



**SCIENTIFIC COMMITTEE
EIGHTEENTH REGULAR SESSION**
10–18 August 2022

Stock assessment of skipjack tuna in the western and central Pacific Ocean: 2022

WCPFC-SC18-2022/SA-WP-01 (REV5)

**C. Castillo Jordán¹, T. Teears¹, J. Hampton¹, N. Davies², J. Scutt Phillips¹,
S. McKechnie¹, T. Peatman³, J. Macdonald¹, J. Day¹, A. Magnusson¹, R. Scott¹,
F. Scott¹, G. Pilling¹, P. Hamer¹**

¹Oceanic Fisheries Programme, Secretariat of the Pacific Community

²TeTakina Ltd

³Private Consultant

Revision 1:

This revision involves an update of [Figure 15.4](#) to include the 2020 retrospective model that was missing from the original retrospective runs, plus the recalculation of Mohns rho with this model included. It also includes corrections to the values of MSY for the 90th percentile, max and diagnostic model in [Table 4](#).

Revision 2:

A new figure ([Figure 15.17](#)) was added to appendix 15.4 to compare the 8 region diagnostic model with the 5 region diagnostic model.

Revision 3:

[Figure 26](#) has been updated due to two versions of the same figure being mistakenly inserted, it now has the correct figure in the bottom panel. [Figure 15.7](#) has also been updated due to a similar issue, noting the updated figure is very similar to the previous figure.

Revision 4:

This revision has additional appendices related to model convergence status, including a section on jitter analyses ([Section 15.2](#)), and discussion of the Hessian diagnostic ([Section 15.3](#)). There is also an additional likelihood profile plot in [Section 15.1](#). Some additional text was added to [Section 10.3.1](#) to account for the addition of the jitter analyses. There is a correction to the y-axis labelling for [Figure 46](#), and a labelling error in [Figure 45](#) where the labels PTTP Z(PH)-5 and PTTP S-5 were assigned to the wrong graphs. Some text was added to [Section 7.4.1](#) to account for the use of relatively uninformative priors for the SSAP reporting groups in this assessment, compared to the 2019 assessment that used more informative priors. [Figure 33](#), [Figure 34](#) and [Figure 35](#) have been updated as the previous versions had excluded tag recaptures that had occurred within the mixing periods. We have also updated [Figure 22](#) to provide separate panels for the pole-and-line and purse seine CPUE indices and added plots of residuals to aid with interpretation.

R Shiny app for exploring the diagnostics and outputs from the 2022 WCPO skipjack stock assessment is available at: <https://ofp-sam.shinyapps.io/GridSKJ2022/>

Revision 5:

This revisions include some minor corrections to captions for [Figure 46](#), [Figure 47](#), [Figure 48](#), [Figure 21](#), and reference to reduction of data conflict within the research recommendations.

Contents

1	Executive Summary	6
2	Author Contributions	9
3	Introduction	10
4	Background	11
4.1	Stock structure	11
4.2	Biological characteristics	12
4.3	Fisheries	13
5	Application of a new catch conditioned approach	14
6	Data compilation	15
6.1	General notes	15
6.2	Spatial stratification	15
6.3	Temporal stratification	16
6.4	Definition of fisheries	16
6.5	Catch and effort data	17
6.5.1	General characteristics	17
6.5.2	Purse seine	18
6.5.3	Longline	19
6.5.4	Pole-and-line	19
6.5.5	Other fisheries	21
6.6	Size data	21
6.6.1	Purse seine	21
6.6.2	Pole-and-line	21
6.6.3	Longline	22
6.6.4	Other fisheries	22
6.7	Re-weighting of size composition data	23
6.8	Tagging data	23
7	Model description	24
7.1	General characteristics	24
7.2	Population dynamics	25
7.2.1	Recruitment	25
7.2.2	Initial population	26
7.2.3	Growth	26
7.2.4	Movement	27
7.2.5	Natural mortality	27
7.2.6	Sexual maturity	28
7.3	Fishery dynamics	28
7.3.1	Selectivity	28
7.4	Dynamics of tagged fish	29

7.4.1	Tag reporting	29
7.4.2	Tag mixing	30
7.5	Likelihood components	32
7.5.1	Survey fishery CPUE likelihood	32
7.5.2	Length frequency: Dirichlet-multinomial likelihood	33
7.5.3	Tagging data	33
7.6	Parameter estimation and uncertainty	34
7.7	Stock assessment interpretation methods	34
7.7.1	Reference points	34
7.7.2	Yield analysis	35
7.7.3	Depletion and fishery impact	36
7.7.4	Kobe analysis and Majuro plots	36
7.7.5	Stock projections from the structural uncertainty grid	36
8	Model runs	37
8.1	Developments from the last assessment	37
8.2	Sensitivity analyses and structural uncertainty	38
8.3	Structural uncertainty	39
9	Results	40
9.1	Consequences of key model developments	40
9.2	Fit of the diagnostic model to data sources	41
9.2.1	Standardized CPUE: survey fisheries	41
9.2.2	Size composition data	42
9.2.3	Tagging data	43
9.3	Model parameter estimates (diagnostic model)	43
9.3.1	Selectivity	43
9.3.2	Movement	44
9.3.3	Natural mortality	44
9.3.4	Maturity-at-Age	45
9.3.5	Tag reporting rates	45
9.3.6	Growth	45
9.4	Stock assessment results	46
9.4.1	Recruitment: diagnostic model	46
9.4.2	Biomass: diagnostic model	46
9.4.3	Depletion: diagnostic model	47
9.4.4	Fished (SB) versus unfished ($SB_{F=0}$) spawning potential: diagnostic model	47
9.4.5	Fishing mortality: diagnostic model	47
9.5	Multi-model inference: sensitivity analyses and structural uncertainty	48
9.5.1	One-off sensitivity analyses	48
9.5.2	Structural uncertainty grid	48

9.5.3	Analyses of stock status	50
10	Discussion and conclusions	51
10.1	Stock Status	51
10.2	Changes to the previous assessment	52
10.3	Structural uncertainties	53
10.3.1	Model diagnostics	54
10.4	Recommendations for further work	55
10.4.1	WCPFC-specific recommendations	55
10.5	Main assessment conclusions	56
11	Acknowledgements	58
12	References	59
13	Tables	66
14	Figures	69
15	Appendices	133
15.1	Likelihood profiles	133
15.2	Convergence status: Jitter analyses	135
15.3	Hessian diagnostic	138
15.4	Retrospective analyses	139
15.5	‘Status quo’ deterministic stock projections for WCPO skipjack tuna	141
15.6	5 region model	145

1 Executive Summary

This paper describes the 2022 stock assessment of skipjack tuna *Katsuwonus pelamis* in the western and central Pacific Ocean. An additional three years of data were available since the previous assessment in 2019, and the model extends through to the end of 2021. The assessment applies the same 8-region model structure that was used for management advice from the 2019 assessment. New developments to the stock assessment include:

- Application of a new MFCL catch conditioned approach to the estimation of fishing mortality, plus inclusion of survey fisheries and a likelihood component for the indices from those survey fisheries.
- Application of a self-scaling approach to estimate effective sample size, the Dirichlet-multinomial likelihood, with growth estimation within the diagnostic model.
- Application of variable tag mixing periods for tag release groups based on simulations using individual based modelling of tag mixing processes.
- Development of an alternative growth model based on tag recapture growth increments and daily aging from otoliths.
- Development of new CPUE indices based on unassociated (free-school) fishing for the purse seine fisheries in equatorial model regions using a novel travel distance effort metric, truncation of the pole-line-index in Region 8, and grouping of selected CPUE indices to inform regional biomass scaling.

This assessment is supported by the analysis of catch and effort data for pole-and-line and purse seine fisheries (Teares et al., 2022), a novel approach to estimating tag mixing periods (Scutt Phillips et al., 2022), a review and new analysis of skipjack growth (Macdonald et al., 2022), re-analysis of tag seeding experiments to inform tag reporting rate priors (Peatman, 2022), and a new analysis of tagger effects (Peatman et al., 2022).

The main influential change in the progression from the 2019 to 2022 diagnostic model was the introduction of grouped survey fisheries with a separate likelihood component as part of the switch to a catch conditioned model. This resulted in a large increase in estimated spawning potential and a more optimistic stock status compared to the 2019 diagnostic model. We note that conversion to the catch conditioned model without the survey fisheries and their related likelihood component, had minimal impact. The other changes and data updates had minor influences compared to the inclusion of survey fisheries.

In addition to the diagnostic model, we report the results of one-off sensitivity models to explore the impact of key data and model assumptions for the diagnostic model on the stock assessment results and conclusions. We also undertook a structural uncertainty analysis (model grid) for consideration in developing management advice that includes combinations of those areas of uncertainty

considered important.

It is recommended that management advice is formulated from the results of the structural uncertainty grid. The results below are based on equal weighting of all models.

Across the 18 models of the structural uncertainty grid run in this assessment, the most important factors when evaluating stock status were the thresholds used to determine the variable tag mixing periods based on the tag mixing simulation studies, and the alternative growth models. The tag mixing scenarios that allocated longer tag mixing periods (more conservative) resulted in scaling up of the spawning potential and estimated a more optimistic stock status compared to those that allocated shorter mixing periods. However, the temporal dynamics of spawning potential and spawning depletion were similar across the scenarios. The application of the externally estimated growth also resulted in increased levels of spawning potential and a slightly less depleted stock trajectory, but showed some difference in the temporal dynamics compared to the diagnostic model where growth was estimated within the model. There were three outlying models in the structural uncertainty grid that estimated higher spawning potential and less spawning depletion. These models involved the most conservative tag mixing scenario and the external growth.

The general conclusions of this assessment are as follows:

- Spawning potential has remained relatively stable, with fluctuations, until around 2010, after which it declines gradually, driven by trends in the equatorial regions. Spawning depletion has declined gradually since the start of the model period. This decline is largely due to the increasing estimates of the unfished spawning potential and recruitment from 1980 to the recent period, informed by stable CPUE survey trends.
- Average fishing mortality rates for juvenile and adult age-classes increase throughout the period of the assessment.
- Overall median depletion from the model grid for the recent period (2018-2021; $SB_{recent}/SB_{F=0}$) is estimated at 0.51 (80 percentile range 0.43-0.64).
- No models from the structural uncertainty grid estimate the stock to be below the LRP of 20% $SB_{F=0}$.
- Recent (2017-2020) median fishing mortality (F_{recent}/F_{MSY}) was 0.32 (80 percentile range 0.18-0.45).

The most notable feature of the assessment is the estimation that the stock is becoming increasingly depleted over time, a trend which is largely driven by the equatorial regions. Importantly, this trend is driven by an increasing trend in the model estimates of the unfished spawning potential over time, rather than a long-term decrease in the estimates of spawning potential. The assessment is indicating that the spawning potential, as informed by a number of CPUE indices, has not changed substantially in the face of the notable increases in catches over the last 20–30 years, and that the

increased catches have been sustained by increased recruitment levels. The interpretation of stock status based of the ($SB_{recent}/SB_{F=0}$) reference point should bear this in mind.

A number of key research needs have been identified in undertaking this assessment that should be investigated either internally or through directed research. These include: 1. Continued work on abundance indices to explore effort metrics related to travel distance for both the purse seine and pole-and-line fisheries, along with exploration of hyperstability and effort creep, focussing on developing effort creep scenarios for the Japanese pole-and-line fishery. 2. Conduct a study to explore the plausibility and evidence for the model predicted long-term increasing trend in skipjack recruitment. 3. Consider epigenetic approaches for improving growth estimation. 4. Further work on the skipjack model to test alternative model structures to obtain a positive definite Hessian solution and reduce data conflicts. 5. Conduct studies to explore meta-population structure of skipjack, particularly focussed on improving understanding of the linkages between populations in the east Asian waters and those in the broader western and central Pacific.

2 Author Contributions

This assessment was very much a team effort. These tuna assessments are large projects requiring a range of technical proficiencies. Ultimately there has to be a first person on this list and we bestow this to **Claudio Castillo Jordán** who conducted and drove the bulk of the core modelling work and the production of the key model outputs and the many figures required to present the assessment. **Paul Hamer** oversaw and provided direction for the overall assessment project and was responsible for writing up the assessment. **John Hampton** provided an invaluable level of support with the modelling work, interpretation of results and provision of advice to the project team plus ad hoc support for the write-up, it is hard to imagine doing this assessment without his input. **Thom Tears** was responsible for generating the inputs to the assessment and did a remarkable job with such a steep learning curve. **Joe Scutt Phillips** lead the work on the tag mixing simulations that was a great contribution to improving the treatment of tag mixing assumptions. **Jed Macdonald, Jemery Day** and **Arni Magnusson** teamed up on the new growth analysis, and were regulars at our weekly discussions, providing great sounding boards and sources of advice as we broached decision points. **Graham Pilling** provided grounding and direction at key points and importantly conducted the stock projections. **Finlay Scott** built the R shiny app for displaying assessment results and diagnostics, and **Rob Scott** was always there in the weekly meetings providing advice and was a general go to for help with MFCL issues, plus fixing R code to run projections and retrospective analyses. **Nick Davies**, our MFCL wizard, was always available when needed to solve MFCL issues and provide advice and tuition, plus run analysis to explore prickly issues such a pesky negative eigenvalues. He worked with Dave Fournier to generate the new catch conditioned feature in MFCL used in this assessment. **Tom Peatman** conducted a range of supporting analysis including the size composition reweighting, tagger effects corrections and analysis of tag seeding for reporting rate priors. Finally, **Sam McKechnie**, a veteran of the skipjack assessment, was critical when we wondered why did they do that last time, or where is the script for that analysis, and was a great help to Thom on generating the data inputs.

3 Introduction

This paper presents the 2022 stock assessment of skipjack tuna (*Katsuwonus pelamis*; SKJ) in the western and central Pacific Ocean (WCPO; west of 150°W)(Figure 1). Since 2000 (Bigelow et al., 2000), assessments for skipjack in the WCPO have been conducted regularly; the most recent assessments are documented in Hoyle et al. (2010), Rice et al. (2014), McKechnie et al. (2016) and Vincent et al. (2019b). Consistent with previous assessments, the 2022 assessment is conducted using the MULTIFAN-CL (MFCL) stock assessment software (<http://www.multifan-cl.org>) (Fournier et al., 1998; Hampton and Fournier, 2001; Kleiber et al., 2019) and continues the development of the WCPO skipjack stock assessment. Each new assessment can involve updates to fishery input data, implementation of new features in the MFCL software, and consideration of new information on biology, population structure and other important assumptions. These changes are part of ongoing efforts to improve the modelling procedures and reduce the uncertainty of estimates of stock status, fishing impacts, biological and population processes. However, they can result in changes to the estimated historical population dynamics, status of the stock and fishing impacts from previous assessments. Advice from the Scientific Committee (SC) on previous assessments, and the annual SPC (Pacific Community) run Pre-assessment Workshop (PAW) (Hamer, 2022) guide this ongoing process. Notable new features of the 2022 assessment are summarised below and described in the methods and supporting papers.

The objectives of this assessment are to estimate population parameters for skipjack in the WCPO, such as time series of recruitment, total biomass, spawning potential, spawning potential depletion and fishing mortality, that indicate the stock status and impacts of fishing. The outcomes of the stock assessment are used to provide a basis for management advice to the Western and Central Pacific Fisheries Commission (WCPFC). We summarize the stock status in terms of reference points adopted by the WCPFC. The methodology used for the assessment is based on the general approach of integrated modelling (Fournier and Archibald, 1982), and implements a size-based, age- and spatially-structured population model in MFCL. Model parameters are estimated by maximizing an objective function, consisting of both likelihood (data) and prior information components. The assessment uses an ‘uncertainty grid’ of models as the basis for management advice. The uncertainty grid is a suite of models that are selected to incorporate important axes of uncertainty that relate to plausible alternative biological assumptions, data inputs and/or data treatment. The variation in estimates of the key management quantities across the uncertainty grid represent the current appreciation of the uncertainty in stock status and should be considered carefully by managers.

This assessment report should be read in conjunction with several supporting papers, specifically the paper on data inputs and preparatory analyses (Tears et al., 2022), the paper that describes the novel simulations of tag mixing periods (Scutt Phillips et al., 2022), the paper on skipjack growth review and new estimation (Macdonald et al., 2022), and the papers on tag reporting rates and tagger effects estimations (Peatman, 2022; Peatman et al., 2022). Finally, the planning for this assessment was informed by the discussion at the 2022 PAW (Hamer, 2022).

New features applied to this assessment, discussed in more detail in relevant sections of this paper, include:

- Application of a new MFCL catch conditioned approach to the estimation of fishing mortality, plus inclusion of survey fisheries and a likelihood component for the CPUE indices from those survey fisheries (Davies et al., 2022).
- Application of a self-scaling approach to estimate effective sample size, the Dirichlet-multinomial likelihood (i.e., Thorson et al. (2017)), with growth estimation within the diagnostic model.
- Application of variable tag mixing periods for tag release groups based on simulations using individual based modelling of tag mixing processes (Scutt Phillips et al., 2022).
- Development of a alternative growth model based on tag recapture growth increments and daily aging from otoliths (Macdonald et al., 2022).
- Development of new CPUE indices bases on unassociated (free-school) fishing for the purse seine fisheries in equatorial model regions using a novel travel distance effort metric, truncation of the pole-line-index in region 8, and grouping of selected CPUE indices to inform regional biomass scaling (Tears et al., 2022).

4 Background

4.1 Stock structure

Skipjack tuna are widely distributed in tropical to sub-tropical waters in all the major oceans, with the populations in each ocean thought to comprise separate stocks (Wild and Hampton, 1994; Artetxe-Arrate et al., 2021).

In the Pacific Ocean, while skipjack tuna have a continuous east-west distribution, there is evidence from genetic and tagging studies for broad stock separation between the WCPO and the Eastern Pacific Ocean (EPO) (Grewe et al., 2019; Moore et al., 2020). In the western Pacific, warm, poleward-flowing currents near northern Japan and southern Australia seasonally extend the skipjack distribution to about 40°N and 40°S. These limits roughly correspond to the 20°C surface isotherm. Some tagged skipjack tuna have shown large west-east movements (Figure 2), however, skipjack tend to show lower rates of long distance movement compared to larger tropical tunas such as yellowfin and bigeye tuna. In general, skipjack movement is highly variable and is thought to be influenced by large-scale oceanographic variability (Lehodey et al., 1997, 2008). The ENSO (El Niño - Southern Oscillation) cycle is influential on east-west distributional patterns in the equatorial region, where in La Niña phases pooling of warm waters towards the western Pacific tends to concentrate skipjack more towards the western Pacific. In the El Niño phase skipjack are more distributed towards the central Pacific as the warm surface waters spread further to the central and eastern Pacific (Senina et al., 2016).

Finer scale population structure of skipjack in the WCPO is poorly understood, although ‘meta-population’ structure is thought to be present (Grewe et al., 2019; Moore et al., 2020). There is a need for further studies of meta-population structure, especially the relationships between skipjack in the east Asian waters and those in the equatorial western and central Pacific waters. Skipjack in the WCPO are considered a single stock for the purpose of this stock assessment. Spatial structure of the assessment is based on hypothesis and data that support sub-regional structure of population processes, the spatial structure inherent in the tagging and size composition data, and operational features of fishing fleets (Kiyofuji and Ochi, 2016).

4.2 Biological characteristics

Skipjack growth is rapid compared to yellowfin, albacore and bigeye tuna. Approximate age estimates from counting daily rings on otoliths suggest that growth may vary between areas of the Pacific. Analyses of tagging-recapture growth increments suggested that growth varies spatially in the eastern Pacific (Maunder, 2001) and the Atlantic (Gaertner et al., 2008). For the WCPO region, samples from the north Pacific region were estimated to reach approximately 40 cm fork length (FL) by 300 days age (Tanabe et al., 2003), whereas fish sampled closer to the equator, near Papua New Guinea, were estimated to reach 42 cm FL in around 150 days (Leroy, 2000). Despite these earlier studies, growth remains a significant biological uncertainty for skipjack (Ochi et al., 2016), largely because there is no method that can reliably estimate growth across their lifespan. There are no clear validated annual increment structures in skipjack otoliths and daily increments cannot be confidently interpreted beyond about 1 year of age. Aging from spines is also not considered reliable and analysis of tag recapture growth increments is typically restricted to the portion of the growth curve that includes the sizes at which fish are large enough to be caught by hook and line, and the predominant sizes of recaptured fish by purse seine fishery (i.e., 35–60 cm FL). Tag-recapture growth analyses are also plagued by poor quality length measurement data. In this assessment we include an alternative growth model to account for this uncertainty based on the analysis presented in the companion paper on skipjack growth by Macdonald et al. (2022).

The maximum age of skipjack tuna is thought to be around 8-10 years, although most fish captured by the industrial purse seine and pole-and-line fisheries are thought to be less than 4 years old. They can reach sexual maturity by approximately 40-50 cm FL (i.e., within 6 months age) (Ashida et al., 2017; Ohashi et al., 2019) and may reach a maximum size of 90-100 cm.

Estimates of natural mortality rate (M) have previously been obtained using a size-structured tag attrition model (Hampton, 2000), which indicated that M was substantially larger for small skipjack (21–30 cm FL, $M=0.8 \text{ mo}^{-1}$) compared to larger skipjack (51–70 cm FL, $M=0.12\text{--}0.15 \text{ mo}^{-1}$). M -at-age is estimated internal to the model in this assessment. The longest period at liberty for a tag-recaptured skipjack to date is approximately 4.5 years.

4.3 Fisheries

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

The Japanese pole-and-line fleets have historically operated over a large area of the WCPO, although effort and the spatial extent of this fishery has declined substantially since the 1980's (Ducharme-Barth et al., 2022). A domestic pole-and-line fishery occurred in Papua New Guinea (PNG) from 1970 to 1985 and in Fiji since 1974, but this fishery is also no longer operating. Pole-and-line fishing in the Solomon Islands has occurred since 1971 but is now operating at a low level.

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

The industrial purse seine fleets usually operate in equatorial waters from 10°N to 10°S; although a Japanese offshore purse seine fleet operates in the temperate North Pacific (model regions 1, 2, 3) (Figure 1), and takes skipjack seasonally in smaller quantities. The distant-water fleets from Japan, Korea, Chinese Taipei, and the USA capture most of the skipjack in the WCPO, although catches by fleets flagged to or chartered by Pacific Island countries, have increased considerably in recent years. The purse seine fishery is usually classified by set type categories - sets on floating objects such as logs and fish aggregation devices (FADs), which are termed “associated sets” and sets on free-swimming schools, termed “unassociated sets”. These different set types can have different spatial distributions, catch per unit effort (CPUE), and size selectivity of skipjack and other tuna.

