres.2016.09.018>
- Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K. R., Maunder, M. N., Taylor, I., Wetzel, C. R., Doering, K., Johnson, K. F., & Methot, R. D. (2021). A cookbook for using model diagnostics in integrated stock assessments. *Fisheries Research*, 240, 105959. <https://doi.org/10.1016/j.fishres.2021.105959>

- Castillo Jordan, C., Teears, T., Hampton, J., Davies, N., Scutt Phillips, J., McKechnie, S., Peatman, T., Macdonald, J., Day, J., Magnusson, A., Scott, R., Scott, F., Pilling, G., & Hamer, P. (2022). *Stock assessment of skipjack tuna in the western and central pacific ocean: 2022* (No. WCPFC-SC18-SA-WP-01). WCPFC.
- Chow, S., Okamoto, H., Miyabe, N., Hiramatsu, K., & Barut, N. (2000). Genetic divergence between atlantic and indo-pacific stocks of bigeye tuna (*thunnus obesus*) and admixture around south africa. *Molecular Ecology*, 9(2), 221–227. <https://doi.org/10.1046/j.1365-294x.2000.00851.x>
- Davies, N., Bouye, F., Kim, K., Magnusson, A., & Hampton, J. (2026a). *Developments in the MULTIFAN-CL software 2025-26* (No. WCPFC-SC22-2026/SA-IP-02). WCPFC.
- Davies, N., Fournier, D. A., Hampton, J., Kleiber, P., Bouye, F., Kim, K., Magnusson, A., & Hoyle, S. (2026b). *MULTIFAN-CL user's guide version 2.2.7.8*. WCPFC.
- Davies, N., Fournier, D., & Hampton, J. (2019). *Developments in the MULTIFAN-CL software 2018-2019* (No. WCPFC-SC15-2019/SA-IP-02). WCPFC.
- Davies, N., Hoyle, S., Harley, S., Langley, A., Kleiber, P., & Hampton, J. (2011). *Stock assessment of bigeye tuna in the western and central pacific ocean* (No. WCPFC-SC7-2011/SA-WP-02). WCPFC.
- Day, J., Magnusson, A., Teears, T., Hampton, J., Davies, N., Castillo Jordán, C., Peatman, T., Scott, R., Scutt Phillips, J., McKechnie, S., Scott, F., Yao, N., Natadra, R., Pilling, G., Williams, P., & Hamer, P. (2023). *Stock assessment of bigeye tuna in the western and central pacific ocean: 2023* (No. WCPFC-SC19-2023/SA-WP-05 (Rev. 2)). WCPFC.
- Deroba, J. J., Butterworth, D. S., Methot, R. D., De Oliveira, J. A. A., Fernandez, C., Nielsen, A., Cadrin, S. X., Dickey-Collas, M., Legault, C. M., Ianelli, J., Valero, J. L., Needle, C. L., O'Malley, J. M., Chang, Y.-J., Thompson, G. G., Canales, C., Swain, D. P., Miller, D. C. M., Hintzen, N. T., ... Hulson, P.-J. F. (2015). Simulation testing the robustness of stock assessment models to error: Some results from the ICES strategic initiative on stock assessment methods. *ICES Journal of Marine Science*, 72(1), 19–30. <https://doi.org/10.1093/icesjms/fst237>
- Ducharme-Barth, N., Peatman, T., Scutt Phillips, J., & Minte-Vera, C. (2026). *Natural mortality estimation for WCPO bigeye tuna: A joint cohort analysis of coral sea and region 4 tagging data* (No. WCPFC-SC22-2026/SA-IP-14). WCPFC.
- Ducharme-Barth, N., Vincent, M., Hampton, J., Hamer, P., Williams, P., & Pilling, G. (2020a). *Stock assessment of bigeye tuna in the western and central pacific ocean* (No. WCPFC-SC16-2020/SA-WP-03 [REV3]). WCPFC.