Skipjack tuna catches in the WCPO increased steadily after 1970, approximately doubling during the 1980s (Figure 3). Catches further increased during the 1980s due to growth in the international purse seine fleet, combined with increased catches by domestic fleets from ID, PH, and VN. The catch was then relatively stable during the early 1990s, approaching 1 million mt per annum. Catches increased again from the late 1990s as the purse seine fishery further developed and have varied between about 1.5 and 2 million mt since 2007, with a record catch estimated at just over 2 million mt taken in 2019.

Pole-and-line fleets, primarily Japanese, initially dominated the fishery, with their catches peaking at 380,000 mt in 1984, but the relative importance of this fishery has declined steadily over time.

Historically, most of the catch has been taken from the equatorial Pacific (model regions 5, 6, 7, and 8) (Figure 4, Figure 5). During the 1990s, combined annual catches from this region fluctuated around 500,000–800,000 mt before increasing sharply to approximately 1.2 million mt in 2007–2009 (Figure 4). Since the late 1990s, there has been a large increase in the purse-seine fishery in the eastern equatorial region of the WCPO (Region 8, Figure 1), although catches from this region have been highly variable among years depending on the ENSO conditions. Around 2014, catches in the central-eastern equatorial region were particularly high, under the influence of strong El Niño conditions from 2014–2016, thought to drive greater eastward displacement of skipjack (Lehodey et al., 1997; Wang et al., 2014). Catches have since decreased in the eastern region of the WCPO as La Niña conditions have predominated in recent years.

5 Application of a new catch conditioned approach

In previous MULTIFAN-CL assessments of skipjack, catch was predicted by the model (termed ‘catch-errors’ model) with observation error allowed, and the standard deviation of the log-catch deviates assumed to be very small (equivalent to a CV of 0.002). This produced very accurate predictions of observed catches and therefore only a small contribution of the catch to the overall objective function. However, the cost of treating the catch in this way was that effort deviation coefficients had to be estimated as model parameters for each catch observation. Additionally, catchability deviation parameters were required for catch-effort observations for fisheries for which time-series changes in catchability were allowed. While these parameters were constrained by prior distributions and estimation was feasible, it resulted in very large numbers of parameters needing to be estimated by the function minimiser. For example, in the 2019 skipjack assessment 7,750 parameters were estimated for the diagnostic case model. Of these, 5,448 were effort deviation coefficients and 319 were parameters relating to catchability.

In an effort to reduce complexity and paramterisation this assessment makes use of a new feature of MULTIFAN-CL version 2.0.8.5 in which catch is assumed to have no error, i.e., the model is ‘catch-conditioned’ (Davies et al., 2022). This makes it possible to solve the catch equation for fishing mortality exactly, using a Newton-Raphson sub-iterative procedure. The main benefit of this approach is that effort deviation coefficients and catchability-related parameters do not require estimation as model parameters. Effort data for extraction fisheries is not required at all but can be used if available to estimate catchability through regressions of fishing mortality and effort, and this is important for making stock projections based on future effort scenarios. As a result of implementing the catch conditioned approach the number of model parameters in this assessment has been reduced to 2253 in the diagnostic case. This has enabled more rapid model convergence and Hessian matrix computation. The only cost of this approach is that missing catches, which could be accommodated in the catch-errors version if there was an accompanying effort observation, are no longer straight forward to account for. However, this is not an impediment for the key WCPO tuna assessments. The catch conditioned approach also allows (but does not require) the specification of

survey fisheries to provide indices of relative abundance, these are discussed in [Section 6.4](#). In the stepwise model development runs conducted for this assessment, the transition from a ‘catch-errors’ to a ‘catch-conditioning’ model, without implementation of the survey fisheries, did not result in any appreciable change in the estimated quantities of relevance to management advice ([Section 9.1](#)).

6 Data compilation

6.1 General notes

Data used in the stock assessment of skipjack tuna using MFCL consist of catch, effort and length-frequencies for the fisheries defined in the analysis and tag-recapture data ([Figure 2](#)). Improvements in these data inputs are ongoing and more detailed summaries of the analyses and methods of producing the necessary input files are given by [Teears et al. \(2022\)](#). In general, data preparation methods for the 2022 assessment largely follow the previous assessment ([Vincent et al., 2019a](#)), with the exception of the purse seine CPUE standardisation for regions 6, 7 and 8, and the approach to re-weighting of the size composition data ([Teears et al., 2022](#)).

The full details of these analyses are not repeated here, rather, a brief overview of the key features is provided and readers are directed to the relevant papers referenced throughout this section. A summary of the data available for the assessment is provided in [Figure 6](#).

6.2 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 eight model regions ([Figure 1](#)) were created as part of the 2019 assessment in an attempt to capture the seasonal movement dynamics and differences in size composition observed in the Japanese pole-an-line fishery in this region ([Kiyofuji and Ochi, 2016](#); [Kiyofuji et al., 2019a](#)). Additional consideration of the region boundaries attempted to ensure there was sufficient tag releases within each region to estimate the biomass within each region. The 8-region model was compared to the previous 5-region model structure in the 2019 assessment ([Vincent et al., 2019b](#)). SC15 preferred to use the 8-region model as the basis for management advice, although management quantities differed slightly between the two model structures. The 8-region model structure was discussed at the 2022 PAW ([Hamer, 2022](#)). Noting it was endorsed as the preferred structure for management advice at SC15, it is maintained as the spatial structure for this assessment. Some participants at the PAW advocated for a sensitivity analysis of a five region model that involves combining regions 1-4 of the 8-region model. Thus a 5-region model analysis is included as an appendix.

Throughout this report we focus on descriptions of the 8-region model unless otherwise noted. We present the 8-region model as the diagnostic model and present the evaluation of the fit to the data for this spatial stratification, with the 5-region model presented in the appendices ([Section 15.6](#)).

6.3 Temporal stratification

The time period covered by the assessment is 1972–2021. Within this period, data were compiled into quarters (1; Jan–Mar, 2; Apr–Jun, 3; Jul–Sep, 4; Oct–Dec). The assessment included data from the most recent full calendar year (2021), which are finalized late in the development of this assessment and may be subject to some change as they are refined. However, recent experience suggests these changes are likely to be relatively minor and not expected to be consequential for the assessment results and management advice. All data are available up until and including 2021 except for the Japanese pole-and-line survey fishery CPUE that could only be updated to 2020 in the time frame for this assessment, due to COVID–19 related delays with data entry from 2021.

6.4 Definition of fisheries

Extraction fisheries: MFCL requires the definition of “fisheries” that consist of relatively homogeneous fishing units. Ideally, the defined fisheries will have selectivity and catchability characteristics that do not vary greatly over time and space, although some allowance can be made for time series variation. For most pelagic fisheries assessments, fisheries are defined according to gear type, fishing method, region, and sometimes by vessel flag or fleet.

The extraction fishery definitions for the 2022 assessment are essentially the same as for the 2019 assessment, with 31 fisheries defined to account for the harvest of skipjack (see [Table 1](#) for fishery descriptions). Equatorial purse seine fishing activity was aggregated over all nationalities, but stratified by region and set type, in order to sufficiently capture the variability in fishing operations. Set types were grouped into ‘associated’ (which includes all of log, FAD, whale, dolphin, and unknown set types) and ‘unassociated’ (free-school) sets. A fishery for the Japanese purse seine fleet for all set types was assumed in Regions 1–3. Purse seine catch in Region 4 by all nationalities was insufficient (less than 1,000 metric tons) to warrant creating a fishery. Additional fisheries were defined for pole-and-line fisheries in each region and miscellaneous fisheries (gillnets, ringnets, handlines etc.) in the western equatorial area, mostly Region 5. A longline fishery was defined in each region to hold the long time series of skipjack length composition data from Japanese research longline cruises in the WCPO and more recently, observer-measured length composition samples. Catch data from the Japanese troll fishery in Region 1 was combined with the pole-and-line fishery in this region. Miscellaneous gears from ID, PH, and VN were aggregated by nationality, with the exception of a fishery created for the domestic ID and PH purse-seine fleets. The catch time series by gear and region that were input to the assessment are shown in [Figure 7](#), [Figure 8](#), [Figure 9](#), [Figure 10](#).

Survey fisheries: The catch-conditioned approach ([Section 5](#)) allows the specification of “survey fisheries” that are used to provide standardised CPUE indices of abundance. Survey fisheries may be the same fisheries as the extraction fisheries, but when used as a survey they do not take any catch, and must have effort data to allow modelling of CPUE. For this assessment, we have defined ten survey fisheries, six based on standardised Japanese pole-and-line CPUE in regions 1–4, 7 and 8, and

four based on standardised purse seine CPUE in regions 5-8 (Tears et al., 2022). Survey fisheries may be grouped if it is felt that the CPUE reflects differences in average abundance among regions. For this assessment, the pole-and-line survey fisheries are grouped, as are the purse seine survey fisheries in regions 6–8. This grouping allows the CPUE to inform on regional as well as temporal abundance changes. The purse seine survey fishery in Region 5 is based on GLM standardised CPUE of smaller Philippines-based vessels that operate differently to the more industrial fleets in regions 6–8 for which a spatiotemporal CPUE model was developed, and therefore is not grouped with the other purse seine survey fisheries (Table 1). The indices from these survey fisheries cover varying periods of time depending on the availability and coverage of data (Figure 6, Figure 11, Figure 12) (Tears et al., 2022).

6.5 Catch and effort data

6.5.1 General characteristics

Catch and effort data were compiled by year and quarter according to the fisheries defined in Table 1. The catches of all fisheries were expressed in weight of fish, with the exception of the longline fishery, the catches of which are very small and set at a nominal low level expressed in numbers of fish (i.e., 500).

Total annual catches by major gear categories for the WCPO are shown in Figure 3 and a regional breakdown is provided in Figure 4. The spatial distribution of catches over the past ten years is provided in Figure 5. Discarded catches are estimated to be minor and were not included in the analysis. Catches in the northern most Region 2 are highly seasonal, mostly occurring in quarters 2 and 3 when surface water temperatures are suitable (warmer) for skipjack. Quarterly catch histories provided to the model for the different fishery groupings are displayed in: ID, PH, VN domestic fisheries (Figure 7), pole-and-line fisheries (Figure 8), purse fisheries in regions 1, 2 and 3 (Figure 9), and associated and unassociated purse seine fisheries in regions 5, 6, 7, 8 (Figure 10).

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

- The development of the Japanese off-shore purse-seine fishery in Regions 1, 2, and 3 since the mid-1990s.
- The virtual cessation of the domestic pole-and-line fisheries in PNG and Fiji and the recent low catches from the Solomon Islands fishery;
- The general decline in the Japanese distant-water pole-and-line fisheries in the equatorial regions, particularly in the eastern Region 8.
- 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 further to the east.
- Large changes in the purse seine fleet composition and efficiency of the fleet.

- The steady increase in catch for the domestic fisheries of Indonesia, Philippines, and Vietnam.
- The record catch of skipjack estimated in 2019.

6.5.2 Purse seine

For the industrial purse seine fisheries predominantly operating in regions 6, 7 and 8, catch by species within each set type (associated or unassociated) is determined by applying estimates of species composition from observer-collected samples to total catches estimated from raised logsheet data (Hampton and Williams, 2016; Peatman et al., 2021). For the Japanese (JP) fleet for which there is greater confidence in species-based reporting, reported catch by species is used. Purse seine catch for PH and ID domestic purse seine fisheries, predominantly operating in Region 5, was derived from raised port sampling data provided by these countries.

Effort data for purse seine fisheries are defined as number of sets (except for the PH/ID domestic purse seine fishery where effort is vessel days), specified by set type (associated or unassociated) and were included in the assessment to provide estimates of catchability over time, which may be of interest for management purposes. Recent estimates of catchability for purse seine fisheries are also required for projections in which effort is specified for the projection period. A change from days fishing and/or searching to number of sets for the main purse seine fisheries in the model was made for the 2019 assessment and did not affect model results but avoids the issue of recent nominal effort creep from changes in reporting practices that have attributed fewer days as searching. The use of sets as the effort metric is also more appropriate for management strategy evaluation and projections (SPC-OFP, 2013).

Survey fisheries—purse seine CPUE: Survey fisheries were created for purse seine fisheries in regions 5, 6, 7, and 8 using standardised CPUE indices developed using different approaches. In Region 5 a standardised CPUE index was developed for the PH domestic purse seine fishery, including data from the PH archipelagic waters and the high seas pocket 1 (HSP-1) (Bigelow et al., 2019). This index was developed using generalized linear models (GLMs) following the methods of Bigelow et al. (2019) that were applied for the 2019 assessment. Briefly, standardized CPUE was estimated with a GLM by removing effects due to vessel and fishing ground (area). The index predicted quarterly CPUE with a YR:QTR, Area (fishing ground) and Vessel effects (Figure 11) (Teears et al., 2022).

CPUE for the purse seine survey fisheries operating within regions 6, 7 and 8 were analysed based on observer data available from the Pacific Islands Regional Fisheries Observer Program since 2010. Observer data were chosen over the longer time series of logbook data due to it being considered more accurate and consistent in the approach used to estimate species composition of the purse seine catches (Hamer, 2022). Exploration of various approaches to purse seine CPUE standardisation were explored with various models including both associated and unassociated set types, similar to that conducted by Vidal et al. (2020). Ultimately it was decided to develop an index based on

unassociated sets only. It was thought that an abundance index that included associated effort would be more prone to hyperstability effects over time due the technological advancements in FAD fishing, and that this would be difficult to account for in a standardisation. Data were therefore filtered to only include unassociated sets and to remove any vessel in a given year with greater than 70% of their sets made on FADs, ie., removing vessels that primarily focussed on FAD fishing. Therefore, the derived indices were essentially based on unassociated fishing without FAD specialists (Teears et al., 2022). The other key consideration for the purse seine index was how to measure effort. It was considered that set as an effort metric in a purse seine CPUE model would also be subject to hyperstability due to the ability of vessels to be selective as to which schools they set on based on a priori information on school sizes. In an attempt to better quantify effort, we used the Vessel Monitoring System (VMS) data to define effort as the daytime distances travelled between unassociated sets, calculated as the sum of the distances travelled hourly during daytime hours with a buffer of 30 minutes prior to sunrise and after sunset (Teears et al., 2022). The effort for the first set was assigned the trip-specific median value.

A spatiotemporal delta GLM approach implemented in VAST (Thorson et al., 2015) was used to develop the unassociated purse seine survey indices for regions 6, 7 and 8. Full details of the CPUE standardisation are in Teears et al. (2022). Briefly, a VAST model with 210 spatial knots with uniform distribution was developed. Building on the previous research (Vidal et al., 2020), the model selection identified three variables as influential on catch rates: species cluster, vessel length, and ENSO (a spatially varying covariate). The resulting indices are displayed in Figure 11.

The purse seine survey fisheries in regions 6, 7, and 8 are assumed to have the same catchability and are grouped to provide information on biomass scaling among those model regions. A time and space invariant penalty weight of 8.3 (equivalent to CV of 0.25) is applied to these indices. A time and space invariant penalty weight of 28.6 (equivalent to CV of 0.13) was applied for the region 5 pure seine index. These penalty weights/CV's were derived from the CPUE standardisation analyses.

6.5.3 Longline

Longline fisheries take a negligible proportion of the total skipjack catch. These fisheries were included in the model solely to utilize the available size frequency data and to attribute a small number of tag returns to. Catches are set at low, arbitrary levels.

6.5.4 Pole-and-line

Pole-and-line catches are provided for the Japanese DW and OS fleets from logbook data since 1972. For other pole-and-line fisheries catches are provided from logbook records provided to SPC. Survey fisheries were created for the Japanese pole-and-line fisheries in regions 1, 2, 3, 4, 7, and 8. These indices were developed consistent with the VAST spatiotemporal delta GLM approach used

for the 2019 assessment (Ducharme-Barth et al., 2019; Kinoshita et al., 2019), and further described in Teears et al. (2022). Nominal fishing-vessel-day was used as the unit of effort for the pole-and-line survey fisheries. The spatiotemporal CPUE model was fit to the DW and OS (fishing closer to Japan) fleet data together and included spatial and environmental components. The following information was included for the CPUE calculation: date, skipjack catches in weight, number of poles, gross registered tonnage (GRT), and vessel identity. The data were further categorized by vessel size with vessels between 20 and 199 GRT defined as OS and vessels greater than or equal to 200 GRT defined as DW. The implementation of important technological innovations are available only in the DW fleet, and these are low temperature live bait tank, the first and second generations of bird radar, sonar, and onboard NOAA meteorological satellite image receiver. In order to have more complete spatiotemporal coverage within the model time period and regions, the DW and OS trips were combined in the CPUE modelling following methods in Kinoshita et al. (2019). Further, Kinoshita et al. (2019) also showed that for nominal CPUE in spatiotemporal strata that were fished by both DW and OS vessels the magnitude and trend of mean catch rates was similar. Given the joint modelling approach of the DW and OS trips, and the fact that device information were unavailable for OS vessels, device covariates were not included in the spatiotemporal VAST model (Kinoshita et al., 2019). The analysis used 280 spatial knots uniformly distributed across the spatial domain, as recommended by Ducharme-Barth et al. (2022) to improve estimation in the case of the contracting spatial coverage of the data overtime. Final model configurations include number of knots, a uniform knot distribution, random effects (RE) for vessel identification, catchability covariates (number of poles, vessel class (OS, DW), vessel GRT), and density covariates (SST). The standardised CPUE time series are displayed in Figure 12. We applied penalty weights of 5.2 for the grouped PL fisheries equivalent to a CV of 0.31.

Note on effort creep: While nominal effort creep is accounted for by using sets as the unit of effort in relation to fishing mortality, for the survey fishery CPUE indices effort creep in terms of increased capture effectiveness per set, due to, for example, increased use of technology over time requires consideration. This type of effort creep has potential to influence rates of harvest per unit of effort (at the set level). Part of the reason for not using associated effort and sets as the effort metric for purse seine CPUE indices in regions 6, 7 and 8, is that it is likely subject to some level of effectiveness creep, however, it is unclear on the rates and dynamics of such effectiveness changes overtime. It is thought that developing purse seine indices based on unassociated fishing removes some clear possibilities for effectiveness creep related to the use of technologies such as FADs with satellite/sonar buoys and their sophisticated associated software. The use of distance between sets as an effort metric for unassociated (free school) fishing is thought to be less prone to hyperstability than sets, and reflects the searching required to find a school of suitable size to set on. While technology can also influence searching efficiency, the purse seine indices are only calculated from 2010, so any effectiveness changes are expected to be of minimal influence on the trend in the index as most modern searching technologies such as sonar, bird radars etc. were implemented by the purse seine fishery prior to 2010. Effort creep was discussed at length at the

PAW (Hamer, 2022), including the difficulty in quantifying this across different effort metrics, gears, fleets and time. For this assessment we conducted a sensitivity analysis that assumes a year-on-year 1% increase in effectiveness of the effort metrics for all survey fisheries. We also note the paper by Matsubara et al. (2022) that explores technology uptake and potential effort creep in the Japanese pole-and-line fishery. This paper indicates that the more sophisticated searching technologies for this fishery; onboard NOAA meteorological satellite image receiver (NOAA receiver), and the first and second generation bird radar and sonar, plus low temperature live bait tanks started to become prevalent in the fishery from 1981. They also suggested at least a 0.2% per quarter increase of fishing effort/effectiveness after installing these devices. The effort creep sensitivity of 1% per year applied in this assessment since 1972 is somewhat consistent with this.

Further work to determine appropriately justified effort creep scenarios is required for all survey fisheries, particularly the pole-and-line indices that cover the longer time period, so that the PAW can provide informed recommendations on alternative scenarios to include in future skipjack assessments.

6.5.5 Other fisheries

Effort data for the ID, PH, and VN surface fisheries and research longline fisheries were unavailable. However, as these fisheries are not part of the survey fisheries, with the new catch-conditioned approached effort data is not necessary for extraction fisheries. Catch estimates for these fisheries are derived from various port sampling programmes dating back to the 1960s for ID and the PH, and early 2000s for VN (Oceanic Fisheries Programme, 2021).

6.6 Size data

6.6.1 Purse seine

Only length frequency data (FL in 2 cm bins) are used in the skipjack assessment. Length frequency data available for the 2022 assessment are summarised in Teears et al. (2022). The purse seine length frequency data are derived from long-term port sampling of primarily US purse seiners in Pago Pago, and samples taken at sea by observers corrected for grab-sample bias (Lawson, 2011). Size data are available for both associated and unassociated set types, with sample numbers for most fisheries increasing from the early 2000s (Teears et al., 2022) (Figure 13). Sample numbers increased greatly in region 5 from 2010 onwards due to increased sampling of the PH and ID purse seine fisheries under improved port sampling programmes.

6.6.2 Pole-and-line

Size composition for pole-and-line fisheries primarily come from port samples or tagging cruises, with the exception of regions 1, 2, 3, 4, and 7 where length data are available from the Japanese offshore and distant-water fleets from the beginning of the model period until 2021 (Teears et al.,

2022). These data are derived from port sampling and sampling onboard research and training vessels. The length frequency for the Japanese pole-and-line fishery in the database held by SPC were revised in 2019 based on the methods described in [Kiyofuji et al. \(2019b\)](#). Length data for the equatorial pole-and-line fisheries are available from both the Japanese distant-water fleet and domestic fleets.

The data from the pole-and-line fishery in Region 8 was dominated by data for the Japanese fleets (1974–2004) with additional data from Fiji in the 1990s. Length data from the pole-and-line fishery in Region 5 was filtered to only include data from ID because it constituted the majority of the catch. Data for ID are sporadically available early in the time series, which precluded weighting samples from other countries based on the catch, but were readily available in 2009–2018. The data from the pole-and-line fishery in Region 6 are a large dataset from multiple countries with the Solomon Island and PNG contributing the majority of the length composition samples. The pole-and-line fisheries in the northern regions generally catch smaller fish than the equatorial fisheries in Regions 5–8, although over the model period, there is a slight increase in the length of fish sampled from the pole-and-line fisheries in the four northern regions, with a substantial change in Region 3 during the early 2000s ([Tears et al., 2022](#)). No systematic trends in the length composition were evident in Regions 5–8.