- Ducharme-Barth, N., Vincent, M., Vidal, T., & Hamer, P. (2020b). *Analysis of pacific-wide operational longline dataset for bigeye and yellowfin tuna catch-per-unit-effort (CPUE)* (No. SC16-SA-IP-07). WCPFC.
- Evans, K., Langley, A., Clear, N. P., Williams, P., Patterson, T., Sibert, J., Hampton, J., & Gunn, J. S. (2008). Behaviour and habitat preferences of bigeye tuna (*thunnus obesus*) and their influence on longline fishery catches in the western coral sea. *Canadian Journal of Fisheries and Aquatic Sciences*, 65, 2427–2443.
- Eveson, P., Vincent, M., Farley, J., Krusic-Golub, K., & Hampton, J. (2020). *Integrated growth*

- models from otolith and tagging data for yellowfin and bigeye tuna in the western and central pacific ocean* (No. SC16-SA-IP-03). WCPFC.
- Farley, J., Eveson, P., Krusic-Golub, K., Sanchez, C., Rouspard, F., McKechnie, S., Nichol, S., Leroy, B., Smith, N., & Chang, S.-K. (2017). *Age, growth and maturity of bigeye tuna in the pacific* (No. WCPFC-SC13- 2017/SA-WP-01). WCPFC.
- Farley, J., Krusic-Golub, K., Eveson, P., Clear, N., Rouspard, F., Sanchez, C., Nicol, S., & Hampton, J. (2020). *Age and growth of yellowfin and bigeye tuna in the western and central pacific ocean from otoliths* (No. SC16-SA-WP-02). WCPFC.
- Fournier, D. A., Hampton, J., & 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, 2105–2116.
- Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielson, A., & 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.
- Fournier, D., & Archibald, C. P. (1982). A general-theory for analyzing catch at age data. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(8), 1195–1207. <https://doi.org/10.1139/f82-157>
- Francis, R. I. C. C. (1992). Use of risk analysis to assess fishery management strategies: A case study using orange roughy (*Hoplostethus atlanticus*) on the chatham rise, new zealand. *Canadian Journal of Fisheries and Aquatic Science*, 49, 922–930.
- Francis, R. I. C. C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, 68, 1124–1138.
- Gonzalez, E. G., Beerli, P., & Zardoya, R. (2008). Genetic structuring and migration patterns of atlantic bigeye tuna, *thunnus obesus* (lowe, 1839). *BMC Evolutionary Biology*, 8(1), 252. <https://doi.org/10.1186/1471-2148-8-252>
- Grewe, P. M., & Hampton, J. (1998). *An assessment of bigeye (thunnus obesus) population structure in the pacific ocean based on mitochondrial DNA and DNA microsatellite analysis*. JIMAR Contribution 98-330.
- Grewe, P. M., Wudianto, C. H., Proctor, C. H., Adam, M. S., Jauhary, A. R., Schaefer, K., Itano, D. G., Evans, K., Killian, A., Foster, S. D., Gosselin, T., Feutry, P., Aulich, J., Gunasekera, R. M., Lansdell, M., & Davies, C. R. (2019). *Population structure and connectivity of tropical tuna species across the indo pacific ocean region* (No. WCPFC-SC15-2019/SA-IP-15). WCPFC.
- Hamel, O. S. (2015). A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. *Ices Journal of Marine Science*, 72(1), 62–69. <https://doi.org/10.1093/icesjms/fsu131>
- Hamel, O. S., & Cope, J. M. (2022). Development and considerations for application of a longevity-based prior for the natural mortality rate. *Fisheries Research*, 256, 106477. <https://doi.org/10.1016/j.fishres.2022.106477>
- Hamer, P. (2026). *Summary report from the SPC pre-assessment workshop - march 2026* (No. WCPFC-SC22-2026/SA-IP-01). WCPFC.

- Hamer, P., Castillo Jordán, C., Panizza, A., & Peatman, T. (2026). *WCPFC Project127/127a - review and reconciliation of size data collected in the WCPFC-CA for stock assessment purposes: Phase 2 progress report* (No. WCPFC-SC22-2026/ST-WP-02). WCPFC.
- Hamer, P., Potts, J., Macdonald, J., & Senina, I. (2023). *Review and analyses to inform conceptual models of population structure and spatial stratification of bigeye and yellowfin tuna assessments in the western and central pacific ocean* (No. SC19-SA-WP-02). WCPFC.
- Hamer, P., Schneider, E., Vidal, T., & Williams, P. (2025). *Progress report: Review and reconciliation of size data collected in the WCPFC-CA for stock assessment purposes (WCPFC project: 127)* (No. WCPFC-SC21-ST-WP-02). WCPFC.
- Hampton, J. (2000). Natural mortality rates in tropical tunas: Size really does matter. *Canadian Journal of Fisheries and Aquatic Science*, 57, 1002–1010.
- Hampton, J., Davies, N., & Hamer, P. (2026a). *Investigation of the impacts of regional growth variation on the 2023 bigeye tuna stock assessment* (No. WCPFC-SC22-2026/SA-IP11). WCPFC.
- Hampton, J., & 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, 937–963.
- Hampton, J., Tears, T., Kim, K., Magnusson, A., Davies, N., & Hamer, P. (2026b). *Investigation of single region fleets as areas models for bigeye and yellowfin tuna in the WCPO* (No. WCPFC-SC22-2026/SA-IP-09). WCPFC.
- Hampton, J., & Williams, P. (2016). *Annual estimates of purse seine catches by species based on alternative data sources* (No. WCPFC-SC12-2016/ST-IP-03). WCPFC.
- Hare, S., Vidal, T., Castillo Jordán, C., Day, J., Hamer, P., Hill, N., Machful, P., Magnusson, A., Nicol, S., Odongo, V., Panizza, A., Peatman, T., Scott, R., Senina, I., Sloyan, B., Teeares, T., & Pilling, G. (2025). *The western and central pacific tuna fishery: 2024 overview and status of stocks* [Tuna Fisheries Assessment Report No. 25]. SPC.
- Hare, S., Vidal, T., Scott, R., & Pilling, G. (2026). *A compendium of fisheries indicators for target tuna stocks in the WCPFC convention area* (No. WCPFC-SC22-2026/SA-WP-07). WCPFC.
- Harley, S. J. (2011). *A preliminary investigation of steepness in tunas based on stock assessment results* (No. WCPFC-SC7-2011/SA-IP-08). WCPFC.
- Harley, S. J., Davies, N., Hampton, J., & McKechnie, S. (2014). *Stock assessment of bigeye tuna in the western and central pacific ocean* (No. WCPFC-SC10-2014/SA-WP-01). WCPFC. <http://wcpfc.int/system/files/SC10-SA-WP-01%20BET%20Assessment.pdf>
- Harley, S. J., Hoyle, S., Hampton, J., & Kleiber, P. (2010). *Stock assessment of bigeye tuna in the western and central pacific ocean* (No. WCPFC-SC6-2010/SA-WP-01). WCPFC.
- Harley, S. J., & Maunder, M. N. (2003). *A simple model for age-structured natural mortality based on changes in sex ratios* (No. SAR-4-01). Inter-American Tropical Tuna Commission.
- Harley, S., Hoyle, S., & Langley, A. (2009). *Stock assessment of bigeye tuna in the western and central pacific ocean*. (No. WCPFC-SC5-2009/SA-WP-04). WCPFC.
- Hasegawa, T., Okamoto, K., Ijima, H., Kim, K., Hamer, P., Wakefield, C., Pons, A., Widdrington, J., & Davies, N. (2026). *New age-at-length data for WCPO bigeye and yellowfin tuna from the*

- temperate north pacific and approaches to address spatial sampling bias in growth estimation within MFCL* (No. WCPFC-SC22-2026/IP-08). WCPFC.
- Hoyle, S. (2024). *Effort creep in longline and purse seine CPUE and its application in tropical tuna assessments* (No. WCPFC-SC20-2024/SA-IP-09). WCPFC.