6.6.3 Longline

Longline fisheries typically do not target skipjack but do catch small numbers of larger skipjack within the 50–90 cm FL range as bycatch, that are usually discarded. Japanese research vessels have routinely collected measurements of the length of skipjack caught by longline since the start of the assessment time period. Japanese research data are only available sporadically in several regions and sample sizes have decreased or ceased to exist in areas where Japanese longline fishery effort has declined. From the 2019 stock assessment, data for other flags conducting longline fishing have been added to these fisheries, resulting in substantial increases in sample size and temporal coverage. Most of these samples were collected by observer programs (since the 2000s). The size data from the longline fisheries are an important component of the model because they provide information regarding the presence of larger-sized skipjack that are not typically caught by the purse seine and pole-and-line surface fisheries. Without this information, the model would have difficulty estimating this “cryptic biomass” component of the population.

6.6.4 Other fisheries

Size composition data for the PH domestic fisheries were collected by a sampling program conducted in the PH in 1993–1995 and augmented with data from the 1980s. In addition, data collected during 1997–2006 under the National Stock Assessment Project (NSAP), and in more recent years under the various West Pacific East Asia (WPEA) projects, were included in the current assessment. The ID domestic fishery is given its own selectivity function as a result of the addition of data from the

large number of measurements from recent sampling under the WPEA project. Similarly, length data for the VN domestic fishery has increased substantially in recent years and is given a separate selectivity in the model (Table 1).

6.7 Re-weighting of size composition data

Statistical correction of size composition data is required as length samples are often collected unevenly in space and time. The methods for re-weighting of the size composition data are detailed in Teears et al. (2022). For the extraction fisheries, re-weighting of composition data is required to ensure that sampling biases in space, time, and the fleets providing data, are minimised so that size composition data better reflect the composition of the overall removals. Strata-specific size data samples are therefore re-weighted by catch for the extraction fisheries (Peatman et al., 2020). For the survey fisheries, re-weighting of composition data is required to ensure that the size composition of the survey fishery indices reflect the size component of the population that is being sampled by the index fisheries through time. Strata specific samples are therefore re-weighted by relative abundance using the standardised CPUE. Given that the same composition data were used for both the extraction and index fisheries, the observed number of size-frequency samples input to the assessment was divided by two after the re-weighting process for both the extraction and survey fisheries where appropriate. A summary of length sample coverage across time for the fishery/region groups is provided in Figure 13.

6.8 Tagging data

A large amount of tagging data is available for incorporation into the assessment. The treatment of the tagging data for this assessment followed the methods described in Vincent et al. (2019a) and is further summarised in Teears et al. (2022). The data were available from SPC's Skipjack Survey and Assessment Program (SSAP) carried out during 1977–80, the Regional Tuna Tagging Project (RTTP) during 1989–92 (including affiliated in-country projects in the Solomon Islands, Kiribati, Fiji and the Philippines), and the Pacific Tuna Tagging Program (PTTP) which has been ongoing since 2006 with the most recent tagging cruise included in the assessment being the Western Pacific 5 (WP5) tagging cruise in 2019 (Figure 2, Figure 14). Tags were released using standard tuna tagging equipment and techniques by trained scientists and technicians. Tags have been returned mostly from purse seine vessels via processing and unloading facilities throughout the Asia-Pacific region.

Tagging data from regular Japanese research tagging cruises were available for the period 1989–2021. Assessments prior to the 2019 assessment did not use Japanese programme tag releases prior to 1998 because the tag releases were not measured at the time of tagging. However, for the 2019 assessment these earlier data were included with release lengths estimated by sampling from the available measured release lengths from the Japanese tagging program (Vincent et al., 2019a). The same approach was used for this assessment.

As in recent tropical tuna assessments, the numbers of tag releases were adjusted for a number of sources of tag loss, e.g., unusable recaptures due to a lack of adequately resolved recapture data, estimates of tag loss (shedding and initial mortality) due to variable skill of taggers, and estimates of base levels of tag shedding/tag mortality, in combination referred to as ‘tagger effects’. The procedures used in re-scaling the releases for tagger effects are described in detail in [Peatman et al. \(2022\)](#). Essentially the re-scaling preserves the recovery rates of tags from the individual tag groups that would otherwise be biased low when an often significant proportion of recaptures cannot be assigned to a recapture category in the assessment.

There is a delay between tagged fish being caught, the tag being reported, and the data being entered into the tagging database. If this delay is significant then reported recapture rates for very recent release events will be biased low and will impact estimates of fishing mortality in the terminal time periods of the assessment. For the Japanese tagging program, tags are generally returned more promptly; thus it was possible to include tag releases to the end of 2019 in the assessment. For the PTPP, efforts have been made to improve the timeliness of tag recaptures being reported to SPC and the validation of recaptures. This has meant that tag releases up to and including the 2019 WP5 tagging cruise can be used in this assessment, with recapture data included up to 2021 for the Japanese tagging programme and the PTPP.

For incorporation into the assessment, tag releases were stratified by release region, year/quarter of release, and length at release using the same size bins as the length-frequency data. Tag release events that had less than 30 tags released per event were removed from the analysis to reduce the computational load otherwise created by the many small release events, and allow for better model convergence. A total of 383,387 effective releases were classified into 328 tag release groups. The returns from each size-class of each tag release group (63,305 total usable tag returns) were then classified by recapture fishery and time period (quarter). A summary of tags recaptured by release and recapture regions is presented in ([Figure 14](#)). Tag return data were aggregated across set types for the purse seine fisheries in each region because tag returns by purse seiners were often not accompanied by information concerning the set type. The assessment model was configured to predict tag recaptures by these fisheries grouped together by region.

7 Model description

7.1 General characteristics

The model can be considered to consist of several components, (i) the dynamics of the fish population; (ii) the fishery dynamics; (iii) the dynamics of tagged fish; (iv) the observation models for the data; (v) the parameter estimation procedure; and (vi) stock assessment interpretations. Detailed technical descriptions of components (i)–(iv) are given in [Hampton and Fournier \(2001\)](#) and [Kleiber et al. \(2019\)](#), but a brief summary will be given below. In addition, we describe the procedures followed for estimating the parameters of the model and the way in which stock status

conclusions are drawn using a series of reference points.

7.2 Population dynamics

The model partitions the population into eight spatial regions and 16 quarterly age-classes. The last age-class comprises a “plus group” in which mortality and other vital rates are assumed to be constant. The population is “monitored” in the model at quarterly time steps, extending through a time window of 1972–2021. The main population dynamics processes are as follows.

7.2.1 Recruitment

Recruitment is defined as the appearance of age-class 1 quarter fish (i.e. fish averaging ~ 23 cm given the current diagnostic model growth curve) in the population. Tropical tuna spawning does not generally follow a clear seasonal pattern but occurs sporadically when food supplies are plentiful (Itano, 2000). The assessment model assumed that recruitment occurs instantaneously at the beginning of each quarter. This is a discrete approximation to continuous recruitment, but provides sufficient flexibility to allow a range of variability to be incorporated into the estimates as appropriate.

The proportion of total recruitment occurring in each region was initially set relative to the variation in average regional catch and then estimated during the later phases of the fitting procedure. 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 terminal recruitments have often been fixed at the mean recruitment of the rest of the model period to prevent the instability that has been detected by retrospective analyses. This approach has been continued here with the terminal two recruitments fixed at the geometric mean, which is appropriate for a log-normally distributed random variable.

Spatially-aggregated recruitment was assumed to have a weak relationship (CV of log-recruitment deviates set to 2.2) with total spawning potential in the preceding quarter, according to a Beverton and Holt stock-recruitment relationship (SRR) with a fixed value of steepness (h). Steepness is defined as the ratio of the equilibrium recruitment produced by 20% of the equilibrium unexploited spawning potential to that produced by the equilibrium unexploited spawning potential (Francis, 1992; Harley, 2011). As has been the practice in other tuna stock assessments, h was fixed at 0.80 in the diagnostic model, and values of 0.65 and 0.95 included in the structural uncertainty grid (ISSF, 2011).

The high CV (2.2) of the log-recruitment deviates, computed annually, ensured that the SRR had negligible impact on the estimation of recruitment and other model parameters, as recommended by Ianelli et al. (2012). The SRR was estimated over the period 1984–2020 to prevent the earlier recruitments, which appear to be part of a less productive recruitment regime, from influencing the relationship.

7.2.2 Initial population

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

7.2.3 Growth

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

In the 2019 skipjack assessment, growth was estimated internally from the length composition data in a preliminary run of the assessment model in which the effective sample sizes of the length frequency data were increased relative to those used in the diagnostic case model. This was necessary in order to obtain stable growth parameter estimates. Those estimates were subsequently input as fixed parameters into the 2019 assessment runs with the effective sample sizes restored to lower levels. In the current assessment, growth was again estimated internally from the length composition data, but we found that the effective sample sizes estimated using the Dirichlet-multinomial likelihood (Figure 20) were sufficient to allow stable estimation of growth. The growth curve estimated in this way was virtually identical to that used for the diagnostic model of the 2019 assessment, but with a more realistic variance across the growth curve (Figure 15).

An alternative externally estimated growth curve was also developed for this assessment following discussions on growth uncertainty at the PAW (Hamer, 2022). This alternative approach does not use length composition data to estimate growth directly but estimates a growth curve outside the integrated assessment framework, combining data from daily otolith aging (Leroy, 2000) and screened growth increment data from tag recaptures. These data are used to estimate a von Bertalanffy growth curve by incorporating separate likelihood components from the otolith data and the lengths of both the tag releases and the tag recaptures, and estimating the age at release for each tagged fish (Figure 15). Full details of the alternative growth model can be found in Macdonald et al. (2022).

Growth modelling from the externally estimated, continuous von Bertalanffy growth curve needs to be indexed appropriately to map length-at-age measured using a continuous time step, to a length-at-age measured using discrete time steps. This assessment has a discrete time step of three months (one quarter). Because the recruitment of individuals to the population in the first age class is somewhat imprecise, and may integrate additional mortality for larvae and very young fish, there

is some room to vary the assignment of fish to the model age class one quarter as a function of their expected biological age (in days) (i.e., from the external growth model). Due to the presence of very small fish (<20cm) in the catch length frequency data, and the implications for the estimated standard deviation that would be required to adequately fit the variability of length-at-age for these younger and smaller fish, the mean length-at-age for fish aged one quarter was set to the mean length at age of 30 days from the externally estimated continuous growth curve (Macdonald et al., 2022). The Schnute (1981) parametrisation of the von Bertalanffy growth curve, requires lengths at two age classes (A_{min} and A_{max}) with the length at age 16 quarters (L_{max} for $A_{max} = 16$). For input to the MFCL model we similarly adjusted the length at 48 months to that estimated at 46 months by Macdonald et al. (2022).

7.2.4 Movement

Movement was assumed to occur instantaneously at the beginning of each quarter. Parameters were estimated for regions that shared a common boundary, but fish can move between non-contiguous regions in a single time step due to the “implicit transition” computational algorithm employed (see Hampton and Fournier, 2001 and Kleiber et al., 2019 for details). Movement is parameterized as the proportion of fish in a given region that move to the adjacent region. Across each regional boundary in the model, movement is possible in both directions for the four quarters. As in the 2019 assessment, movement was assumed to be constant across ages for this assessment because the increased number of parameters does not yield sufficient improvement in model fit to the data to warrant their inclusion (Vincent et al., 2019b). The seasonal pattern of movement persists from year to year with no allowance for inter annual variation in movement. A prior mean of 0 was assumed for all movement coefficients, but the penalty was weak for deviations from the mean. Starting parameter estimates for the movement coefficients were specified based on the proportion of tags returned within each region and quarter relative to total tags returned for releases in the same region.

Movement among model regions is still considered a significant uncertainty and may not be incredibly well informed by the data, particularly for regions where less tag release effort has occurred. As an alternative to the internally estimated movement, similar to the 2021 South Pacific albacore assessment (Castillo Jordan et al., 2021), we conduct a sensitivity that specifies movement using externally estimated movement coefficients from the skipjack SEAPODYM (Senina et al., 2020b; Lehodey et al., 2008) model (see Section 8.2). Different to the internally estimated movement, the SEAPODYM movement coefficients are specified for quarterly age groups, taking advantage of SEAPODYM’s more highly resolved approach to modelling skipjack movement.

7.2.5 Natural mortality

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

7.2.6 Sexual maturity

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

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

where the proportion mature for each length bin was calculated using Equation 1 and the center of the length bin (Figure 16). For the 2022 assessment we continued with the maturity-at-length curve for the tropical region as was used in the 2019 assessment (Figure 16).

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

7.3 Fishery dynamics

7.3.1 Selectivity

Selectivity is often modelled as a functional relationship with age to reduce the number of parameters estimated within the stock assessment model. Examples include a logistic curve to model monotonically increasing selectivity and various dome-shaped curves to model fisheries that select neither the youngest nor oldest fish. Modelling selectivity with separate age-specific coefficients (with a range of 0-1), constrained with smoothing penalties, allows more flexibility but has the disadvantage of requiring more parameters. Instead, we have used a method based on a cubic spline interpolation. This is a form of smoothing, but the number of parameters for each fishery is the number of cubic spline “nodes” that are deemed to be sufficient to characterize selectivity over the age range. The number of nodes varied by fishery and were selected because they provided an improvement in Akaike Information Criterion (AIC), increased model stability by reducing the number of parameters, or removed unreasonable trends in the selectivity-at-age.

All selectivities were modelled with cubic splines with the number of nodes indicated in Table 1. In all cases, selectivity was assumed to be time-invariant and fishery-specific. However, a single selectivity function could be “shared” among a group of fisheries that have similar length compo-

sitions or were assumed to operate in a similar manner. This grouping facilitates a reduction in the number of parameters estimated and can provide insight into the regional abundance of fish of specific sizes. Selectivity groupings are indicated in [Table 1](#).

The primarily Japanese pole-and-line fisheries in Regions 1-4, 7, and 8 were assumed to share selectivity (4 node cubic spline function). Selectivity for the equatorial purse seine fisheries was grouped by set type, i.e., selectivity was shared among regions separately for associated and unassociated sets fisheries. The ID, PH and VN domestic fisheries in Region 5 were given separate selectivity functions with differing numbers of nodes, and the longline fisheries were modelled with three node splines ([Table 1](#)).

7.4 Dynamics of tagged fish

Tagged fish are modelled as discrete cohorts based on the region, year, quarter, and length at release for the first 12 quarters after release. The tags released are assigned to a quarterly age bin according to the length at release and the estimated growth curve and its associated standard deviations. Subsequently, the tagged fish are pooled into a common group in order to limit memory and computational requirements.

7.4.1 Tag reporting

In theory, tag-reporting rates can be estimated internally within the stock assessment model. In practice, experience has shown that independent information on tag-reporting rates for at least some fisheries tends to be required for reasonably precise estimates to be obtained. We provided reporting rate priors for all reporting groups that reflect independent estimates of the reporting rates and their variance ([Peatman, 2022](#)). We also made some assumptions regarding fisheries that were similar to those with independent estimates, but increased the prior variance. For others where we felt there was very little information to inform priors and variance, uninformative priors were allocated. The prior reporting rates and penalty terms were informed from analyses of tag seeding experiments reported in [Peatman \(2022\)](#). For the RTTP and PTTP purse seine fisheries in equatorial regions 6, 7 and 8, relatively informative priors were formulated given the larger extent of tag seeding information available. Similarly, an informative prior was provided for PTTP program purse seine fishery in region 5 due to more recent tag seeding experiments. Priors for the SSAP are very uncertain as this program predates any tag seeding experiments; unlike the 2019 assessment that provided informative priors for the SSAP, we provided relative uninformative priors centred on 0.5 with large variance. All reporting rates within a tagging program were assumed to be time-invariant, but the recent analysis of tag seeding has been used to update the PTTP priors for this assessment.

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.

7.4.2 Tag mixing

The population dynamics of the fully recruited tagged and untagged populations are governed by the same model structures and parameters. The populations differ in respect of the recruitment process, which for the tagged population is the release of tagged fish, i.e., an individual tag and release event is the recruitment for that tagged population. Implicitly, we assume that the probability of recapturing a given tagged fish is the same as the probability of catching any given untagged fish in the same region and time period. For this assumption to be valid, either the distribution of fishing effort must be random with respect to tagged and untagged fish and/or the tagged fish must be randomly mixed with the untagged fish. The former condition is unlikely to be met because fishing effort is almost never randomly distributed in space. The second condition is also unlikely to be met soon after release because of insufficient time for mixing to have taken place.

Depending on the distribution of fishing effort in relation to tag release sites, the probability of capture of tagged fish soon after release may be different to that for the untagged fish. It is therefore desirable to designate one or more time periods after release as “pre-mixed” and compute fishing mortality for the tagged fish based on the actual recaptures, corrected for tag reporting and tagging effects, rather than use fishing mortalities based on the general population parameters. This, in effect, desensitizes the likelihood function to tag recaptures in the specified pre-mixed periods while correctly removing fish from the tagged population that is present after the “pre-mixed” period.

In the 2019 diagnostic model it was assumed that tagged skipjack gradually mix with the untagged population at the regional level and that this mixing process is complete by the end of the quarter of release (referred to as a mixing period of ‘one quarter’). The assessment also included an alternative scenario where tags were assumed to be mixed by the end of the second quarter after they were released in (mixing period of ‘two quarters’). These two fixed mixing period options were included as an axis in the structural uncertainty grid and had the most influence on management quantities. It is recognised that the mixing period assumptions is a key uncertainty. Further, given the variation in locations of tag releases, and environmental influences on skipjack movement over time, the spatial extent of assessment regions, and spatiotemporal variability in fishing effort at the sub-regional scale, mixing periods are likely to be variable among release groups. Further research to better inform assumptions of tag mixing periods was recommended (Vincent et al., 2019b), prompting the development of the tag mixing simulation modelling study described in Scutt Phillips et al. (2022).

For the 2022 assessment we have moved away from the unlikely assumptions of fixed mixing periods across all tag release events and implemented variable mixing periods based on tag mixing simulations by Scutt Phillips et al. (2022). The approach simulated mixing periods specifically for

each release group, taking into account the unique locational and temporal (environmental, fishing effort) contexts of each release event constituting the group that may result in different rates of mixing of released fish. It applies an individual based Lagrangian model (Ikamoana) (Scutt Phillips et al., 2018) to track movement of individual fish (particles) and quantify the fishing pressure that individuals experience across their dispersal trajectories. Ikamoana uses the forcings and parameters such as fishing mortality, growth and natural mortality that have been estimated from real data by the Eulerian model SEAPODYM (Lehodey et al., 2008; Senina et al., 2020a). The individual based modelling approach simulates post-tagging movement and probability of capture for all individual tagged fish in a release group while also simulating the broader untagged population as the simulated ‘truth’ with which to compare fishing pressure to the tagged groups. The key results from the individual simulations are trajectories of survival and capture probabilities that can be compared between tagged and untagged populations. By comparing distributions of capture probabilities for the tagged and untagged fish it is possible to use criteria to estimate at what period after release, fish from particular tagging events are experiencing sufficiently similar fishing mortality as the untagged fish for particular model regions to be considered to be fully mixed.

The simulation studies provided summary distributions of recapture probability after increasing time-at-liberty from an assumed mixing period for both the tagged fish and the untagged population. By comparing these distributions of final recapture probability for individual release events assuming different mixing periods, i.e., 0 quarter, 1 quarter, 2 quarter, 3 quarters after release, it is possible to make a judgement on whether the tagged fish are likely to be mixed (i.e., have sufficiently similar probability of recapture as the untagged population). For determining what constitutes sufficiently mixed we used the non-parametric Kolmogorov-Smirnov D statistic to indicate the degree of similarity between the distributions of recapture probability for the tagged and the untagged (reference) population. Because formal statistical tests of differences in distribution are overly sensitive for our application we chose to use three levels of similarity (D 0.1, 0.2 and 0.3) for determining alternative mixing periods for each tag release group. We chose these three D values based on what we considered covered a range of similarity levels that could be considered mixed based on qualitative comparison of the distributions of recapture probabilities between tagged and untagged groups, see examples in Figure 17. D of 0.1 can be considered as indicating that the distributions are extremely similar, D of 0.2 indicates the distributions are very similar, and D of 0.3 indicates the distributions are quite similar. For each simulated release group the D statistics were calculated at mixing periods of 0, 1, 2, and 3 quarters.

In the previous assessment the minimum mixing period applied was one quarter. The approach of determining an appropriate mixing period using the results from tag mixing simulations can, in some cases, suggest a mixing period of zero quarters. Whilst in reality, instantaneous mixing is unlikely, here it can be considered an indication that the estimated mixing of a particular release group was so rapid, or the distribution of fishing effort within the release region was sufficiently homogeneous, that the evolution through time of its experienced fishing pressure was comparable to

the untagged population. In these cases, assuming an extra quarter for mixing resulted in negligible improvements to mixing statistics.

The D statistics for each time period (0, 1, 2, and 3 quarters) after release then formed the basis for classifying the release groups as mixed. The lower the D value the more similar the recapture probability distributions need to be between tagged and untagged fish before a tag group will be considered mixed, therefore lower D values have the implication of requiring longer mixing periods, and consequently, of admitting less recapture data into the model (i.e., more tags are recaptured before they satisfy the mixing period) (Figure 18). As the tagging data are very influential on parameter estimation, removing tagging data from the model can have notable impacts on estimation. The distributions of the assumed mixing periods across the tag release groups under each of the D values are shown in Figure 19. Because there is some subjectivity in deciding the D value to apply, we included the mixing periods for the three D values as an axis in the structural uncertainty grid for this assessment, and for the diagnostic model apply the D value of 0.2, which results in a distribution of mixing periods that is closer to the fixed one quarter assumption from the 2019 assessment (i.e., most mixing periods are 0 or 1 quarter) (Figure 19).

While most of the recent PTTP release events were simulated (i.e., this from 2006-2017), not all tag release groups could be simulated due to the environmental forcing models not being updated for recent years or not available for earlier years at the required resolution. The JPTP has many release groups often with small numbers of fish, and it was not practical to simulate all of these; therefore a subset was simulated to provide a sample of mixing periods from 1998 to 2017. For those release groups not simulated we simply applied the median mixing period values for tag release groups in that region for each of the D values for JPTP, the historical SSAP/RTTP and very recent PTTP tagging events.

7.5 Likelihood components

There are three data components that contribute to the log-likelihood function for the skipjack stock assessment - the survey fishery CPUE data, the length-frequency data, and the tagging data.