- Hoyle, S. D. (2008). *Adjusted biological parameters and spawning biomass calculations for south pacific albacore tuna, and their implications for stock assessments* (No. WCPFC-SC4-2008/ME-WP-02). WCPFC.
- Hoyle, S., & Nicol, S. (2008). *Sensitivity of bigeye stock assessment to alternative biological and reproductive assumptions* (No. WCPFC-SC4-2008/ME-WP-01). WCPFC.
- Hurtado-Ferro, F., Szuwalski, C. S., Valero, J. L., Anderson, S. C., Cunningham, C. J., Johnson, K. F., Licandeo, R., McGilliard, C. R., Monnahan, C. C., Muradian, M. L., Ono, K., Vert-Pre, K. A., Whitten, A. R., & Punt, A. E. (2015). Looking in the rear-view mirror: Bias and retrospective patterns in integrated, age-structured stock assessment models. *Ices Journal of Marine Science*, 72(1), 99–110. <https://doi.org/10.1093/icesjms/fsu198>
- Ianelli, J., Maunder, M. N., & Punt, Andre. E. (2012). *Independent review of the 2011 WCPO bigeye tuna assessment* (No. WCPFC-SC8-2012/SA-WP-01). WCPFC.
- Ijima, H., & Jusup, M. (2023). *Tuna and billfish larval distributions in a warming ocean*. arXiv. <http://arxiv.org/abs/2304.09442>
- ISSF. (2011). *Report of the 2011 ISSF stock assessment workshop* (No. ISSF Technical Report 2011-02). ISSF.
- Itano, D. (2000). *The reproductive biology of yellowfin tuna (Thunnus albacares) in hawaiian waters and the western tropical pacific ocean: Project summary* (JIMAR Contribution 00-328 No. SOEST 00-01). University of Hawai'i.
- Kim, K., & Neubauer, P. (2025). A quantitative evaluation of the pollock *Gadus chalcogrammus* restocking programme in south korea. *Canadian Journal of Fisheries and Aquatic Sciences*, 82, 1–20. <https://doi.org/10.1139/cjfas-2024-0375>
- Kim, K., Sibanda, N., Arnold, R., & A'mar, T. (2024). Enhancing data-limited assessments with random effects: A case study on korea chub mackerel (*Scomber japonicus*). *Canadian Journal of Fisheries and Aquatic Sciences*, 81(10), 1433–1455. <https://doi.org/10.1139/cjfas-2023-0358>
- Kolody, D., & Hoyle, S. (2015). Evaluation of tag mixing assumptions in western pacific ocean skipjack tuna stock assessment models. *Fisheries Research*, 163, 127–140. <https://doi.org/10.1016/j.fishres.2014.05.008>
- Langley, A., Hampton, J., Kleiber, P., & Hoyle, S. (2008). *Stock assessment of bigeye tuna in the western and central pacific ocean, including an analysis of management options* (No. WCPFC-SC4/SA-WP-1). WCPFC.
- Maunder, M. N., & Piner, K. R. (2015). Contemporary fisheries stock assessment: Many issues still remain. *ICES Journal of Marine Science*, 72(1), 7–18. <https://doi.org/10.1093/icesjms/fsu015>
- McKechnie, S., Hampton, J., Abascal, F., Davies, N., & Harley, S. J. (2015a). *Sensitivity of WCPO stock assessment results to the inclusion of EPO dynamics within a pacific-wide analysis* (No. WCPFC-SC11-2015/SA-WP-04). WCPFC.
- McKechnie, S., Ochi, D., Kiyofuji, H., Peatman, T., & Caillot, S. (2016). *Construction of tagging*

- data input files for the 2016 skipjack tuna stock assessment in the western and central pacific ocean* (No. WCPFC-SC12-2016/SA-IP-05). WCPFC.
- McKechnie, S., Pilling, G., & Hampton, J. (2017a). *Stock assessment of bigeye tuna in the western and central pacific ocean* (No. WCPFC-SC13-2017/SA-WP-05). WCPFC.