7.5.1 Survey fishery CPUE likelihood

In previous catch-errors models, abundance indices were constructed for extraction fisheries by assuming that catchability remained constant over time. In catch-conditioned models, a new approach has been implemented to model directly the CPUE for ‘survey’ fisheries. While such survey fisheries exist in the model and in the data inputs as defined ‘fisheries’, they differ from the regular extraction fisheries in that no catch is extracted and their CPUE is modelled directly as a lognormal likelihood contribution. The likelihood weighting of the different survey fisheries is controlled by specified weights for each. Currently, grouped survey fisheries must be given the same weights. Time-series variation in the weighting of CPUE observations is not currently supported. The weights used for the survey fisheries in this assessment are based on estimates from the respective

CPUE standardisations (Tears et al., 2022) (Section 6.5) .

7.5.2 Length frequency: Dirichlet-multinomial likelihood

Previous MULTIFAN-CL assessments have generally used a robust lognormal likelihood for modelling length and weight frequency data. This method suffered from the requirement to subjectively specify the ratio of effective sample size (ESS) to observed sample size (OSS). In this assessment, in order to have a more objective weighting of size composition data, we have implemented a Dirichlet-multinomial (DM) likelihood for the length frequency data. Our implementation is similar to the DM implementation in Stock Synthesis (Thorson et al., 2017), which has been shown to be capable of estimating ESS for compositional data and performs similarly to iterative re-weighting methods. In the MULTIFAN-CL implementation, two categories of parameter are estimated - an exponent for an ESS multiplier and an exponent for a sample size covariate. The ESS multiplier parameter estimates the relationship between the OSS and ESS. With thousands of length-frequency samples in the model, it is obviously not feasible to estimate this parameter independently for each sample. Therefore, we estimate a sample size covariate parameter, which defines how ESS varies with OSS within a fishery, or group of fisheries. This keeps the parameter estimation tractable, with just two parameters estimated for each fishery, or group of fisheries. In this assessment, we have defined three fishery groups for the estimation of DM parameters according to gear type – large purse seine and pole-and-line, longline, and the small-fish miscellaneous gear fisheries in Indonesia, Philippines and Vietnam. The estimated relationships between OSS and ESS for these three fishery groups are shown in Figure 20. The average ESS for samples in the three groups are 19.6, 18.6 and 41.3, respectively. This represents average down-scaling from the OSS of 0.50%, 6.81% and 0.26%, respectively for the three fishery groups.

7.5.3 Tagging data

A log-likelihood component for the tag data was computed using a negative binomial distribution. The negative binomial is preferred over the more commonly used Poisson distribution because tagging data often exhibit more variability than can be attributed by the Poisson. We have employed a parameterization of the overdispersion parameter (τ) such that as it approaches 1, the negative binomial approaches the Poisson. Therefore, if the tag return data show high variability (for example, due to contagion or non-independence of tags), then the negative binomial is able to recognize this. This should then provide a more realistic weighting of the tag return data in the overall log-likelihood and allow the variability in tag returns to impact the confidence intervals of estimated parameters. Therefore, we allowed the overdispersion parameter (τ) to be estimated by the assessment model. A complete derivation and description of the negative binomial likelihood function for tagging data is provided in Kleiber et al. (2019).

7.6 Parameter estimation and uncertainty

The parameters of the model were estimated by maximizing the log-likelihood of all data components plus the log of the probability density functions of the priors and penalties specified in the model. The maximization to a point of model convergence was performed by an efficient optimization using exact derivatives with respect to the model parameters (auto-differentiation, [Fournier et al. \(2012\)](#)). Estimation was conducted in a series of phases, the first of which used relatively arbitrary starting values for most parameters. A bash shell script, “doitall”, implements the phased procedure for fitting the model. Some parameters were assigned specified starting values consistent with available biological information. The values of these parameters are provided in the skj.ini input file.

In this assessment, despite good convergence criteria (maximum parameter gradients for the grid models being of the order 1E-03 to 1E-05, one model was 1E-02), none of the Hessians for the grid models was positive definite, with between 1 and 20 negative eigenvalues across the 18 models. Under these circumstances, we were unsure of the reliability of the Hessian-based estimation uncertainty, which for the key stock status reference point variable $SB_{recent}/SB_{F=0}$ indicated CV's consistently around 0.02. Therefore we chose not to use these estimates of uncertainty for the individual models in the grid, but focussed on the factorial grid of model runs which incorporated important structural and biological uncertainties. This structural uncertainty grid attempts to describe the main sources of structural and parameter uncertainty in the assessment. Previous experience has shown that overall uncertainty is dominated by the structural uncertainty grid. We are continuing to explore simplified model structures to try to understand what aspects of the current model are preventing the attainment of positive definite Hessians.

For highly complex population models fitted to large amounts of often conflicting data, it is common for there to be difficulties in estimating absolute abundance. Therefore, a likelihood profile analysis was undertaken of the marginal posterior likelihood in respect of population scaling, following the procedure outlined by [McKechnie et al. \(2017\)](#) and [Tremblay-Boyer et al. \(2017\)](#). This results of this procedure are presented in the the appendices (Appendices [Section 15.1](#)).

Retrospective analyses were conducted as a general test of the stability of the model, as a robust model should produce similar output when rerun with data for the terminal year/s sequentially excluded ([Cadigan and Farrell, 2005](#)). The retrospective analyses for the 2022 diagnostic model are presented in the appendices (Appendices [Section 15.4](#)).

7.7 Stock assessment interpretation methods

7.7.1 Reference points

The unfished spawning potential ($SB_{F=0}$) in each time period was calculated given the estimated recruitments and the Beverton-Holt SRR ([Section 7.2.1](#)). This offers a basis for comparing the exploited population relative to the population subject to natural mortality only. The WCPFC

adopted 20% $SB_{F=0}$ as a limit reference point (LRP) for the skipjack stock where $SB_{F=0}$ for this assessment is calculated as the average over the period 2011–2020. The interim target reference point (iTRP) for this stock as indicated in CMM 2021-01 is 50% $SB_{F=0}$ (time period as defined for the LRP). Stock status was referenced against these points by calculating the reference points; $SB_{recent}/SB_{F=0}$ and $SB_{latest}/SB_{F=0}$ where $SB_{F=0}$ is calculated over 2011–2020 and SB_{recent} and SB_{latest} are the mean of the estimated spawning potential over 2018–2021, and 2021 respectively (Table 4).

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} ; 2017–2020 for this stock assessment) divided by the fishing mortality producing MSY which is a product of the yield analysis and was detailed in Section 7.7.2.

For this assessment we also add a reference point for the ratio of the values of recent spawning depletion $SB_{recent}/SB_{F=0}$ to that for year 2012 $SB_{2012}/SB_{F=0}$. This is added in response to a request from SC17.

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

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

Fishing mortality-at-age (F_a) for the yield analysis was determined as the mean over a recent period of time (2017–2020). We do not include 2021 in the average as fishing mortality tends to have high uncertainty for the terminal data year of the analysis and the catch and effort data for this terminal year are potentially incomplete. Additionally, recruitments for the last two quarters of the terminal year of the model are constrained to be the geometric mean across the entire time series, which affects the F for the youngest age classes.

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

2020). 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.

7.7.3 Depletion and fishery impact

Fishery depletion was calculated by computing the unexploited spawning potential time series (at the region level) using the estimated model parameters, but assuming that fishing mortality was zero. Both the estimated spawning potential SB_t (with fishing) and the unexploited spawning potential $SB_{F=0[t]}$ incorporate recruitment variability. Therefore, the ratio of these two quantities at each quarterly time step (t) of the analysis $SB_t/SB_{F=0[t]}$ can be interpreted as an index of fishery depletion. The computation of unexploited spawning potential includes an adjustment in recruitment to acknowledge the possibility of reduction of recruitment in exploited populations through stock-recruitment effects. To achieve this, the estimated recruitment deviations are multiplied by a scalar based on the difference in the equilibrium recruitment between the fished and unfished spawning potential estimates.

A similar approach was used to estimate depletion associated with specific fisheries or groups of fisheries. Here, fishery groups of interest (purse seine unassociated sets, pure seine associated sets, purse seine unidentified, pole-and-line, longline, and miscellaneous fisheries), are removed in-turn in separate simulations. The changes in depletion observed in these runs are then indicative of the depletion (fishing impact) caused by each of the removed fisheries.

7.7.4 Kobe analysis and Majuro plots

For the standard yield analysis (Section 7.7.2), the fishing mortality-at-age, F_a , is determined as the average over some recent period of time (2017–2020). In addition to this approach the MSY-based reference points (F_t/F_{MSY}), and SB_t/SB_{MSY}) and the depletion-based reference point ($SB_t/SB_{F=0[t]}$) were also computed by repeating the yield analysis for each year in turn. This enabled temporal trends in the reference points to be estimated and a consideration of the differences in MSY levels under historical patterns of age-specific exploitation. This analysis is presented in the form of dynamic Kobe plots and “Majuro plots”, which have been presented for all stock assessments in recent years.

7.7.5 Stock projections from the structural uncertainty grid

Projections of stock assessment models can be conducted within MFCL to ensure consistency between the fitted model and the simulated future dynamics, and the framework for performing this exercise is detailed in Pilling et al. (2016). Typically, stochastic 30 year projections of recent catch and effort (2019-2021) are conducted from each assessment model within the uncertainty grid developed. For each model, 100 stochastic projections, which incorporate future recruitments randomly sampled from historical deviates, are performed. At the time of writing the functions in of MFCL for running stochastic projections for catch conditioned models were still being completed.

Deterministic projections are provided in this report (Section 15.5), stochastic projections will be conducted as soon as this capability is functioning and added to a revision.

8 Model runs

8.1 Developments from the last assessment

The progression of model development (referred to as the ‘stepwise’) from the 2019 assessment reference case model to the 2022 diagnostic model is described in the following section. Initial steps in the model development process involved using data inputs from the 2019 assessment and implementing the main changes to the modelling methods used for the 2022 assessment. Following these steps, the final step for the development of the 2022 diagnostic case involved updating the data/inputs until 2021, the final year of this assessment (discussed in more detail below). While this final step involved a ‘package’ of updates, these were all essentially determined through considerable planing discussions and preparatory work, and as such they were applied together. Ultimately this ‘packaged’ step was more efficient, and was not overly influential compared to the previous steps.

Stepwise model development

1. The 2019 diagnostic model. [*Diag2019*]
2. The 2019 diagnostic model implemented with the new MFCL 2.0.8.5 executable. [*S1NewExe*]
3. Conversion to the catch-conditioned model, no survey fisheries. [*S2CatchCond*]
4. Catch-conditioned with addition of survey fisheries and related likelihood components. Survey fisheries for pole-and-line regions 1, 2, 3, 4, 7 and 8 ungrouped, Region 5 and Region 6 estimated as per 2019. [*S3Survey CPUE and likelihood-ungrouped*]
5. Catch-conditioned with pole-and-line survey indices grouped, Region 5 and Region 6 survey indices as per 2019 [*S4Survey CPUE and likelihood-grouped*]
6. Apply the Dirichlet-multinomial likelihood for estimating effective sample size of length composition data, and estimate growth internal to the model (rather than input as external parameters). [*S5Dirichlet and Intgrowth*]
7. 2021 inputs: update catch, length composition (re-weighted), tagging data (with releases to 2019 included and updated tagger effects corrections), updated tag reporting rate priors, and updated survey fishery CPUE. Updating the survey CPUE involved: 1. re-estimating the survey fishery standardised CPUE for the Japan pole-and-line fisheries (regions 1, 2, 3, 4, 7, 8) with the updated data (to 2020, 2021 data was incomplete due to COVID implications) using the same spatiotemporal VAST model as applied in the previous assessment (Teears et al., 2022). Adding the new spatiotemporal VAST ‘unassociated’ purse seine survey indices, which replaced the previous combined set types GLM index for region 6 (Vidal et al., 2019) and provided new purse seine survey indices for regions 7 and 8. Updating the Region 5 PH

purse seine index, using the same GLM standardisation as for the 2019 assessment (Tears et al., 2022). Finally, as recommended during the 2022 PAW - due to the spatial contraction of Japanese pole-and-line fisheries, particularly from Region 8, the survey time series for this region was truncated by removing years with less than 20% sample coverage (Tears et al., 2022). As per previous step, the survey fisheries were grouped for the pole-and-line in regions 1, 2, 3, 4, 78, and for the purse seine survey fisheries in regions 6, 7, 8. Finally, the fixed tagged mixing assumption of one quarter used in the 2019 diagnostic case was replaced with the variable mixing periods estimated from the simulation studies based on the D statistic of 0.2. At D 0.2 the majority of release groups were considered mixed at zero or one quarter after release (Figure 19), similar to the assumed fixed 1 quarter mixing period applied to the 2019 diagnostic case. As per the 2019 diagnostic model, steepness was fixed at 0.8 [*S62021inputs-2022diag*]

8.2 Sensitivity analyses and structural uncertainty

Several uncertainties were recommended for inclusion in a structural uncertainty model grid at the 2022 Pre-assessment Workshop (Hamer, 2022) based on their influence in the 2019 assessment (i.e., tag mixing period, growth), or as standard practice in tuna assessments (i.e., steepness). Some others were suggested for exploration as additional one-off sensitivities. To inform the axes of the structural uncertainty grid the following one-off sensitivity analyses were therefore conducted:

1. Variable tag mixing from tag mixing simulations for KS D = 0.1 (i.e., results in longer tag mixing periods and allows less tagging data into the model)
2. Variable tag mixing from tag mixing simulations for KS D = 0.3 (i.e., results in shorter tag mixing periods and allows more tagging data into the model)
3. Growth - applying the externally estimated growth curve from otolith daily aging and tag-recapture growth increments.
4. Steepness 0.65
5. Steepness 0.95

The 2022 Pre-Assessment Workshop also recommended exploring three other sensitivities with the diagnostic model:

6. **Effort creep:** There was considerable discussion at the 2022 Pre-assessment Workshop around the significance of effort/efficiency creep in skipjack tuna fisheries in the WCPO. However, it was acknowledged that there is poor quantitative understanding of rates of effort/efficiency creep and their variation over time for any of the skipjack fisheries, and that more work is required to develop informed scenarios to apply in assessment models used for management advice. We conducted an additional sensitivity analysis that assumed a 1% year-on-year increase in the catchability (efficiency) of each of the survey fisheries. For ex-

ample, by applying this to the pole-and-line survey fisheries from 1972 to 2020, we assume catchability of this fishery has increased by 48% over this period. The 1% year-on-year effort creep scenario was modelled simply by multiplying the standardised index values (by quarter) for each sequential year from the start of the time series by a scalar starting at 1 and reducing by 0.01 each year.

7. **Movement:** For spatially structured assessments internal model estimation of how abundance shifts among model regions is very important. The diagnostic model attempts to estimate movement coefficients internally to fit the patterns in the data and has substantial freedom to do this. However, even with tag-recapture data in the model, the distributions of tag releases and recaptures are conditioned on the locations of tagging cruises and the spatio-temporal patterns of fishing. Movement among model regions is still considered a significant uncertainty and may not be well informed by the data, particularly for regions where less tag release effort has occurred. As an alternative to the internally estimated movement, a sensitivity in which we used movement estimates based on the skipjack SEAPODYM (Senina et al., 2020a) model was specified. SEAPODYM is highly spatially resolved and provides predictions on spatio-temporal exchange of fish by age class, forced by environmental/habitat variables. For the SEAPODYM sensitivity, we computed movement probability matrices based on estimates from the SEAPODYM model. The probability matrices were computed by year quarter and quarterly age class, and averaged over years to match the MULTIFAN-CL movement structure. The matrices were specified in MULTIFAN-CL as fixed parameters.
8. **Tagger effects corrections:** The companion paper by Peatman et al. (2022) describes the estimation of correction factors for tagger effects for the 2022 skipjack assessment. These estimated correction factors have an associated uncertainty. Discussion at the 2022 PAW recommended that this uncertainty in tagger effects adjustment be explored (Hamer, 2022), therefore we have included one-off sensitivity models that apply the 10th and 90th percentile values from the distributions of the estimated tagger effects corrections.

Each one-off sensitivity model was created by making a single change to the 2022 diagnostic model.

5–region model: A simpler five region model was also developed based on a recommendation from the PAW (Hamer, 2022). The 5-region model was based on merging of regions 1-4 of the 8-region model, a similar extraction fisheries structure (some fisheries had to be merged), and recalculating the pole-and-line survey index and size composition re-weightings for the single northern region. The 5-region model analysis is included as an appendix to inform continued discussion on the regional structure of the WCPO skipjack assessment.

8.3 Structural uncertainty

Stock assessments of pelagic species in the WCPO use an approach to assess the structural uncertainty in the assessment model by running a “grid” of models that explore the interactions among

selected “axes” of uncertainty. The grid contains all combinations of levels of several model quantities, or assumptions, and allows the sensitivity of stock status and management quantities to this uncertainty to be determined and factored into management advice. The axes are generally selected from factors explored in the one-off sensitivities with the aim of providing an approximate understanding of variability in model estimates due to assumptions in model structure not accounted for by statistical uncertainty estimated in a single model run, or over a set of one-off sensitivities.

The structural uncertainty grid for the 2022 skipjack stock assessment was constructed from 3 axes of uncertainty with 2–3 levels for each, resulting in a total of 18 models (Table 2). The reduction in number of grid model from the 2019 assessment (54 models) was due to the use of the Dirichlet-multinomial likelihood which removed the need for the axis of arbitrary size composition weighting divisors.

The values for the diagnostic model are in bold and the levels used in the grid are directly comparable to those presented in Section 8.2 through identical notation. The levels of the grid are:

1. Steepness [*0.65*, **0.8**, *0.95*]
2. Growth [Model estimated (**GrowthModel**), and External estimated (*GrowthExternal*)]
3. Assumed tag mixing period [T1, (*Mix D 0.1*), T2 (**Mix D 0.2**) and T3 (*Mix D 0.3*)]

9 Results

9.1 Consequences of key model developments

The progression of model development from the 2019 to the 2022 diagnostic model is described in Section 8.1 and the results are displayed in Figure 21. A summary of the consequences of this progression through the model focussing on the key management quantities of spawning potential SB and dynamic spawning potential depletion $SB/SB_{F=0}$ is as follows:

1. The reference case model for the 2019 assessment refit with the latest version of MFCL (M1; Figure 21) produced a nearly identical result to the 2019 MFCL version (2019 diagnostic, M0, is obscured by the M1 line) for both SB and $SB/SB_{F=0}$.
2. The conversion to the catch conditioned model (M2) had a minor and variable effect on $SB/SB_{F=0}$ over the time series and the terminal year was nearly identical to the previous step, however, SB was higher throughout the time series (M2; Figure 21).
3. The introduction of the survey fishery indices (ungrouped) and the associated likelihood term resulted in the first notable difference in the stepwise development, with a large increase in SB to almost double that of the 2019 diagnostic case in the terminal year, and a considerable increase in $SB/SB_{F=0}$ (less depleted) compared to the previous step, the difference was also greater for the more recent years of the model time period (M3; Figure 21).

4. The grouping of the pole-and-line survey fisheries led to a further increase in $SB/SB_{F=0}$ (less depleted) but a decrease in SB (M4; [Figure 21](#)).
5. The next and final step using the 2019 inputs was to implement the Dirichlet-multinomial likelihood and estimate the growth within the model. This resulted in a notable reduction of both the estimated $SB/SB_{F=0}$ (more depleted) and the SB from the previous step. Overall, implementing the new modelling approaches for the 2019 inputs resulted in estimation of a less depleted stock and higher spawning potential in the terminal years than the 2019 diagnostic case (M7 [Figure 21](#)).
6. The next step in the model development (Mdc2022) was to update the data inputs and the survey indices, the assumptions for tagger effects and tag reporting rates based on [Peatman \(2022\)](#); [Peatman et al. \(2022\)](#), and implement the variable mixing periods from the tag mixing simulations for KS D value of 0.2. This 'packaged' step created the 2022 diagnostic model, and had a moderate effect on $SB/SB_{F=0}$ from the last step involving the 2019 inputs, estimating a less depleted stock status across the recent two decades (M7 versus Mdc2022, [Figure 21](#)). The SB in recent years was estimated at a similar level to the final step for the 2019 inputs, but showed some differences in the variability over time, with lower levels estimated from the late 1990s back to the start of the model period (M1 and M7 versus Mdc2022; [Figure 21](#)).

Overall, the 2022 diagnostic model estimated a less depleted stock status than the 2019 diagnostic, with $SB/SB_{F=0}$ of approximately 0.5 for the 2022 diagnostic versus approximately 0.4 for the 2019 diagnostic for the overlapping recent years of the model time period. There was also a clearer downward trend of increasing depletion in recent years for the 2022 diagnostic. The 2022 diagnostic model estimated higher spawning potential from the late 1990s onwards than the 2019 diagnostic, but similar or lower spawning potential prior to the that (M1 versus Mdc2022, [Figure 21](#)). Both diagnostic models indicated a declining trend in spawning potential since around 2006.

The most influential steps in the development of the 2022 diagnostic model were the introduction of survey fishery CPUE and associated likelihood term which led to a less depleted stock and higher spawning potential, followed by the application of the Dirichlet-multinomial likelihood with internal growth estimation, that had the effect of increasing the depletion and reducing the spawning potential slightly ([Figure 21](#)).

9.2 Fit of the diagnostic model to data sources

9.2.1 Standardized CPUE: survey fisheries

Pole-and-line CPUE regions 1, 2, 3, 4, 7 and 8: There was substantial seasonal variability in the CPUE indices for the pole-and-line survey fisheries, but a stable trend over time. Overall the model-estimated CPUE predicted the seasonal variation in the survey indices relatively well and also the stable long-term trends, with exception of Region 7 where the model-estimated CPUE did poorly at fitting the decline in the pole-and-line survey CPUE from 2010 onwards ([Figure 22](#)). While

the model-estimated CPUE largely captured the seasonal variability observed, for the temperate and equatorial pole-and-line fisheries (i.e., P-ALL-3, P-ALL-4, P-ALL-7, and P-ALL-8) the fit to the variability in the survey fishery CPUE during the 1990s was relatively poor for the highest observations but better for the lower observations (Figure 22).

Purse seine CPUE Region 5: The model-estimated CPUE showed a poor fit to the higher variability in the PH purse seine survey CPUE in the early part of the time series from 2005-2010, but after 2010 the model predicted the survey CPUE very well, including the observed decreasing trend in CPUE (Figure 22).

“Unassociated” purse seine CPUE regions 6, 7, 8: For the unassociated purse seine CPUE the model underestimated the survey CPUE in Region 6 from 2010-2014, but estimated it well for the remaining years. For Region 7, the model CPUE fit the observed CPUE quite well, both the short-term variation and the long-term decreasing trend. For Region 8 the model tended to overestimate the observed CPUE until the end of the model period, however it captured the dynamics of the observed CPUE reasonably well (Figure 22).

9.2.2 Size composition data

ID, VN, PH fisheries: The model estimations of the composite length composition of the domestic (DOM) fisheries in ID, VN and PH were generally very good. There was some over estimation of smaller fish for Dom-ID-5 (Fishery 11), and an unusual bimodal feature to the estimated composition with a dip that corresponded to the mode of the observed composition (Figure 23). Across the model time series the model estimates of median lengths were very consistent with the temporal variation in the observed data (Figure 28).

Pole-and-line fisheries: The model estimation of the length composition for the pole-and-line fisheries were generally very good, although for Region 4 there was some over estimation of smaller fish (Figure 24). Across the time series the model estimates of median lengths were generally consistent with observations and predicted the temporal variation in the observed data, although for region 3 the model tended to overestimate the lengths in the earlier/middle part of the time series, and underestimate it in the recent years. However, Region 3 has relatively low observed data for much of the time series (Figure 13). Model estimates for regions 5, 7 and 8 fit through the observed data well, despite the observed data being sporadic and with low sample numbers in many year/quarters (Teears et al., 2022) (Figure 13, Figure 29).

Pure seine fisheries regions 1, 2, 3: The model estimation of the length composition for the purse seine fisheries for regions 1, 2, and 3 were generally very good, although for regions 1 and 2 there was some underestimation of the proportions in the modal lengths of the observed data (Figure 25). Across the time series the model estimates of median lengths were similar to and consistent with the temporal variation in the observed data, noting the low coverage of the observed data across the time period (Teears et al., 2022) (Figure 13, Figure 30).

Pure seine fisheries ‘associated’ and ‘unassociated’ regions 5, 6, 7, 8: The model estimation of the length composition for the associated and unassociated purse seine fisheries for region 5, 6, 7, and 8 were very good for all regions, albeit with some slight over estimation of the proportions of smaller fish for the unassociated component in regions 7 and 8 (Figure 26). Across the time series the model estimates of median lengths were consistent with the observed data, but noting the very stable size composition of the catches from these fisheries over time (Teears et al., 2022) (Figure 31).

Longline fisheries: The model estimations of the length compositions for the longline fisheries, noting the relatively low sample sizes, were reasonable for all regions, with some slight over estimation of proportions of smaller fish in regions 7 and 8 (Figure 27). Across the time series the model estimates of median lengths were consistent with the observed data, noting the limited observations in regions 1, 2, 3, and 5, and some larger sample lengths for Region 5 at the end of the time series that are not well estimated by the model (Figure 32).

9.2.3 Tagging data

On the aggregated scale, the model tag attrition estimates fit the observed tagging data very well (Figure 33), but on the tag program scale there are some differences in the quality of fit (Figure 34). The model generally predicts the number of tags returned by the JPTP to be larger than the observed values, whereas the PTTP and RTTP are fit very well. The model under predicts the tag returns for the SSAP for times at liberty of 3–5 quarters. At the regional scale the model predicts the number of tags returns very well, especially for the recaptures after one quarter at liberty (Figure 35).

For the aggregated data the model predicted tag returns over time shows relatively good agreement with observed data, though some of the years with high recaptures are underestimated by the model (Figure 36). The predicted tag returns by the ID, VN and PH fisheries fit the observed tag returns reasonable well, but likewise underestimated the quarters with higher tag returns (Figure 37). For the pole-and-line fisheries, there is considerable range among regions in the numbers of observed returns, and the model does well to predict the relative rates of observed returns across the range, but still underestimated the higher tag returns (Figure 38). Similar to the other fisheries, the dynamics and variation in the model predicted tag returns for the purse seine fisheries over time were reasonably consistent with the observed data, with the main difference again being for some of the quarters with higher observed returns (Figure 39).

9.3 Model parameter estimates (diagnostic model)

9.3.1 Selectivity

A range of selectivity patterns are shown by the different fisheries in the model and can be largely classified by gear type. The age-specific selectivity coefficients are displayed in Figure 40 and

length-specific selectivity curves are displayed in [Figure 41](#). The pole-and-line fisheries (including the survey fisheries) select mostly younger fish that are three to seven quarters old and 40-70 cm FL, whereas the longline fisheries catch the largest and oldest fish, with their selectivity asymptoting at around 70 cm FL and around 8 quarters, with very low selectivity below 60 cm. The purse seine associated and unassociated fisheries (and the related unassociated index fishery) have slightly different selectivities, with the associated fishery mostly selecting smaller fish in the length range 40–60 cm and ages two to seven quarters, compared to the unassociated fishery that selects fish in the length range 50–70 cm and ages three to eight quarters. The purse seine fisheries in regions 1, 2, and 3 are largely unassociated fisheries, and similar to the equatorial unassociated fisheries, select fish in the length range 50–70 cm and ages three to eight quarters. Fisheries in Region 5, mostly select smaller fish around 30–50 cm, with the exception of the PH domestic fishery (Z-PH-5) that catches some larger fish.

9.3.2 Movement

Observed patterns of tag releases and returns among regions are compared to the diagnostic model estimated movement coefficients among regions for each quarter in [Figure 42](#). The tag return data show generally low movement between the northern regions 1-4 and the equatorial regions 5-8, but some movement of tagged fish is observed from regions 3 and 4 to regions 5 and 7, particularly in quarters three and four. Model estimated movement coefficients similarly predict that most movement occurs during quarters three and four, and similar patterns of low movement between the northern regions 1-4 and the equatorial regions 5-8 are predicted. However, movement from model regions 3 and 4 to equatorial model regions 6, 7, and 8 is notable and occurs mostly in quarter four.

9.3.3 Natural mortality

The estimated M-at-age for the diagnostic model, and the other grid models that vary in their combination of growth curve and tag mixing estimation are shown in [Figure 43](#), noting that varying the steepness has no influence on the M-at-age estimation. There is a notable difference in the M-at-age curves that is driven by the choice of growth curve. The M-at-age curves for the externally estimated growth (G2) have a U-shaped form with higher M at younger and older ages, with the exception of the curve for the model with the tag mixing D value of 0.1, where M-at-age flattens after 12 quarters. The M-at-age curves for models that use the internally estimated growth all show a common shape with higher M at young ages, lowest M at ages between four to eight quarters, increasing M for ages 6 to 11 quarters, and then decreasing M for ages beyond 11 quarters. The different tag mixing assumptions also have an effect, mostly for the T1 assumption (longer mixing period and less tag data used) which increases the M estimates, most notable for ages 9–12 quarters. It is worth noting the skipjack population and catches are dominated by fish of less than eight quarters age ([Figure 55](#)). The estimates of M-at-age, although variable depending on the model, are not well informed for ages greater than eight quarters, and have little influence on

model estimates of biomass as so few fish occur in these older age classes.

9.3.4 Maturity-at-Age

Maturity-at-age calculated by the 2022 diagnostic model (model T2G10.8 in [Figure 44](#)) is similar to the 2019 assessment due to the almost identical mean growth curves and the use of same length at maturity relationship. Comparison between maturity-at-age determined for the two growth models used in the uncertainty grid shows they both predict 50% maturity at just less than four quarters age. The diagnostic model predicts slightly lower proportions of mature fish below about four quarters and higher proportions of mature fish above four quarters compared to the model based on the external growth estimation.

9.3.5 Tag reporting rates

The estimated tag reporting rates by fishery recapture group (see groupings in [Table 1](#)) are displayed in [Figure 45](#). Unlike the 2019 assessment we have not estimated reporting rates where zero tag returns have been recorded. This removed 10 fishery/tag program groups from the reporting rate estimates, and there are no reporting rates at zero in [Figure 45](#). Further, as some tag reporting rates were estimated on the upper bounds of 0.9 we re-ran the diagnostic model with the upper bound set at 0.99, that resulted in four of the 10 reporting rates that were estimated at the 0.9 bound being estimated at between 0.9 and <0.99 and these are denoted in [Figure 45](#).

The most important groups for scaling the population size are the purse seine fisheries in regions 5–8 that account for the majority of recaptures from the SPC tagging programs (i.e., SSAP, RTTP, PTTP). The pole-and-line and purse seine fisheries in regions 1, 2, 3, and 4 recapture most of the releases from the JPTP (Japanese Tagging Programme), which releases fewer tags in the other regions. The reporting rates for the equatorial purse seine fisheries in region 6, 7, and 8, and reported under the more recent SPC tagging programmes (RTTP, PTTP), were all below the 0.9 bound except for the earlier RTTP S-8 that was on the 0.9 upper bound but was below the 0.99 upper bound. All except two of the JPTP reporting rates were below the upper bound, although three were close to the lower bounds, but these were for fisheries in regions 5 and 6 where the JPTP does not release tags or actively promote tag returns.

9.3.6 Growth

In the 2019 assessment a range of scalars were used to adjust the weighting of the length composition data to derive alternative growth curves that were then applied as external growth parameters to fit the actual stock assessment ([Vincent et al., 2019b](#)). Using the Dirichlet-multinomial likelihood approach the model was able to effectively estimate a growth curve similar to that used in the last assessment, but with a more realistic increasing variance of length at age with age. Hence this year variability in length-at-age for the older age classes is greater than it was in 2019 where the standard deviation (SD) was fixed across ages. This estimated diagnostic growth curve started

from a mean length of 23 cm at age-class one quarter. Mean length-at-age increased quickly until about age-class 8 quarters, after which growth slowed down until reaching a length of 84 cm (L_{inf}) in the oldest fish in the model (age-class 16) (Figure 15).

As previously noted, an additional growth curve was estimated as an alternative growth hypothesis using empirical data from daily otolith aging and tag-recapture growth increments (Macdonald et al., 2022). This empirical growth curve estimated a larger size at one quarter (31 cm v 23 cm) and slower growth rate with a lower mean size at 16 quarters of 73 cm (Figure 15).

9.4 Stock assessment results

9.4.1 Recruitment: diagnostic model

The estimated recruitment aggregated across all regions (Figure 46) shows high interannual variation, however, the trend in recruitment is relatively stationary over the first decade of the assessment period. From the early 1980s the trend in recruitment shows a gradual increase until around 2005, after which it stabilizes until 2015. From 2015, while still highly variable at short time scales, the estimated recruitment shows a downward trend until the end of the time series (Figure 46). As in the 2019 assessment, the final two estimates of recruitment at the aggregated WCPO scale are constrained to equal the geometric mean of recruitments over the entire assessment period (Rice et al., 2014). This has little impact on the spawning potential or other reference points as recruits from these most recent quarters have not entered the fishery. At the regional scale highest recruitments over the past two decades were estimated to have occurred in regions 5, 7, and 8, followed by region 6, although historically, relatively high recruitments were also estimated in the other regions (Figure 47, Figure 48). Recruitment in Region 8 increased around 2015, which is likely related to the observed large catch in this region around that time (Figure 48). In general, the overall trends in spawning potential and total biomass are consistent with the trends in recruitment (Figure 48).

The estimated stock recruitment relationship for the diagnostic model is presented in Figure 49, which shows the general increasing levels of recruitment over time.

9.4.2 Biomass: diagnostic model

The 2022 diagnostic model predicted that total biomass and spawning potential aggregated across the model regions increased from early 1970s to the mid-2000s, after which it has steadily declined (Figure 48, Figure 50). For the northern model regions 1–4, despite some variation among regions early in the time series, the trends in spawning potential have been relatively stable since around 1990. For the equatorial regions 5–8, Region 6 has shown a stable trend in spawning potential from 1980, but regions 5, 7, and 8 show slight increasing trends from the 1980s until the mid-2000's, after which the spawning potential declines (Figure 50).

9.4.3 Depletion: diagnostic model

The 2022 diagnostic model predicted that spawning potential depletion (here after referred to as ‘spawning depletion’) aggregated across the model regions has declined steadily since the beginning of the time series (Figure 51). For the northern model regions the spawning depletion has been very stable since the 1980s, becoming slightly less depleted in the past decade for regions 1 and 2. The equatorial regions 5–8, all show long-term decreasing trends, becoming more depleted overtime. Compared to regions 5, 7 and 8, Region 6 shows a more stable depletion trajectory since 2010 (Figure 51).

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

In order to better interpret the trends in spawning depletion it is useful to compare the individual trends in spawning potential (SB_t) with the spawning potential that would be predicted in the absence of fishing ($SB_{F=0}$) (Figure 52). At the scale of the entire model region, this comparison indicates that the long-term decreasing trend in the spawning depletion ($SB/SB_{F=0}$) value is largely driven by the long-term increasing trend in the prediction of the unfished spawning potential, and the stability of the fished spawning potential across time. Further, since 2005 the unfished spawning potential has been stationary, but the fished spawning potential has begun to decline. These patterns are largely driven by the equatorial regions 5-8. In contrast, for the northern regions 1–4, the variation in fished and unfished spawning potential tend to track each other closely (Figure 52). These comparisons, and the predicted increasing recruitment over time for the equatorial regions (Figure 47), indicate that the model is estimating that the stock has become more productive overtime, at least in the equatorial regions. Given that the information from the various longer term input data, most notably the pole-and-line CPUE indices and length composition series are indicative of a stable stock biomass over time, the model is supporting the observed large increases in catches over time by increasing the recruitment. The plausibility of an increased recruitment trend in the face of increased fishing pressure and other environmental factors, plus the assumption that information on stock trends from fishery dependent CPUE are reliable, are discussed further in the Section 10.

9.4.5 Fishing mortality: diagnostic model

Average fishing mortality rates for juvenile and adult age-classes have increase continually and at a similar rate throughout the time series (Figure 53).

The temporal trends in fishing-mortality-at-age vary by region, but the overall fishing-mortality-at-age increases over time (Figure 54). Fishing mortality in regions 1 and 3 is estimated to be very high seasonally during the first decade of the assessment, but then decreases and remains at the same level until the end of the time series. The fishing mortality in regions 1–4 are highly seasonal which is consistent with the CPUE for these regions (Figure 22). The fishing mortality for the equatorial regions 5–8 generally increases throughout the time series, particularly in the last two

decades of the assessment. Fishing mortality on the youngest age-class is only significant in Region 5 where the miscellaneous fisheries catch very small skipjack.

Changes in fishing mortality-at-age and the population age structure are shown for decadal time intervals in [Figure 55](#). Since the 1980s, the increase of fishing mortality to the current levels is due to the increases in catches of both juvenile and adult fish by the equatorial purse seine fisheries and the mixed gear fisheries of ID, PH, and VN in Region 5. Fishing mortality on ages 3–6 quarters also increased through time consistent with the increased fishing mortality from the purse seine fishery.

9.5 Multi-model inference: sensitivity analyses and structural uncertainty

9.5.1 One-off sensitivity analyses

Comparisons of the spawning depletion and spawning potential trajectories for the diagnostic model and the related one-off sensitivity models are provided in [Figure 56](#). The steepness sensitivities are not included in [Figure 56](#) to avoid crowding of the figure as they had very little influence on spawning potential and a minor influence on spawning depletion (i.e., steepness 0.95 slightly less depleted, 0.65 slightly more depleted) compared to the other sensitivities. These comparisons show that estimates of both spawning depletion and spawning potential, were most sensitive to the choice of D values for estimating the tag mixing periods (i.e., Mixing T1 = D value 0.1, Mixing T3 = D value 0.3), and the growth model used. Under the more conservative tag mixing scenario (Mixing T1), which excludes more tagging data, the spawning depletion and spawning potential are both higher than the diagnostic case, and opposite for the Mixing T3 scenario. For the external fixed growth option, spawning potential is estimated to be higher than the diagnostic case, similarly for spawning depletion, but more so since 2000 ([Figure 56](#)).

The other sensitivities had relatively minor effects compared to those discussed above. For spawning depletion, the effort creep scenario was slightly more depleted than the diagnostic model, the SEAPODYM sensitivity was almost the same as the diagnostic case, as were the higher and lower tagger effects correction models. For spawning potential, all four of the previously discussed sensitivities estimated slightly lower spawning potential than the diagnostic model ([Figure 56](#)).

Due the minor influence of these sensitivities we chose to maintain the original proposed structural uncertainty grid, with axis for tag mixing, growth and steepness. While steepness may not be that influential on spawning potential or spawning depletion reference points it can be quite influential on MSY based reference points and as such it is standard practice to include this uncertainty axis in tuna assessments.

9.5.2 Structural uncertainty grid

Results of the structural uncertainty analysis are summarized in box and violin plots of F_{recent}/F_{MSY} and $SB_{recent}/SB_{F=0}$ for the different levels of each of the three axes of uncertainty ([Figure 57](#)).

The distribution of recruitment across model regions and quarters for all models in the structural uncertainty grid is summarised in [Figure 58](#). Time series of spawning depletion ($SB_{recent}/SB_{F=0}$) and spawning potential SB_t across grid models are shown in [Figure 59](#) and [Figure 60](#). Majuro and Kobe plots showing the estimates of F_{recent}/F_{MSY} , $SB_{recent}/SB_{F=0}$, SB/SB_{MSY} , along with $SB_{latest}/SB_{F=0}$ and SB_{latest}/SB_{MSY} across all models in the grid ([Figure 61](#), [Figure 62](#)). The averages and quantiles across the 18 models in the grid for all of the reference points and other quantities of interest ([Table 4](#)).

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

- The grid contains 18 models that display a moderate range of estimates of stock status and reference points, and suggest that, overall, the stock is moderately less depleted and with higher spawning potential than the estimates from the 2019 assessment ([Table 4](#)). The grid has two outlying models that influence the range of uncertainty towards higher values of spawning potential ([Figure 60](#)) and spawning depletion ([Figure 59](#)). These two models involve the combinations of fixed externally estimated growth and the more conservative tag mixing assumptions. The primary influence being the conservative tag mixing assumption that predicts lower fishing mortality, and therefore higher spawning potential and a less depleted stock.
- The most influential axis was the tag mixing period, consistent with the 2019 assessment. We suggest that the approach used to allocate variable tag mixing periods by release groups for this assessment is a substantial improvement on the previous fixed mixing period assumption. Similar to the 2019 structural uncertainty analysis, the scenarios that in general assumed longer mixing periods, and therefore exclude more tag-recapture data from the model estimation, estimate a less depleted stock and lower levels of fishing mortality ([Figure 57](#)).
- The second most influential axis of uncertainty in the grid was the growth. The F_{recent}/F_{MSY} was lower, when the externally estimated growth was assumed and the $SB_{recent}/SB_{F=0}$ was higher.
- As expected, steepness had a more notable influence on F_{recent}/F_{MSY} than $SB_{recent}/SB_{F=0}$, with lower estimates of F_{recent}/F_{MSY} and slightly higher estimates of $SB_{recent}/SB_{F=0}$ for the higher steepness ([Figure 57](#)).
- The reference points calculated from the uncertainty grid suggest that median $SB_{recent}/SB_{F=0}$ is just above the interim target reference point of 50% $SB_{F=0}$, and the $SB_{latest}/SB_{F=0}$ is just below the interim target reference point ([Table 4](#), [Figure 61](#)).
- Overfishing is not estimated to be occurring for any of the models in the structural uncertainty grids, nor is the stock estimated to be in an overfished state ([Figure 61](#)).
- Spawning depletion in the region 5 and in the equatorial Region 7 is approaching the limit reference point 20% $SB_{F=0}$ in the recent model years for some of the grid models ([Figure 59](#)).

However, as discussed, this appears mostly due to the estimation of increased levels of unfished spawning potential $SB_{F=0}$, rather than notable decreases in the estimates of spawning potential (SB). Declining trends in spawning potential are however estimated for regions 5, 7 and 8 over the last five years (Figure 60).

- The northern model regions 1–4, show very stable spawning depletion Figure 59 and spawning potential Figure 60 across the model time period.
- Recruitment is predicted to be highest in Region 5 and generally higher in regions 5–8 than regions 1–4, with recruitment occurring in all quarters for the equatorial regions. Recruitment is also suggested to be higher in region 2 in quarters 1 and 4, perhaps related to the strong seasonality of availability in this region (Figure 58).

9.5.3 Analyses of stock status

There are several ancillary analyses related to stock status that are typically undertaken on the diagnostic model (e.g., dynamic Majuro and Kobe analyses, fisheries impacts analyses etc.). We do not present the results of all analysis for all models in the stock assessment paper. In this section, we rely largely on the tabular results of the structural uncertainty grid (Table 4) and the dynamic spawning depletions and spawning potential plots for the models in the structural uncertainty grid (Figure 59, Figure 60). We also refer to the fished and unfished spawning potential trajectories for the diagnostic model discussed previously (Figure 52) and the dynamic Majuro and Kobe plots (Figure 63).

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

Points: The section summarizing the structural uncertainty grid (Section 9.5.2) presents terminal estimates of stock status in the form of Majuro plots. Further analyses can estimate the time-series of stock status in the form of Majuro and Kobe plots, the methods of which are presented in Section 7.7.4. The dynamic Majuro and Kobe plots for the diagnostic model models are presented in Figure 63.

Both the dynamic Majuro and Kobe plots show the steady increase in depletion of the stock since the 1070s. The dynamic Majuro plot indicates that the stock ends below the interim target reference point. The terminal spawning potential is well above B_{MSY} , the fishing mortality is well below F_{MSY} (Figure 63).

The summary of reference points in Table 4, indicates that both $SB_{recent}/SB_{F=0}$ and $SB_{latest}/SB_{F=0}$, are above the limit reference point across the models in the grid and the median $SB_{recent}/SB_{F=0}$ is just above the interim target reference point. Overall, the reference point values are consistent with the stock not being overfished and overfishing not occurring.

Fishing impact: In addition to the above analysis, it is possible to attribute the fishery impact with respect to depletion levels to specific fishery components (i.e., grouped by gear-type), in

order to estimate which types of fishing activity have the most impact on the spawning potential (Figure 64). The early impacts on the population were primarily driven by pole-and-line fishing, which is still a dominant source of fishing impacts in regions 1-3, and less so in region 4. For the other regions, and the stock as whole, the impact of that gear has declined to the extent that it is relatively small in regions 6, 7, and 8. Equatorial purse seine fishing is estimated to have had the most significant impact on spawning potential within regions 6, 7, and 8. Overall the associated and unassociated fishing methods have had similar impact over the last decade. The miscellaneous fisheries were estimated to have a significant impact on spawning potential in Region 5.

Yield analysis: The yield analyses conducted in this assessment incorporates the spawner recruitment relationship (Figure 49) into the equilibrium biomass and yield computations. Importantly, in the diagnostic model, the steepness of the SRR was fixed at 0.8 so only the scaling parameter was estimated. Other models in the one-off sensitivity analyses and structural uncertainty analyses assumed steepness values of 0.65 and 0.95.