- McKechnie, S., Pilling, G., & Vidal, T. (2026). *Summaries of bigeye tuna catch and effort to support management procedure discussions* (No. WCPFC-SC22-2026/SA-IP-12). WCPFC.
- McKechnie, S., Tremblay-Boyer, L., & Harley, S. J. (2015b). *Analysis of pacific-wide operational longline CPUE data for bigeye tuna* (No. WCPFC-SC11-2015/SA-WP-03). WCPFC.
- McKechnie, S., Tremblay-Boyer, L., & Pilling, G. (2017b). *Background analyses for the 2017 stock assessments of bigeye and yellowfin tuna in the western and central pacific ocean* (No. WCPFC-SC13-2017/SA-IP-06). WCPFC.
- Minte-Vera, C. V., Maunder, M. N., & Aires-da-Silva, A. (2021). Auxiliary diagnostic analyses used to detect model misspecification and highlight potential solutions in stock assessments: Application to yellowfin tuna in the eastern Pacific Ocean. *ICES Journal of Marine Science*, 78(10), 3521–3537. <https://doi.org/10.1093/icesjms/fsab213>
- Mohn, R. (1999). The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. *Ices Journal of Marine Science*, 56(4), 473–488. <https://doi.org/10.1006/jmsc.1999.0481>
- Moore, B. R., Bell, J. D., Evans, K., Farley, J., Grewe, P. M., Hampton, J., Marie, A. D., Minte-Vera, C., Nicol, S., Pilling, G. M., Scutt Phillips, J., Tremblay-Boyer, L., Williams, A. J., & Smith, N. (2020). Defining the stock structures of key commercial tunas in the pacific ocean i: Current knowledge and main uncertainties. *Fisheries Research*, 230, 105525. <https://doi.org/10.1016/j.fishres.2020.105525>
- Moore, B. R., Lestari, P., Cutmore, S. C., Proctor, C., & Lester, R. J. G. (2019). Movement of juvenile tuna deduced from parasite data. *ICES Journal of Marine Science*, 76(6), 1678–1689. <https://doi.org/10.1093/icesjms/fsz022>
- Natasha, J., Stockwell, B. L., Marie, A. D., Hampton, J., Smith, N., Nicol, S., & Rico, C. (2022). No population structure of bigeye tunas (*thunnus obesus*) in the western and central pacific ocean indicated by single nucleotide polymorphisms. *Front. Mar. Sci.*, 9, 799684. <https://doi.org/10.3389/fmars.2022.799684>
- Nicol, S., Hoyle, S., Farley, J., Muller, B., Retalmai, S., Sisior, K., & Williams, A. (2011). *Bigeye tuna age, growth, and reproductive biology (project 35)* (No. WCPFC-SC7-2011/SA-WP-01). WCPFC.
- Nishikawa, Y., Honma, M., Ueyanagi, S., & Kikawa, S. (1985). Average distribution of larvae of oceanic species of scombroid fishes, 1956-1981. *Far Seas Fish.Res.Lab.*, 99 p.
- Peatman, T., Fukofuka, S., Park, T., Williams, P., Hampton, J., & Smith, N. (2019). *Better purse seine catch composition estimates: Progress on the project 60 work plan*. [Technical Report]. SPC.
- Peatman, T., Gislard, S., Scutt Phillips, J., Potts, J., & Nicol, S. (2026a). *Analysis of tag seeding data for the 2026 yellowfin & bigeye assessments: Reporting rates for purse seine fleets* (No. WCPFC-SC22-2026/SA-IP-05). WCPFC.
- Peatman, T., Kim, K., Magnusson, A., Castillo Jordan, C., & Hamer, P. (2026b). *Analysis of size*

- frequency data for the 2026 bigeye and yellowfin assessments* (No. WCPFC-SC22-SA-IP06). WCPFC.
- Peatman, T., Scutt Phillips, J., Potts, J., & Nicol, S. (2026c). *Analysis of tagging data for the 2026 yellowfin & bigeye assessments: Corrections to tag releases for tagging conditions* (No. WCPFC-SC22-2026/SA-IP-04). WCPFC.