The yield distributions under different values of fishing effort relative to the current effort are shown in Figure 65 for models representing different axes of the structural uncertainty grid (i.e, the alternative tag mixing scenarios and the externally estimated growth). For the diagnostic model, it is estimated that MSY would be achieved by approximately doubling fishing mortality, although the resulting increase in yield would be relatively small, and such an increase in fishing effort would unlikely be practical or economically viable. The right-hand arm of the yield curve displays a gradual decline in yield with increasing fishing mortality. The different example models shown display a similar pattern over the scale of fishing mortality although the absolute value of the yield curve differs significantly. The models with more conservative (longer) mixing periods and the external growth suggest higher yields.

10 Discussion and conclusions

10.1 Stock Status

The 2022 WCPO skipjack tuna stock assessment estimated that the median recent spawning depletion ($SB_{recent}/SB_{F=0}$) at the stock-wide scale is just above the interim target reference point and all models in the grid are well above the limit reference point (Figure 59). The spawning potential shows relative stability over the long-term, although at the regional scale recent declining trends are indicated for the equatorial regions 5, 7 and 8 (Figure 60). Overall the outcomes of this assessment suggest that the skipjack stock in the WCPO is not overfished or undergoing overfishing

The most notable feature of the assessment is the estimation that the stock is becoming increasingly depleted over time, a trend which is largely driven by the equatorial regions. Importantly, this trend is driven by an increasing trend in the model estimates of the unfished spawning potential over time, rather than a long-term decrease in the estimates of spawning potential (Figure 52). The assessment is indicating that the spawning potential has not changed substantially in the face of

the major increases in catches over the last 30 years, and that the increased catches have been sustained by increased recruitment levels (Figure 46). The interpretation of stock status based of the ($SB_{recent}/SB_{F=0}$) reference point should bear this in mind.

This model prediction of increased recruitment in response to increased catches poses some questions that require further consideration. The first one relates to aspects of the data inputs that provide information on biomass trends, most importantly the survey fishery CPUE indices. The main CPUE indices that inform the model on long-term biomass trends are those from the Japanese pole-and-line fisheries. These indices are both very stable over time, and have limited spatial coverage of the equatorial regions where most of the skipjack catch is taken. If the stability of these indices is a true reflection of biomass trends, in both the northern and equatorial regions, then the model estimates of increased recruitment likely reflect real changes in the pelagic ecosystem of the equatorial western Pacific that have driven increased recruitment since at least the late 1980s. If the CPUE indices are not a true reflection of biomass trends and suffer from hyperstability, then it is likely that the estimated trajectory of the skipjack biomass is also biased towards stability. Concerns that the pole-and-line CPUE indices used for the skipjack assessment are hyperstable and poorly represent the equatorial region have been expressed in previous assessments and CPUE analyses (i.e., Ducharme-Barth et al. (2022); Vincent et al. (2020)). Options for improvements in the development of CPUE indices are discussed below, along with further studies to explore the plausibility of a long-term increasing recruitment trend.

10.2 Changes to the previous assessment

The addition of three more years of data (tagging, catch, effort, length compositions) and several other model changes were introduced to the 2022 assessment. For the first time using MFCL, a catch conditioned model was applied with survey fishery CPUE abundance indices and a separate likelihood component for these indices. We also grouped survey fisheries considered to have the same selectivity and catchability to provide the model with information on relative biomass scale among regions. New unassociated purse seine CPUE indices for regions 6, 7 and 8 were added and to account for the contraction of the Japanese pole-and-line fishery we removed the last 23 years of data from the Region 8 pole-and-line CPUE time series. We implemented a novel approach using individual based modelling to simulate tag mixing processes and provide variable tag mixing values. The Dirichlet-multinomial likelihood was applied to weight the length composition data and removed an axis from the structural uncertainty grid, while also allowing estimation of growth within the model. Finally, to improve the representation of growth uncertainty in the model grid, we included a new growth curve developed from analysis of otolith daily aging and tag-recapture growth increments. These changes to the modelling approach resulted in a large reduction of estimable model parameters from 7642 for the 2019 diagnostic model to 2253 for the 2022 diagnostic model. The switch to the Dirichlet-multinomial had the impact of reducing the grid of models used in this assessment from 54 in 2019 to 18.

Despite these changes, the stock status outcomes of the 2022 assessment, and the estimation of historic trends, are not substantially different from the 2019 assessment. The depletion reference points for the 2022 assessment suggest a slightly less depleted status than for the 2019 assessment.

10.3 Structural uncertainties

The main structural uncertainties in the current assessment continued to be the tag mixing and growth assumptions. A considerable effort was made in response to recommendations from previous assessments to improve the way in which tag mixing periods were specified. This was achieved by implementing the variable mixing periods based on the simulation modelling by [Scutt Phillips et al. \(2022\)](#). However, while this provides a better approach to incorporating tag mixing uncertainty it does not remove the uncertainty altogether as it still requires some choices among options for classifying tag release groups as being mixed. Tag mixing remains a key and very influential uncertainty for the skipjack assessment due to its influence on biomass scale. The more conservative tag mixing scenario (applying the $K_s D$ statistic of 0.1) meant most of the tag-recaptures were recaptured within the specified mixing periods and could not contribute to model estimates. The stock status from these model are notably more optimistic than when the less conservative D statistic thresholds were applied and more tag-recaptures were available to the model. However, it is apparent that the tag mixing assumption is more influential on scaling of biomass and depletion rather than altering historical trends. Therefore relativity among time periods is largely consistent among the alternative mixing scenarios. This should be considered in relation to how these uncertainties influence the provision of management advice, particularly when reference points are based on particular reference years.

In relation to growth, considerable effort was made by [Macdonald et al. \(2022\)](#) to develop an empirical growth curve as an alternative to the internal model estimated growth. Again this was in response to the recommendations from previous assessments, and the 2022 PAW, to improve on the inclusion of growth uncertainty in the skipjack assessment. The study by [Macdonald et al. \(2022\)](#) made good use of empirical data available for estimating skipjack growth relevant to the WCPO region, and applied a very thoughtful and analytical sound approach to provide an alternative growth model. While this does not remove or necessarily reduce this uncertainty, it provides a better representation of growth uncertainty. Clearly, a lot more work is required to improve our estimation of skipjack growth, and we expect growth will continue to be an important uncertainty in skipjack stock assessments. While the external growth option scaled up the biomass and suggested a less depleted stock than the internal growth model, there were also some difference in the trends that point to a more dynamic influence of the growth assumptions on the model estimation of the key management reference points.

Application of the Dirichlet-multinomial likelihood was an improvement to this assessment that allowed removal of a grid axis for size data weighting for which values were rather arbitrarily allocated in previous assessments.

10.3.1 Model diagnostics

Consistent with 2019 assessment, it is clear from the likelihood profile (Section 15.1) on total biomass and the structural uncertainty grid that the tagging data remain an important source of information for the stock assessment. It is worth noting that the tagging data mostly provides information on movement, biomass scale and fishing mortality, but because of the sporadic spatiotemporal nature of this data, it does not provide much information on biomass trend. Biomass trend is largely informed by the CPUE indices. The length composition is influential on biomass scaling and to a lesser extent trends, as it provides information on recruitment and mortality. Overall the diagnostic model fits to the key data components were good, with exception of some of the CPUE fits, in particular, that could potentially be improved if a time varying penalty could be applied in future (this is currently not available with the new MFCL catch conditioned method). The retrospective analyses (Section 15.4) indicated that the model was quite stable in its estimation of spawning potential (biomass) and spawning depletion levels. However, there are some data conflicts evident from the likelihood profile on average biomass. In contrast with the 2019 assessment, the tagging data favour a substantially lower average biomass than the CPUE data, but a more similar biomass level to that favoured by the length composition. The overall likelihood prefers a biomass more closely aligned with CPUE and length composition data. The jitter analyses for the diagnostic model (Section 15.2) also showed that better model fits (improved total likelihoods) could be achieved, however, they resulted in negligible changes to the estimates of the key management reference point, $SB_{recent}/SB_{F=0}$. The likelihood surface in the region of best fit appears to be complex and it may be difficult to find a global minima for the complex skipjack models. Further simplification may be required to achieve this.

It is generally considered to be stock assessment ‘best practice’ that a fully converged model should produce a positive definite Hessian matrix (PDH). While not necessarily guaranteeing a plausible model, a PDH is good evidence that the model as specified has converged, and that the inverse Hessian provides a good approximation to the variance-covariance matrix of the model parameters, conditional on the specified model structure. As noted in section 6.6, in spite of generally good levels of model convergence as indicated by maximum parameter gradients, we were not successful in obtaining a PDH for any of our grid models. From the available diagnostics, there was no consistent picture of which parameter types were primarily responsible for the negative eigenvalues. Movement parameters and regional recruitment deviations were frequent contributors; however when models were simplified with regards to movement and recruitment distribution, other parameter types then emerged to produce negative eigenvalues. Unfortunately, there was insufficient time to fully diagnose this issue, but it remains an important area of follow-up work for this assessment.

10.4 Recommendations for further work

10.4.1 WCPFC-specific recommendations

The critical areas of work required to improve confidence in the outcomes of the skipjack assessment centre on the issues raised around the interpretation of the spawning depletion status and long-term stable spawning potential, as follows:

- **Indices of abundance:** The development of the unassociated purse seine index using VMS travel distance as an effort metric was promising and warrants further development, specifically to explore whether suitable proxies for travel distance between sets can be developed. If so this may allow the use of historical logbook CPUE to extend the equatorial purse seine indices back further in time (acknowledging some uncertainty of logbook-based catch estimates that may also require work to correct). Likewise, exploration of a distance between ‘active fishing’ metric for the Japanese pole-and-line vessels is recommended. VMS data should exist for the Japanese pole-and-line vessels going back several years, and this could be analysed in a similar way to that done for the unassociated purse seine index. Pole-and-line CPUE indices developed with the approach used in this assessment could then be contrasted, at least for recent years, with an alternative version using a distance based effort metric (i.e., travel distance between fishing events) to see if the trends differ.

It may not be possible to develop alternative distance based effort metrics for the historical time period, and the potential hyperstability in the CPUE trend for the pole-and-line fishery may be difficult to quantify. However, trends could be imposed on the indices to account for assumed effort creep. At the time of this assessment, while there was strong speculation that effort creep was influencing the long-term stability of the pole-and-line indices, there was insufficient time to do a thorough exploration of this using Japanese held operational data. The study by [Matsubara et al. \(2022\)](#) has provided a basis for a more thorough investigation to determine defensible effort creep scenarios to apply to the CPUE indices, and we recommend that this work continues, noting also that the 1% per annum effort creep scenario applied as a sensitivity in this assessment had little impact. In the meantime it would be worth conducting a series of models with increased levels of effort creep adjustment imposed on the CPUE indices to determine if increasing the effort creep assumptions will eventually remove the model estimation of the increasing recruitment over time. It could then be considered whether the amount of effort creep required to remove the recruitment and unfished biomass trend is technically feasible or consistent with other information.

Ultimately the use of fishery dependent abundance indices for a highly targeted school based fishery such as skipjack is not ideal and alternatives should be explored. Large scale fishery independent surveys of skipjack abundance are unlikely to be feasible. One approach that has come to prominence as a method to estimate tuna population sizes is the genetic method of Close Kin Mark Recapture (CKMR) ([Bravington et al., 2016, 2021](#)). This is currently being

explored for South Pacific albacore. Pending the feasibility work on albacore, the feasibility of this approach could be explored for skipjack in the WCPO, although current understanding suggests it may be problematic based on the current model based estimates of the skipjack population size.

- **Exploration of the increased recruitment hypothesis:** Attempts should be made to explore the plausibility and evidence for the model predicted long-term increasing trend in skipjack recruitment. Data sources including climatic, oceanographic, biological and even anecdotal from experienced skippers etc. should be included in this exploration. There is some evidence of consistency between the estimated total recruitment and evolution of the area of the western Pacific warm pool over several decades. Analyses should, however, attempt to use data independent from the stock assessment outputs to provide evidence or otherwise for the occurrence of increased skipjack recruitment in the equatorial Pacific over the last 20–30 years.
- **Growth:** Despite the thorough estimation of growth using available empirical data, more work on skipjack growth is important. Given the difficulties with aging skipjack using traditional methods (i.e., otoliths and spines) we recommend an exploration of the epigenetic aging approach using samples from individual tag-recapture skipjack.
- **Model simplification:** Further work on the skipjack model to explore the sources of negative eigenvalues and test alternative model structures to see what is required to obtain a positive definite Hessian solution, and reduce data conflicts is also recommended.
- **Tag mixing simulations:** The simulation modelling framework for exploring tag mixing has value for numerous other investigations related to skipjack behavioural ecology and strategic planning of tagging effort. We suggest that a workshop to discuss these possibilities be convened to develop both proposals, and a strategy for its further development and application to other tuna species.
- **Population structure:** There remains a poor understanding of meta-population structuring of skipjack in the Western Pacific, particularly on the linkages between populations in the east Asian waters and those in the broader western and central Pacific. The fisheries in the east Asian region catch large numbers of smaller skipjack. Understanding if and how important this region is for replenishment of skipjack more broadly in the western and central Pacific is important for considering how this large catch is included within the WCPO skipjack assessments and for other management related modelling work, such as the development of harvest strategies.

10.5 Main assessment conclusions

The general conclusions of this assessment are as follows:

- Spawning potential, as informed by several stable CPUE abundance indices, has remained relatively stable, with fluctuations, until around 2010, after which it declines gradually, driven by trends in the equatorial regions. Spawning depletion has declined gradually since the start of the model period. This decline is largely due to the increasing estimates of the unfished spawning potential and recruitment from 1980 to the recent period.
- Average fishing mortality rates for juvenile and adult age-classes increase throughout the period of the assessment.
- Overall median depletion from the model grid for the recent period (2018-2021; $SB_{recent}/SB_{F=0}$) is estimated at 0.51 (80 percentile range 0.43-0.64).
- No models from the structural uncertainty grid estimate the stock to be below the LRP.
- Recent (2017-2020) median fishing mortality (F_{recent}/F_{MSY}) was 0.32 (80 percentile range 0.18-0.45).

We conclude that the skipjack stock in the WCPO is not overfished, nor undergoing overfishing.

11 Acknowledgements

We thank the various fisheries agencies and regional fisheries observers for their support with data collection, provision and preparatory analysis. In particular we especially thank Yoshi Aoki, Naoto Matsubara and Yuichi Tsuda for their support on developing the pole-and-line CPUE indices and the provision and analysis of the JPTP tagging and length composition data used in this analysis. We also thank Keith Bigelow who again provided standardized CPUE indices for the Philippines purse seine fishery, and Inna Senina for providing the movement coefficients for the SEAPODYM sensitivity. We thank participants at the preparatory stock assessment workshop for their contributions to the assessment. We especially thank the SPC data management team for their hard work and support to provide the data fuel for the assessment, and make special mention of the support from Manu Schneider and Peter Williams, plus Helene Ixeko, Christine Nguyen, Stephanie Nguyen Cuu and Sylvie Pironnec for their great team effort to validate 100s of tag returns at the last minute. We appreciate the advice from Cole Monnahan, Andre Punt and Ian Taylor on 'Life, the universe and positive definite Hessians'. Finally, we thank Fabrice Bouye for ensuring smooth operation of the computing resources required to do the assessment. Funding from Pacific-European Union (EU) Marine Partnership (PEUMP) supported the work on CPUE analysis and data preparation by Thom Teears.

12 References

References

- Artetxe-Arrate, I., Fraile, I., Marsac, F., Farley, J. H., Rodriguez-Ezpeleta, N., Davies, C. R., Clear, N. P., Grewe, P., and Murua, H. (2021). A review of the fisheries, life history and stock structure of tropical tuna (skipjack *Katsuwonus pelamis*, yellowfin *Thunnus albacares* and bigeye *Thunnus obesus*) in the Indian Ocean. In *Advances in Marine Biology*, volume 88, pages 39–89. Elsevier.
- Ashida, H., Tanabe, T., and Suzuki, N. (2017). Difference on reproductive trait of skipjack tuna *Katsuwonus pelamis* female between schools (free vs FAD school) in the tropical western and central Pacific Ocean. *Environ Biol Fish*, 100(8):935–945.
- Bigelow, K., Garvilles, E., Drusila, E., and Cecilio, C. (2019). Relative abundance of skipjack tuna for the purse seine fishery operating in the Philippines Moro Gulf (Region 12) and High Seas Pocket #11. Technical Report WCPFC-SC15-2019/ SA-IP-08.
- Bigelow, K., Hampton, J., and Fournier, D. A. (2000). Preliminary application of the MULTIFAN CL model to skipjack tuna in the tropical WCPO. SCTB13 Working Paper SKJ-2.
- Bravington, M., Nicol, S., and Anderson, G. (2021). Feasibility of Close-Kin-Mark-Recapture for albacore in the WCPO (project 100b). Technical Report WCPFC-SC17/SA-IP-14.
- Bravington, M. V., Grewe, P. M., and Davies, C. R. (2016). Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. *Nat Commun*, 7(1):13162.
- Cadigan, N. G. and Farrell, P. J. (2005). Local influence diagnostics for the retrospective problem in sequential population analysis. *ICES Journal of Marine Science: Journal du Conseil*, 62(2):256–265.
- Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K. R., Maunder, M. N., Taylor, I., Wetzell, C. R., Doering, K., Johnson, K. F., and Methot, R. D. (2021). A cookbook for using model diagnostics in integrated stock assessments. *Fisheries Research*, 240:105959.
- Castillo Jordan, C., Hampton, J., Ducharme-Barth, N., Xu, H., Vidal, T., Williams, P., Scott, F., Pilling, G., and Hamer, P. (2021). Stock assessment of South Pacific albacore tuna. Technical Report WCPFC-SC17-2021/SA-WP-02.
- Davies, N., Fournier, D., Bouye, F., and Hampton, J. (2022). Developments in the Multifan-CL software 2021-2022. Technical Report WCPFC-SC18-2022/SA-IP-03.
- Ducharme-Barth, N., Vincent, M., Pilling, G., and Hampton, J. (2019). Simulation analysis of pole and line CPUE standardization approaches for skipjack tuna in the WCPO. Technical Report WCPFC-SC15-2019/SA-WP-04, Pohnpei, Federated States of Micronesia 12-20 August 2019.

- Ducharme-Barth, N. D., Grüss, A., Vincent, M. T., Kiyofuji, H., Aoki, Y., Pilling, G., Hampton, J., and Thorson, J. T. (2022). Impacts of fisheries-dependent spatial sampling patterns on catch-per-unit-effort standardization: A simulation study and fishery application. *Fisheries Research*, 246:106169.
- Fournier, D. and Archibald, C. P. (1982). A general-theory for analyzing catch at age data. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(8):1195–1207.
- Fournier, D., Hampton, J., and Sibert, J. (1998). MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences*, 55:2105–2116.
- Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielson, A., and Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, 27(2):233–249.
- Francis, R. I. C. C. (1992). Use of risk analysis to assess fishery management strategies: A case study using orange roughy (it *Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. *Canadian Journal of Fisheries and Aquatic Science*, 49:922–930.
- Gaertner, D., gado de Molina, A., Ariz, J., Pianet, R., and Hallier, J. P. (2008). Variability of the growth parameters of the skipjack tuna (it *Katsuwonus pelamis*) among areas in the eastern Atlantic: analysis from tagging data within a meta-analysis approach. *Aquatic Living Resources*, 21:349–356.
- Grewe, P. M., Wudianto, C., 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., and Davies, C. R. (2019). Population Structure and Connectivity of Tropical Tuna Species across the Indo Pacific Ocean Region. Technical Report WCPFC-SC15-2019/SA-IP-15.
- Hamer, P. (2022). Report from the SPC Pre-assessment Workshop - March 2022. WCPFC-SC18-2022/SA-IP-02.
- 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. and Fournier, D. (2001). A spatially-disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Marine and Freshwater Research*, 52:937–963.
- Hampton, J. and Williams, P. (2016). Annual estimates of purse seine catches by species based on alternative data sources. Technical Report WCPFC-SC12-2016/ST-IP-03, Bali, Indonesia, 3–11 August 2016.

- Harley, S. J. (2011). A preliminary investigation of steepness in tunas based on stock assessment results. Technical Report WCPFC-SC7-2011/SA-IP-08, Pohnpei, Federated States of Micronesia, 9–17 August 2011.
- Hoyle, S., Kleiber, P., Davies, N., Harley, S., and Hampton, J. (2010). Stock assessment of skipjack tuna in the western and central Pacific Ocean. Technical Report WCPFC-SC6-2010/ST-IP-02, Nuku'alofa, Tonga, 10–19 August 2010.
- 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., and 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.
- Ianelli, J., Maunder, M. N., and Punt, A. E. (2012). Independent review of the 2011 WCPO bigeye tuna assessment. Technical Report WCPFC-SC8-2012/SA-WP-01, Busan, Republic of Korea, 7–15 August 2012.
- ISSF (2011). Report of the 2011 ISSF stock assessment workshop. Technical Report ISSF Technical Report 2011-02, Rome, Italy, March 14–17.
- Itano, D. (2000). The reproductive biology of yellowfin tuna (it *Thunnus albacares*) in Hawaiian waters and the western tropical Pacific Ocean: Project summary. JIMAR Contribution 00-328 SOEST 00-01.
- Kinoshita, J., Aoki, Y., Ducharme-Barth, N., and Kiyofuji, H. (2019). Standardized catch per unit effort (CPUE) of skipjack tuna of the Japanese pole-and-line fisheries in the WCPO from 1972 to 2018. Technical Report WCPFC-SC15-2019/SA-WP-14, Pohnpei, Federated States of Micronesia.
- Kiyofuji, H., Aoki, Y., Kinoshita, J., Okamoto, S., Masujima, M., Matsumoto, T., Fujioka, K., Ogata, R., Nakao, T., Sugimoto, N., and Kitagawa, T. (2019a). Northward migration dynamics of skipjack tuna (*Katsuwonus pelamis*) associated with the lower thermal limit in the western Pacific Ocean. *Progress in Oceanography*, 175:55–67.
- Kiyofuji, H. and Ochi, D. (2016). Proposal of alternative spatial structure for skipjack stock assessment in the WCPO. Technical Report WCPFC-SC12-2016/SA-IP-09, Bali, Indonesia, 3–11 August 2016.
- Kiyofuji, H., Ohashi, S., Kinoshita, J., and Aoki, Y. (2019b). Overview of historical skipjack length and weight data collected by the Japanese pole-and-line fisheries and Research vessel (R/V) from 1953 to 2017. Technical Report WCPFC-SC15-2019/SA-IP-12.
- Kleiber, P., Fournier, D., Hampton, J., Davies, N., Bouye, F., and Hoyle, S. (2019). MULTIFAN-CL User's Guide. Technical report.

- Lawson, T. (2011). Purse-seine length frequencies corrected for selectivity bias in grab samples collected by observers. Technical Report WCPFC-SC7-2011/ST-IP-02, Pohnpei, Federated States of Micronesia, 9–17 August 2011.
- Lehodey, P., Bertignac, M., Hampton, J., Lewis, A., and Picaut, J. (1997). El Nino Southern Oscillation and tuna in the western Pacific. *Nature*, 389:715–718.
- Lehodey, P., Senina, I., and Murtugudde, R. (2008). A spatial ecosystem and populations dynamics model (SEAPODYM) – Modeling of tuna and tuna-like populations. *Progress in Oceanography*, 78(4):304–318.
- Leroy, B. (2000). Preliminary results on skipjack (it *Katsuwonus pelamis*) growth. Technical Report SKJ-1, 13th Meeting of the Standing Committee on Tuna and Billfish. Noumea, New Caledonia.
- Macdonald, J., Day, J., Magnusson, A., Maunder, M., Aoki, Y., McKechnie, S., Tears, T., Castillo Jordan, C., Hampton, J., and Hamer, P. (2022). Review and new analyses of skipjack growth in the Western and Central Pacific Ocean. Technical Report WCPFC-SC18-2022/SA-IP-06.
- Matsubara, N., Aoki, Y., and Tsuda, Y. (2022). Historical development of fishing devices in Japanese pole-and-line fishery. Technical Report WCPFC-SC18-2022/SA-IP-16.
- Maunder (2001). Growth of skipjack tuna (*Katsuwonus pelamis*) in the easter Pacific Ocean, as eestimate from tagging data. Technical Report Bulletin of the Inter-American Tropical Tuna Commision, vol. 22.
- McKechnie, S., Hampton, J., Pilling, G. M., and Davies, N. (2016). Stock assessment of skipjack tuna in the western and central Pacific Ocean. Technical Report WCPFC-SC12-2016/SA-WP-04, Bali, Indonesia, 3–11 August 2016.
- McKechnie, S., Pilling, G., and Hampton, J. (2017). Stock assessment of bigeye tuna in the western and central Pacific Ocean. Technical Report WCPFC-SC13-2017/SA-WP-05, Rarotonga, Cook Islands, 9–17 August 2017.
- Methot, Richard D., J. and Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, 142:86–99.
- Moore, B. R., Bell, J. D., Evans, K., Farley, J., Grewe, P. M., Hampton, J., Marie, A. D., Minte-Vera, C., Nicol, S., Pilling, G. M., Scutt Phillips, J., Tremblay-Boyer, L., Williams, A. J., and Smith, N. (2020). Defining the stock structures of key commercial tunas in the Pacific Ocean I: Current knowledge and main uncertainties. *Fisheries Research*, 230:105525.
- Oceanic Fisheries Programme, P. C. (2021). Western and Central Pacific Fisheries Commission Tuna Fisheries Yearbook 2020. Technical report.

- Ochi, D., Ijima, H., and Kiyofuji, H. (2016). A re-consideration of growth pattern of skipjack on the western central Pacific. Technical Report WCPFC-SC12-2016/SA-IP-08, Bali, Indonesia, 3–11 August 2016.
- Ohashi, S., Aoki, Y., Tanaka, F., Fujioka, K., Aoki, A., and Kiyofuji, H. (2019). Reproductive traits of female skipjack tuna *Katsuwonus pelamis* in the western central Pacific Ocean (WCPO). Technical report.
- Peatman, T. (2022). Analysis of tag seeding data and reporting rates for purse seine fleets. Technical Report WCPFC-SC18-2022/SA-IP-19.
- Peatman, T., Ducharme-Barth, N., and Vincent, M. (2020). Analysis of purse-seine and longline size frequency data for bigeye and yellowfin tuna in the WCPO. Technical Report SC16-SA-IP-18.
- Peatman, T., Scutt Phillips, J., Potts, J., and Nicol, S. (2022). Analysis of tagging data for the 2022 skipjack tuna assessment: corrections for tagging conditions. Technical Report WCPFC-SC18-2022/SA-IP-20.
- Peatman, T., Williams, P., and Nicol, S. (2021). Project 60: Progress towards achieving SC16 recommendations. Technical Report WCPFC-SC17-2021/ST-IP-04.
- Pilling, G. and Brouwer, S. (2019). Report from the SPC Pre-assessment Workshop, Noumea, April 2019. Technical Report WCPFC-SC15-2019/SA-IP-01, Pohnpei, Federated States of Micronesia.
- Pilling, G., Scott, R., Davies, N., and Hampton, J. (2016). Approaches used to undertake management projections of WCPO tuna stocks based upon MULTIFAN-CL stock assessments. Technical Report WCPFC-SC12-2016/MI-IP-04, Bali, Indonesia, 3–11 August 2016.
- Rice, J., Harley, S., Davies, N., and Hampton, J. (2014). Stock assessment of skipjack tuna in the Western and Central Pacific Ocean. Technical Report WCPFC-SC10-2014/SA-WP-05, Majuro, Republic of the Marshall Islands, 6–14 August 2014.
- Schnute, J. (1981). A Versatile Growth Model with Statistically Stable Parameters. *Can. J. Fish. Aquat. Sci.*, 38(9):1128–1140.
- Scutt Phillips, J., Lehodey, J., Hampton, J., Senina, I., and Nicol, S. (2022). Quantifying Rates of Mixing in Tagged, WCPO Skipjack Tuna. Technical Report WCPFC-SC18-2022/SA-WP-04.
- Scutt Phillips, J., Sen Gupta, A., Senina, I., Seville, E., Lange, M., Lehodey, P., J, H., and Nichol, S. (2018). An individual-based model of skipjack tuna *Katsuwonus pelamis* movement in the tropical Pacific Ocean. *Progress in Oceanography*, 164:63–74.
- Senina, I., Lehodey, P., Calmettes, B., Nichol, S. J., Caillot, S., J, H., and P, W. (2016). Predicting skipjack tuna dynamics and effects of climate change using SEAPODYM with fishing and tagging data. Technical Report WCPFC-SC12-2016/EB WP01.

- Senina, I., Lehodey, P., Sibert, J., and Hampton, J. (2020a). Integrating tagging and fisheries data into a spatial population dynamics model to improve its predictive skills. *Can. J. Fish. Aquat. Sci.*, 77(3):576–593.
- Senina, I. N., Lehodey, P., Hampton, J., and Sibert, J. (2020b). Quantitative modelling of the spatial dynamics of South Pacific and Atlantic albacore tuna populations. *Deep Sea Research Part II: Topical Studies in Oceanography*, 175:104667.
- SPC-OFP (2013). Purse seine effort: a recent issue in logbook reporting. Technical Report WCPFC-TCC9-2013-18, Pohnpei, Federated States of Micronesia.
- Tanabe, T., Kayama, S., and Ogura, M. (2003). Precise age determination of young to adult skipjack tuna (*Katsuwonus pelamis*) with validation of otolith daily increment. Technical Report SCTB16/SKJ-08, 16th Meeting of the Standing Committee on Tuna and Billfish, Mooloolaba, Queensland, Australia, 9–16 August.
- Tears, T., Aoki, Y., Matsubura, N., Tsuda, Y., Castillo Jordon, C., Hampton, J., Schneiter, E., Scutt Phillips, J., Peatman, T., and Hamer, P. (2022). Background analyses and data inputs for the 2022 skipjack tuna stock assessment in and central Pacific Ocean. Technical Report WCPFC-SC18-2022/SA-IP-05.
- Thorson, J. T., Johnson, K. F., Methot, R. D., and Taylor, I. G. (2017). Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. *Fisheries Research*, 192:84–93.
- Thorson, J. T., Shelton, A. O., Ward, E. J., and Skaug, H. J. (2015). Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *Ices Journal of Marine Science*, 72(5):1297–1310.
- Tremblay-Boyer, L., McKechnie, S., Pilling, G., and Hampton, J. (2017). Stock assessment of yellowfin tuna in the Western and Central Pacific Ocean. Technical Report WCPFC-SC13-2017/SA-WP-06, Rarotonga, Cook Islands, 9-17 August 2017.
- Vidal, T., Hamer, P., Escalle, L., and Pilling, G. (2020). Assessing trends in skipjack tuna abundance from purse seine catch and effort data in the WCPO. Technical Report WCPFC-SC16-2020/SA-IP-09.
- Vidal, T., Pilling, G., Tremblay-Boyer, L., and Usu, T. (2019). Standardized CPUE for skipjack tuna *Katsuwonus pelamis* from the Papua New Guinea archipelagic purse seine fishery. Technical Report WCPFC-SC15-2019/SA-IP-05, Pohnpei, Federated States of Micronesia.
- Vincent, M., Aoki, Y., Kiyofuji, H., Hampton, J., and Pilling, G. M. (2019a). Background analyses for the 2019 stock assessment of skipjack tuna. Technical Report WCPFC-SC15-2019/SA-IP-04.

- Vincent, M., Ducharme-Barth, N., Hamer, P., Hampton, J., Williams, P., and Pilling, G. (2020). Stock assessment of yellowfin tuna in the western and central Pacific Ocean. WCPFC-SC16-2020/SA-WP-04-Rev2.
- Vincent, M., Pilling, G., and Hampton, J. (2019b). Stock assessment of skipjack tuna in the WCPO. Technical Report WCPFC-SC15-2019/SA-WP-05, Pohnpei, Federated States of Micronesia.
- Wang, X., Chen, Y., Truesdell, S., Xu, L., Cao, J., and Guan, W. (2014). The Large-Scale Deployment of Fish Aggregation Devices Alters Environmentally-Based Migratory Behavior of Skipjack Tuna in the Western Pacific Ocean. *PLoS ONE*, 9(5):e98226.
- Wild, A. and Hampton, J. (1994). A review of the biology and fisheries for skipjack tuna, *Katsuwonus pelamis*, in the Pacific Ocean. FAO Fisheries Technical Paper (FAO) 336(2):1–51.

13 Tables

Table 1: Definition of fisheries for the 2022 MULTIFAN-CL skipjack tuna stock assessment in the WCPO.

Fishery Number	Gear	Model Code-Fleets	Flags	Model Region	Sel Group	Sel Nodes	Recapt Group	Report Group
F1	PL	P-ALL-1	ALL	1	1	4	1	1
F2	PS	S-ALL-1	ALL	1	2	4	2	1
F3	LL	L-ALL-1	ALL	1	3	3	3	1
F4	PL	P-ALL-2	ALL	2	1	4	4	1
F5	PS	S-ALL-2	All	2	3	4	5	1
F6	LL	L-ALL-2	ALL	2	3	3	6	1
F7	PL	P-ALL-3	ALL	3	1	4	7	1
F8	PS	S-ALL-3	ALL	3	3	4	8	1
F9	LL	L-ALL-3	ALL	3	3	3	9	1
F10	Dom	Z-PH-5	PH	5	4	6	10	2
F11	Dom	Z-ID-5	ID	5	5	5	11	3
F12	PS	S-ID.PH-5	ID.PH	5	6	5	12	4
F13	PL	P-ALL-5	ALL	5	7	6	13	1
F14	PSAssoc	SA-DW-5	ALL	5	8	4	14	5
F15	PSUnassoc	SU-DW-5	ALL	5	9	4	14	5
F16	Dom	Z-VN-5	VN	5	10	5	15	6
F17	LL	L-ALL-5	ALL	5	11	3	16	1
F18	PL	P-ALL-6	ALL	6	12	4	17	1
F19	PSAssoc	SA-ALL-6	ALL	6	8	4	18	7
F20	PSUnassoc	SU-ALL-6	ALL	6	9	4	18	7
F21	LL	L-ALL-6	ALL	6	11	3	19	1
F22	PL	P-ALL-4	ALL	4	1	4	20	1
F23	LL	L-ALL-4	ALL	4	3	3	21	1
F24	PL	P-ALL-7	ALL	7	1	4	22	1
F25	PSAssoc	SA-ALL-7	ALL	7	8	4	23	8
F26	PSUnassoc	SU-ALL-7	ALL	7	9	4	23	8
F27	LL	L-ALL-7	ALL	7	11	3	24	1
F28	PL	P-ALL-8	ALL	8	1	4	25	1
F29	PSAssoc	SA-ALL-8	ALL	8	8	4	26	9
F30	PSUnassoc	SU-ALL-8	ALL	8	9	4	26	9
F31	LL	L-ALL-8	ALL	8	13	3	27	1
Survey fisheries								
F32	PL	PL-ALL-1	INDEX	1	1	4	-	-
F33	PL	PL-ALL-2	INDEX	2	1	4	-	-
F34	PL	PL-ALL-3	INDEX	3	1	4	-	-
F35	PL	PL-ALL-4	INDEX	4	1	4	-	-
F36	PS	PS-PH-5	INDEX	5	6	5	-	-
F37	PSUnassoc	PS-ALL-6	INDEX	6	9	4	-	-
F38	PSUnassoc	PS-ALL-7	INDEX	7	9	4	-	-
F39	PL	PL-ALL-7	INDEX	7	1	4	-	-
F40	PSUnassoc	PS-ALL-8	INDEX	8	9	4	-	-
F41	PL	PL-ALL-8	INDEX	8	1	4	-	-

Table 2: Structural uncertainty grid for the 2022 WCPO skipjack tuna stock assessment. Bold values indicate settings for the diagnostic. For information on the mixing axis refer to [Section 7.4.2](#), and for growth refer [Section 7.2.3](#)

Axis	Levels	Option 1	Option 2	Option 3
Tag mixing	3	T1, D=0.1 (longer period)	T2 , D=0.2 (intermediate)	T3, D=0.3 (shorter)
Growth	2	G1 , Internally estimated (Dirichlet-multinomial)	G2, Externally estimated (otolith and tagging data)	
Steepness	3	0.65	0.8	0.95

Table 3: Description of symbols used in the yield and stock status analyses.

Symbol	Description
C_{latest}	Catch in the last year of the assessment (2021)
F_{recent}	Average fishing mortality-at-age for a recent period (2017–2020)
$Y_{F_{recent}}$	Equilibrium yield at average fishing mortality for a recent period (2017–2020)
f_{mult}	Fishing mortality multiplier at maximum sustainable yield (MSY)
F_{MSY}	Fishing mortality-at-age producing the maximum sustainable yield (MSY)
MSY	Equilibrium yield at F_{MSY}
F_{recent}/F_{MSY}	Average fishing mortality-at-age for a recent period (2017–2020) relative to F_{MSY}
SB_{latest}	Spawning biomass in the latest time period (2021)
SB_{recent}	Spawning biomass for a recent period (2018–2021)
$SB_{F=0}$	Average spawning biomass predicted in the absence of fishing for the period 2011–2020
SB_{MSY}	Spawning biomass that will produce the maximum sustainable yield (MSY)
$SB_{MSY}/SB_{F=0}$	Spawning biomass that produces maximum sustainable yield (MSY) relative to the average spawning biomass predicted to occur in the absence of fishing for the period 2011–2020
$SB_{latest}/SB_{F=0}$	Spawning biomass in the latest time period (2021) relative to the average spawning biomass predicted to occur in the absence of fishing for the period 2011–2020
SB_{latest}/SB_{MSY}	Spawning biomass in the latest time period (2021) relative to that which will produce the maximum sustainable yield (MSY)
$SB_{recent}/SB_{F=0}$	Spawning biomass for a recent period (2018–2021) relative to the average spawning biomass predicted to occur in the absence of fishing for the period 2011–2020
SB_{recent}/SB_{MSY}	Spawning biomass for a recent period (2018–2021) relative to the spawning biomass that produces maximum sustainable yield (MSY)
$20\%SB_{F=0}$	WCPFC adopted limit reference point – 20% of spawning biomass in the absence of fishing average over years $t - 10$ to $t - 1$ (2011–2020)

Table 4: Summary of reference points over the 18 individual models in the structural uncertainty grid.

	mean	median	min	10%ile	90%ile	max	diagnostic model
C_{latest}	1530209	1530208	1530207	1530207	1530212	1530212	1530207
F_{MSY}	0.23	0.23	0.18	0.19	0.27	0.28	0.24
$fmult$	3.61	3.18	1.88	2.22	5.54	8.08	2.86
F_{recent}/F_{MSY}	0.32	0.32	0.12	0.18	0.45	0.53	0.35
MSY	2933489	2648400	2046000	2167840	4777200	4868000	2416000
SB_0	7958888	7204500	5317000	5611000	12842000	14390000	5686000
$SB_{F=0}$	8073171	7616930	5953338	6156944	12310363	12744728	6147339
SB_{latest}/SB_0	0.48	0.48	0.37	0.41	0.56	0.60	0.48
$SB_{latest}/SB_{F=0}$	0.47	0.46	0.35	0.38	0.60	0.61	0.44
SB_{latest}/SB_{MSY}	2.82	2.68	1.65	1.95	3.81	4.62	2.54
SB_{MSY}	1419366	1335000	806300	870530	1984600	2925000	1073000
SB_{MSY}/SB_0	0.18	0.18	0.13	0.13	0.22	0.22	0.19
$SB_{MSY}/SB_{F=0}$	0.17	0.17	0.11	0.13	0.22	0.23	0.17
$SB_{recent}/SB_{F=0}$	0.52	0.51	0.41	0.43	0.64	0.66	0.50
SB_{recent}/SB_{MSY}	3.12	2.98	1.92	2.20	4.22	4.97	2.88
$Y_{F_{recent}}$	1896888	1892400	1621600	1683880	2116000	2282800	1762400
$(SB_{recent}/SB_{F=0})/(SB_{2012}/SB_{F=0})$	0.84	0.85	0.82	0.82	0.86	0.87	0.85

14 Figures

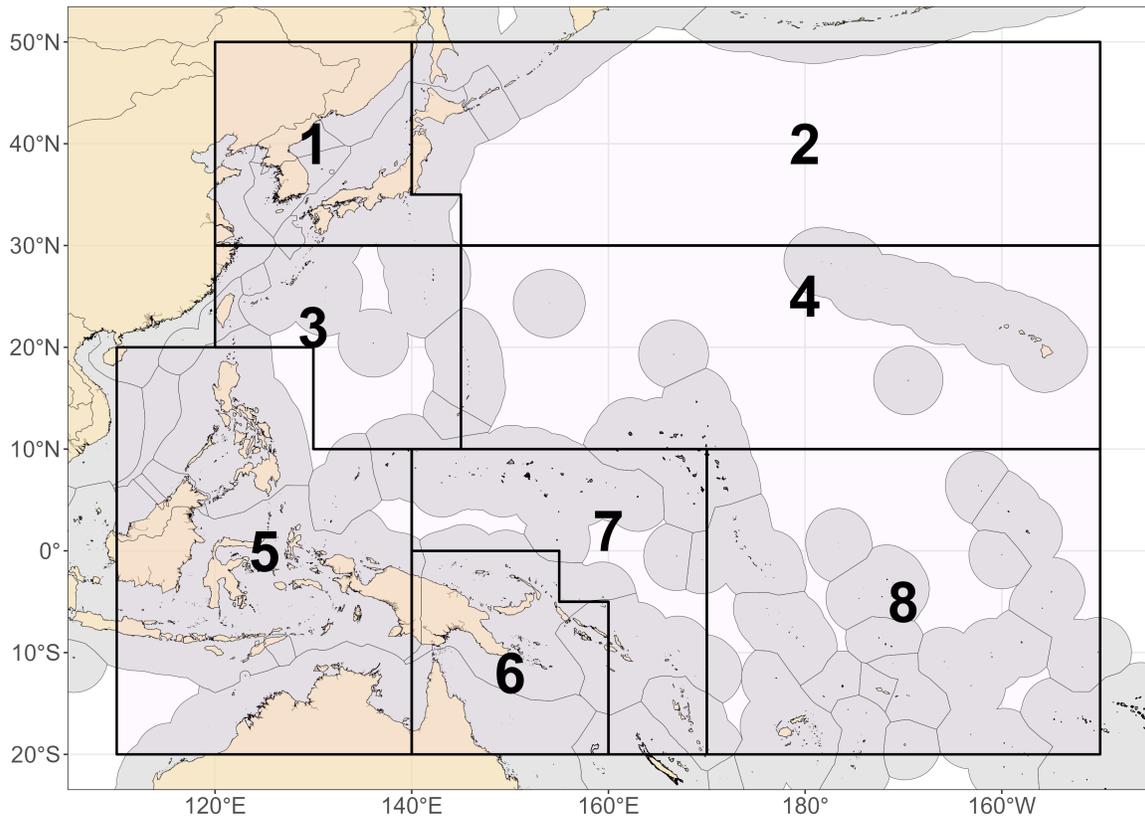


Figure 1: The geographical area covered by the stock assessment and the boundaries of the eight model regions used for 2022 WCPO skipjack assessment.