- Peatman, T., Williams, P., & Nicol, S. (2021). *Project 60: Progress towards achieving SC16 recommendations* (No. WCPFC-SC17-2021/ST-IP-04). WCPFC.
- Peatman, T., Williams, P., & Nicol, S. (2023). *Project 60: Progress report - improving purse seine species composition* (WCPFC Scientific Committee Paper No. SC19-ST-IP-03).
- Pilling, G., Scott, R., Davies, N., & Hampton, J. (2016). *Approaches used to undertake management projections of WCPO tuna stocks based upon MULTIFAN-CL stock assessments* (No. WCPFC-SC12-2016/MI-IP-04). Bali, Indonesia, 3-11 August 2016.
- Piner, K., & Lee, H. (2011). *Meta-analysis of striped marlin natural mortality* (No. ISC/11/BILLWG-1/10). ISC.
- Proctor, C. H., Lester, R. G. J., Clear, N. P., Grewe, P. M., Moore, B. R., Eveson, J. P., Lestari, P., Wudji, A., Taufik, M., Wudianto, Lansdell, M. J., Hill, P. L., Dietz, C., Thompson, J. M., Cutmore, S. C., Foster, S. D., Gosselin, T., & Davies, C. R. (2019). *Population structure of yellowfin tuna (Thunnus albacares) and bigeye tuna (T. obesus) in the indonesian region*. (No. Final Report as output of ACIAR Project FIS/2009/059; pp. 139 pp.). ACIAR.
- Punt, A. E., Maunder, M. N., & Ianelli, J. (2023). *Independent review of recent WCPO yellowfin tuna assessment* (No. SC19-SA-WP-01). WCPFC.
- Rooker, J. R., David Wells, R. J., Itano, D. G., Thorrold, S. R., & Lee, J. M. (2016). Natal origin and population connectivity of bigeye and yellowfin tuna in the pacific ocean. *Fish. Oceanogr.*, 25(3), 277–291. <https://doi.org/10.1111/fog.12154>
- Schaefer, K. M., Fuller, D. W., & Miyabe, N. (2005). *Reproductive biology of bigeye tuna (Thunnus obesus) in the eastern and central pacific ocean* (No. 23(1). 35p.). IATTC Bulletin.
- Schaefer, K., Fuller, D., Hampton, J., Caillot, S., Leroy, B., & Itano, D. (2015). Movements, dispersion, and mixing of bigeye tuna (*Thunnus obesus*) tagged and released in the equatorial central pacific ocean, with conventional and archival tags. *Fisheries Research*, 161, 336–355.
- Scutt Phillips, J., Lehodey, J., Hampton, J., Senina, I., & Nicol, S. (2022). *Quantifying rates of mixing in tagged, WCPO skipjack tuna* (No. WCPFC-SC18-2022/SA-WP-04). WCPFC.
- Scutt Phillips, J., Sen Gupta, A., Senina, I., Sebille, E., Lange, M., Lehodey, P., J, H., & Nichol, S. (2018). An individual-based model of skipjack tuna katsuwonus pelamis movement in the tropical pacific ocean. *Progress in Oceanography*, 164, 63–74.
- Scutt Phillips, J., Senina, I., & Bonin, L. (2026). *Estimation of tag mixing periods for the 2026 WCPO tuna stock assessments* (No. WCPFC-SC22-2026-SA-IP-10). WCPFC.
- Senina, I., Bonnini, L., Forestier, R., Lengaigne, M., Guihou, K., Temple-Boyer, J., & Nicol, S. (2026). *Updated SEAPODYM reference models for four pacific tuna species* (No. WCPFC-SC22-2026/EB-IP-04). WCPFC. <https://meetings.wcpfc.int/node/32435>
- Senina, I., Lehodey, P., Nicol, S., Scutt Phillips, J., & Hampton, J. (2020). *SEAPODYM: Revisiting bigeye reference model with conventional tagging data* (No. WCPFC-SC16-2020/EB-IP-05). WCPFC.

- Subbey, S. (2018). Parameter estimation in stock assessment modelling: Caveats with gradient-based algorithms. *ICES Journal of Marine Science*, 75(5), 1553–1559. <https://doi.org/10.1093/icesjms/fsy044>
- Sun, C.-L., Chu, S.-L., & Yeh, S.-Z. (2006). *Reproductive biology of bigeye tuna in the western and central pacific ocean* (No. WCPFC-SC2-2006/BI-WP-01). WCPFC.
- Teears, T. (2026). *Background analyses and data inputs for the 2026 WCPO yellowfin and bigeye tuna stock assessment* (No. WCPFC-SC22-2026-SA-IP03). WCPFC.
- Teears, T., Castillo Jordán, C., Davies, N., Day, J., Hampton, J., Magnusson, A., Peatman, T., Pilling, G., Xu, H., Vidal, T., Williams, P., & Hamer, P. (2024). *Stock assessment of south pacific albacore: 2024* (No. WCPFC-SC20-2024/SA-WP-02-Rev 2). WCPFC.
- Teears, T., Davies, N., Hamer, P., Magnusson, A., Peatman, T., Scutt Phillips, J., Vidal, T., & Hampton, J. (2025). *Stock assessment of skipjack tuna in the western and central pacific ocean: 2025 (rev 3)* (No. WCPFC-SC21-2025/SA-WP-02_rev_3). WCPFC. <https://meetings.wcpfc.int/node/26679>
- Teears, T., Day, J., Hampton, J., Magnusson, A., McKechnie, S., Peatman, T., Scutt Phillips, J., Williams, P., & Hamer, P. (2023). *CPUE analysis and data inputs for the 2023 bigeye and yellowfin tuna assessments in the WCPO* (No. WCPFC-SC19-2023/SA-WP-03). WCPFC.
- Thorson, J. T., Johnson, K. F., Methot, R. D., & Taylor, I. G. (2017). Model-based estimates of effective sample size in stock assessment models using the dirichlet-multinomial distribution. *Fisheries Research*, 192, 84–93.
- Thorson, J. T., Shelton, A. O., Ward, E. J., & 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. <https://doi.org/10.1093/icesjms/fsu243>
- Tremblay-Boyer, L., McKechnie, S., Pilling, G., & Hampton, J. (2017). *Stock assessment of yellowfin tuna in the western and central pacific ocean*. (No. WCPFC-SC13-2017/SA-WP-06). WCPFC.
- Vidal, T., Heather, F., Satria, F., Hai, V., Barcoma, S., Duong, N., Sadiyah, L., Anas, M., Januar, A., Rachmadi, D., & Suadela, P. (2026). *Data improvements in the west pacific east asia (WPEA) region of WCPFC* (No. WCPFC-SC22-2026/ST-IP-05). WCPFC.
- Vincent, M., Ducharme-Barth, N., Hamer, P., Hampton, J., Williams, P., & Pilling, G. (2020). *Stock assessment of yellowfin tuna in the western and central pacific ocean* [WCPFC-SC16-2020/SA-WP-04-Rev2]. WCPFC.
- Vincent, M., Pilling, G., & Hampton, J. (2018). *Incorporation of updated growth information within the 2017 WCPO bigeye stock assessment grid, and examination of the sensitivity of estimates to alternative model spatial structures* (No. WCPFC-SC14-2018/ SA-WP-03). Oceanic Fisheries Programme, The Pacific Community.
- Wakefield, C., Widdrington, J., Krusic-Golub, K., Pons, A., Reis-Santos, P., Gillanders, B., Farley, J., Nicol, S., & Hamer, P. (2026). *Project 129 update: Building an age-length data stream for tuna assessments* (No. WCPFC-SC22-2026/SA-WP-12). WCPFC.
- Williams, P., & Ruaia, T. (2023). *Overview of tuna fisheries in the Western and Central Pacific Ocean, including economic conditions - 2022*.