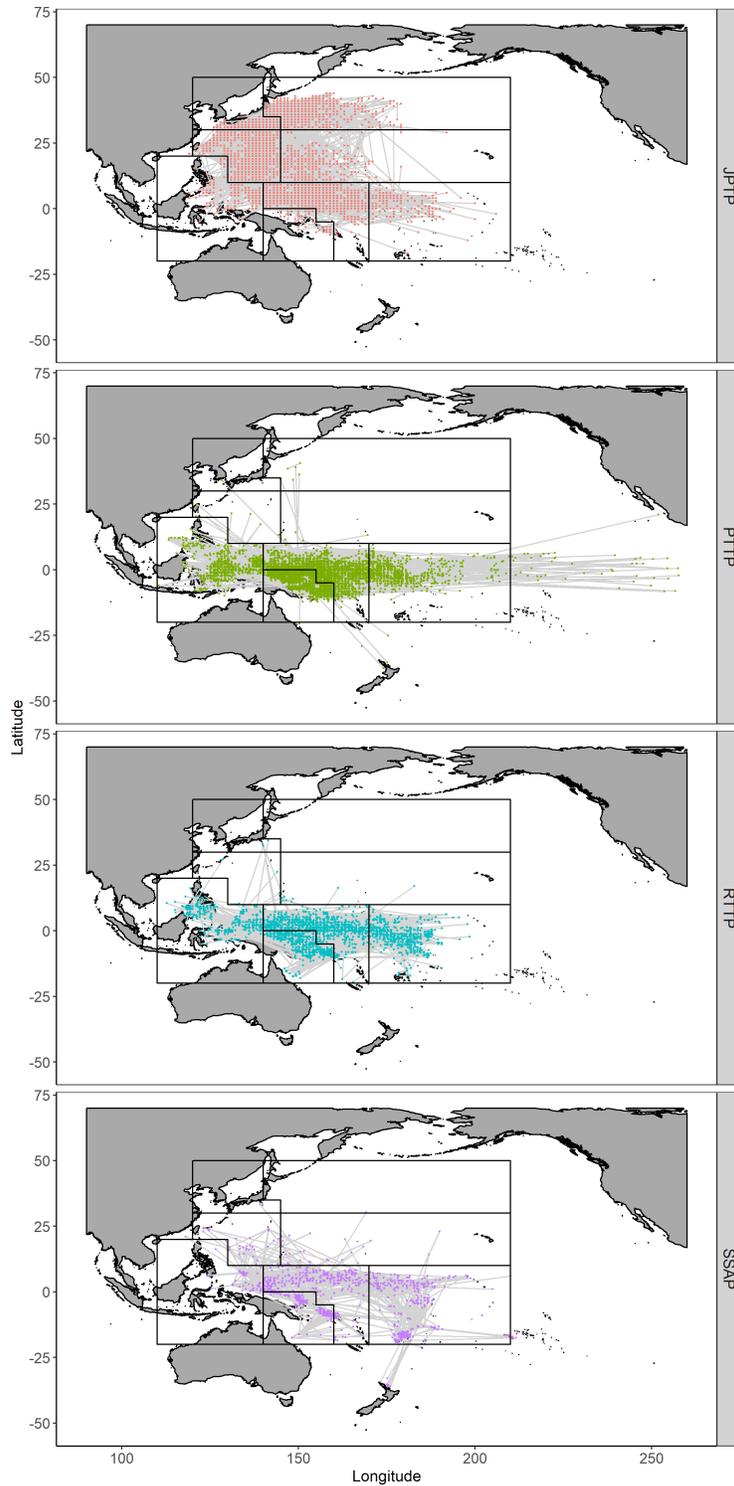


Figure 2: Map of tag recaptures. The panel shows the distributions of release and recapture displacements for the different tagging programs. Pacific Tuna Tagging Program (PTTP), Regional Tuna Tagging Program (RTTP), Skipjack Survey and Assessment Program (SSAP), and the Japanese Tagging Program (JPTP).

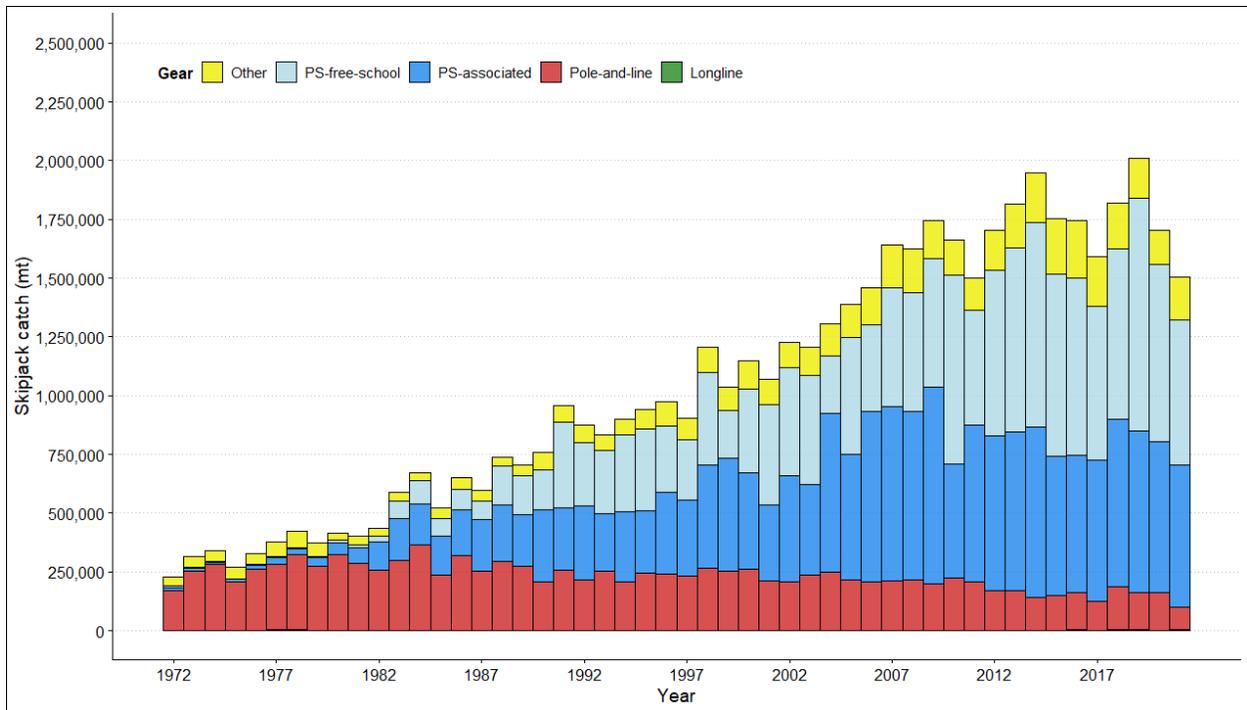


Figure 3: Annual catches of skipjack by gear type in the WCPO area covered by the assessment.

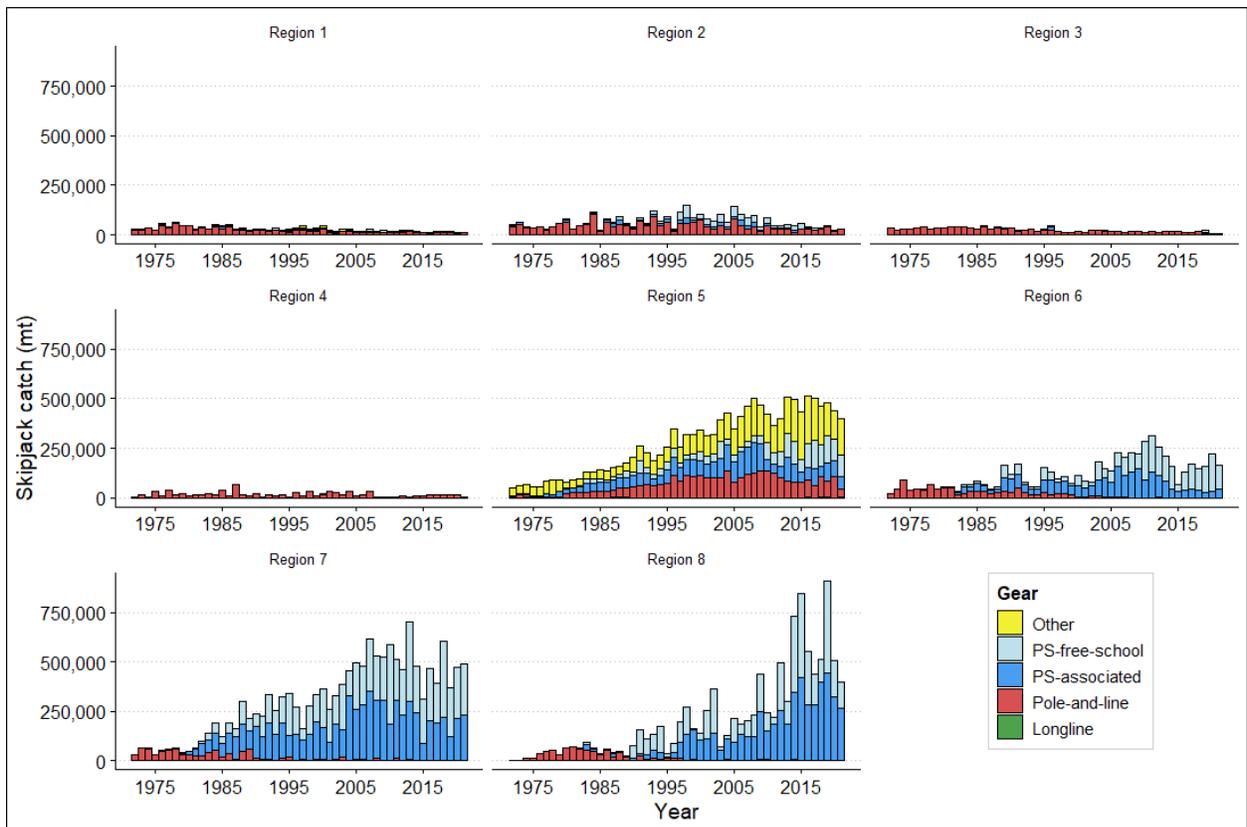


Figure 4: Annual catches of skipjack by gear type for each of the eight model regions.

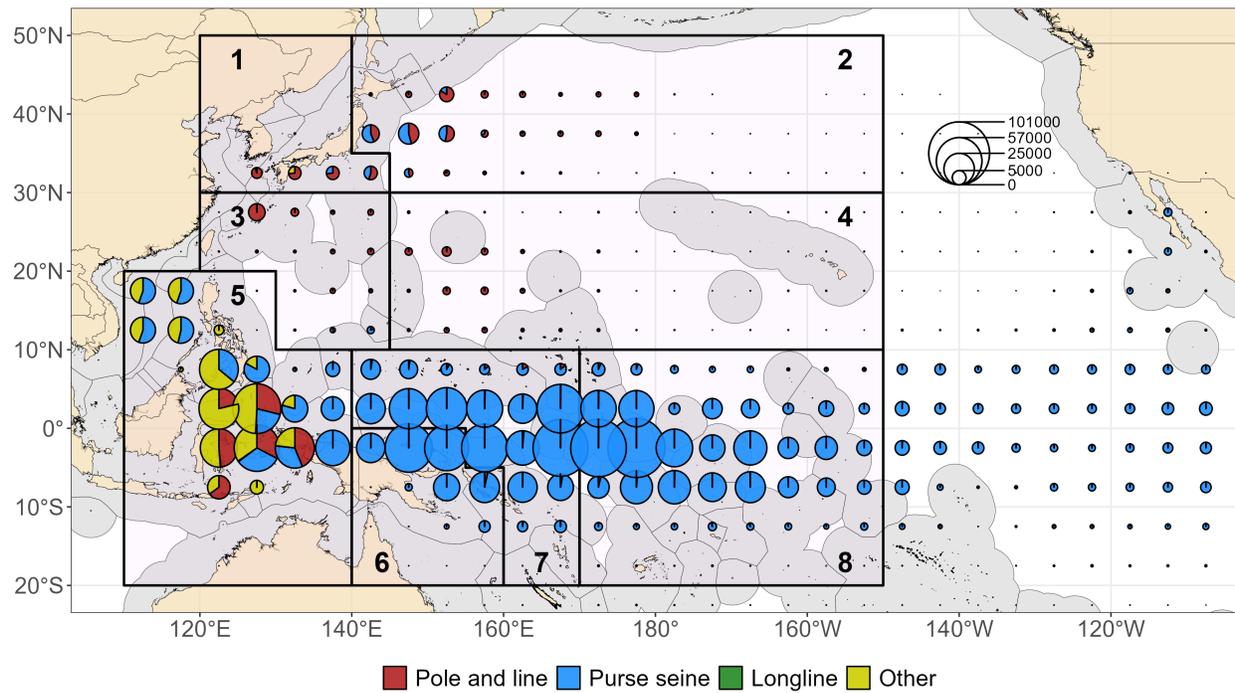


Figure 5: Distribution and magnitude of skipjack catches (mt) by gear type summed over the last 10 years (2012-2021) for 5 x 5 degree cells.

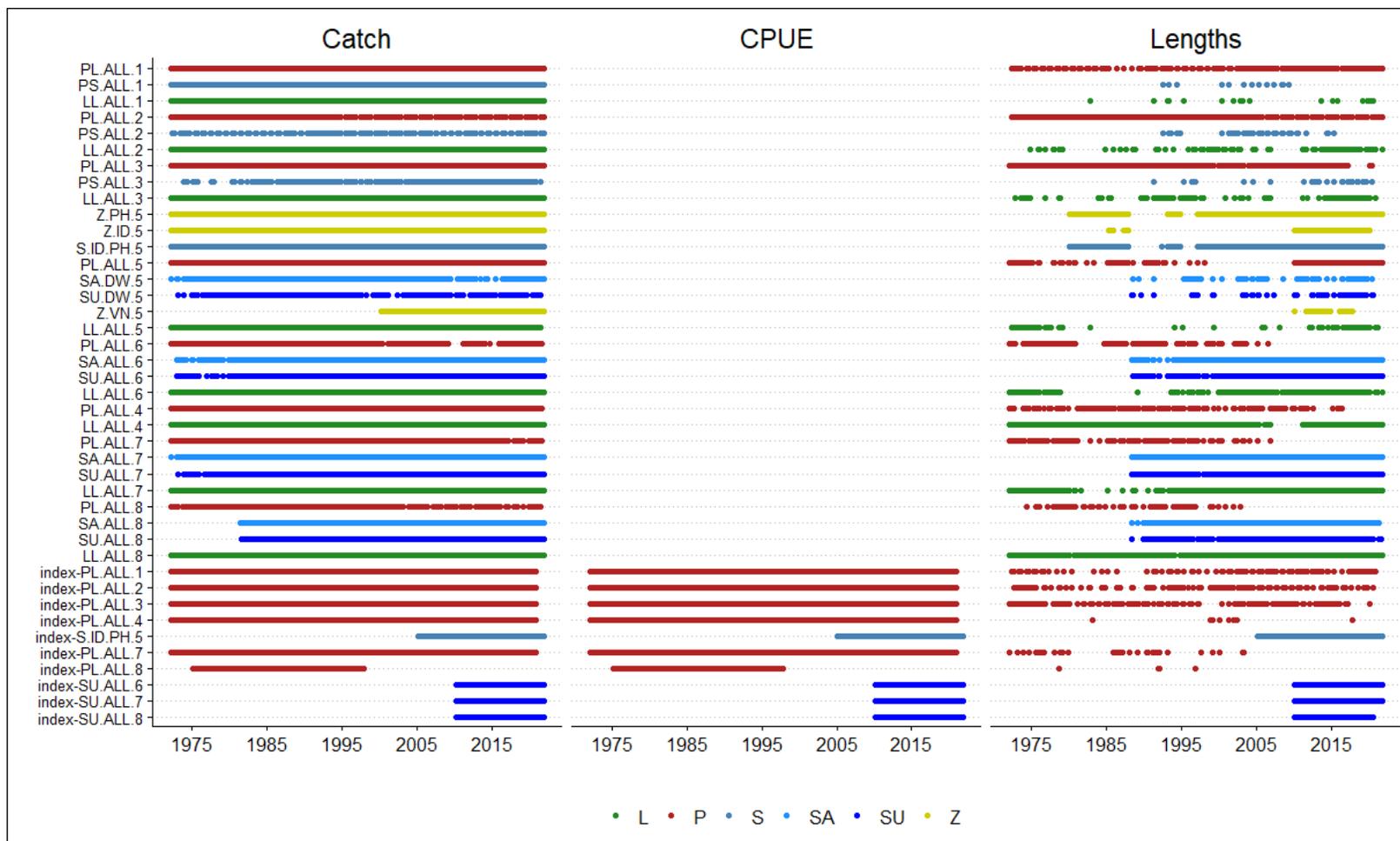


Figure 6: Summary of data coverage by fishery for the WCPO 2022 skipjack assessment.

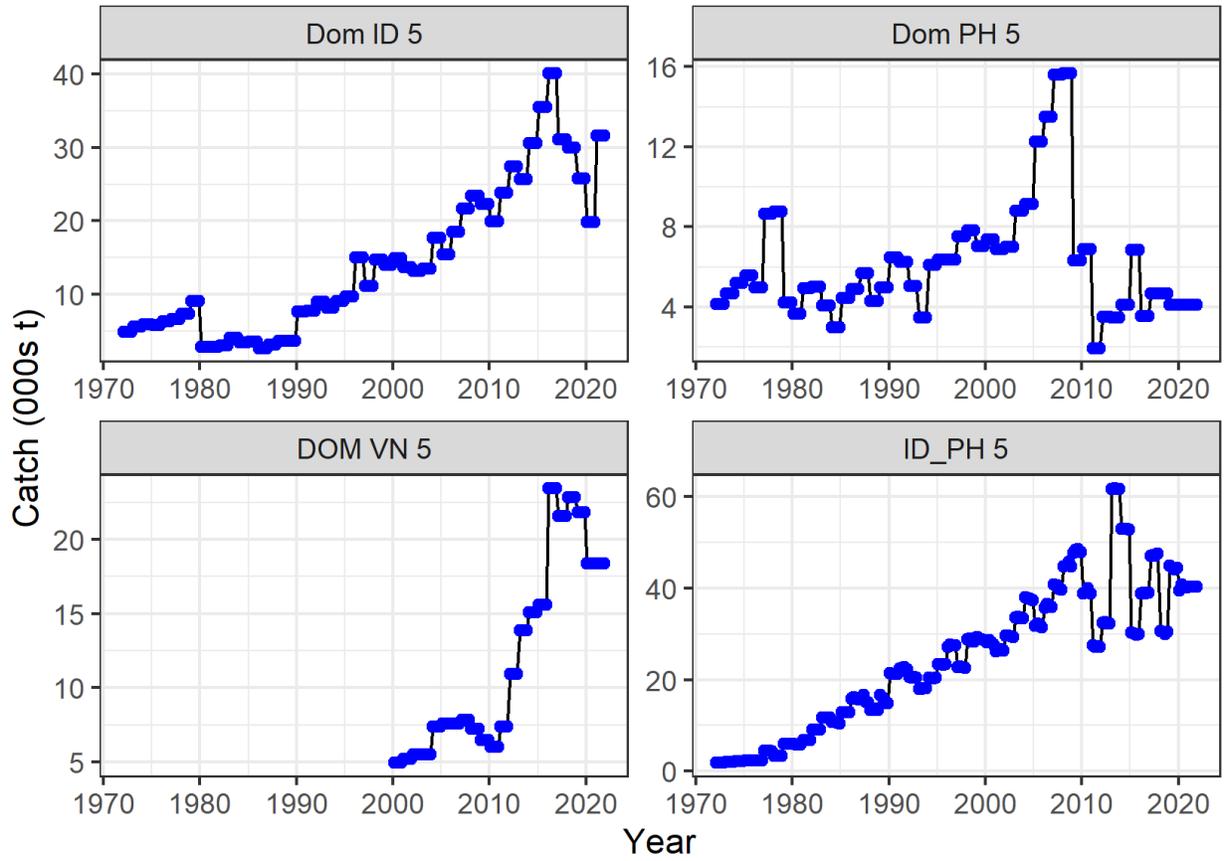


Figure 7: Time series of quarterly catches (mt) by fishery and regions: ID, VN and PH.

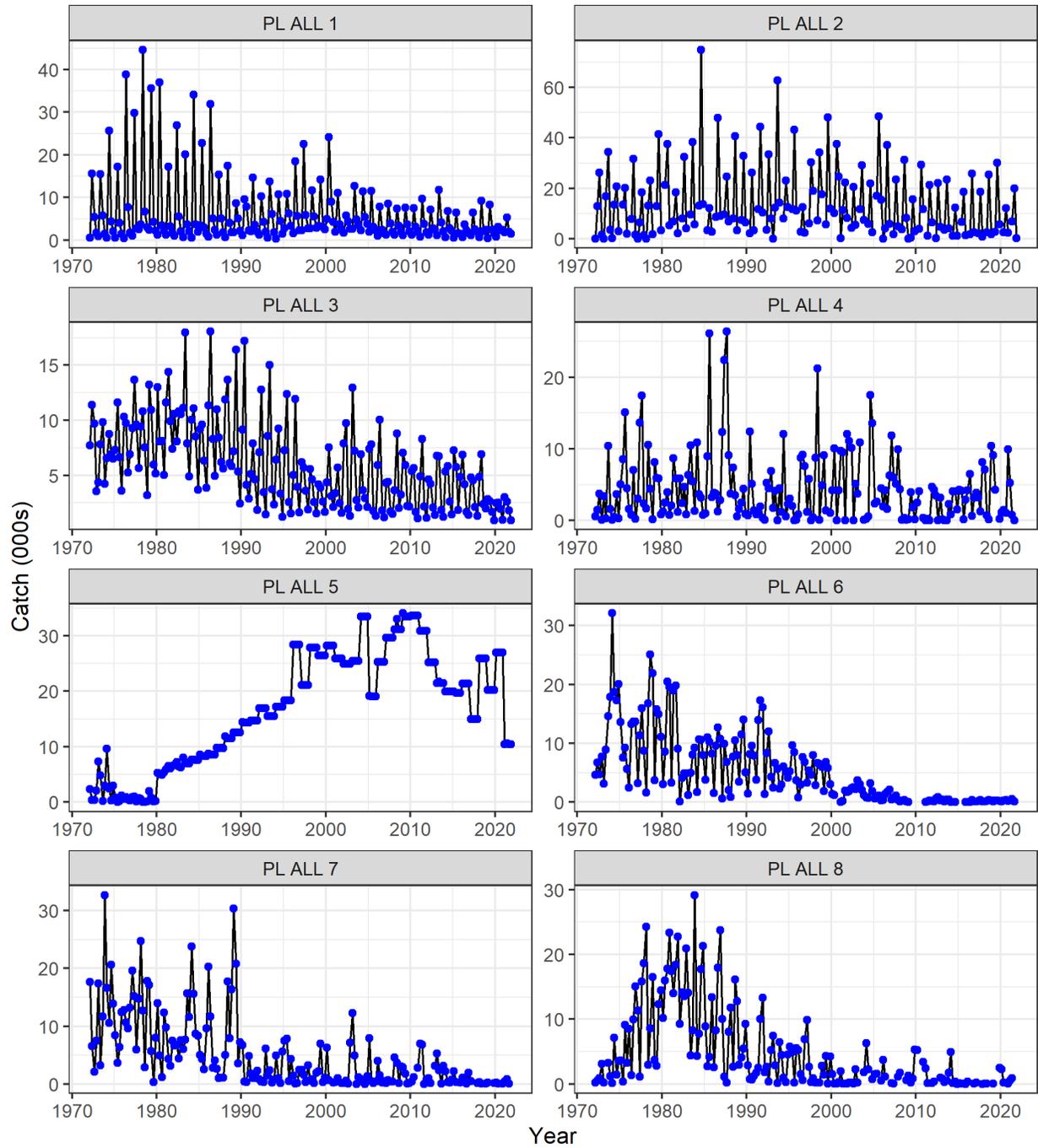


Figure 8: Time series of quarterly catches (mt) by fishery and regions: PL-all regions 1-8.

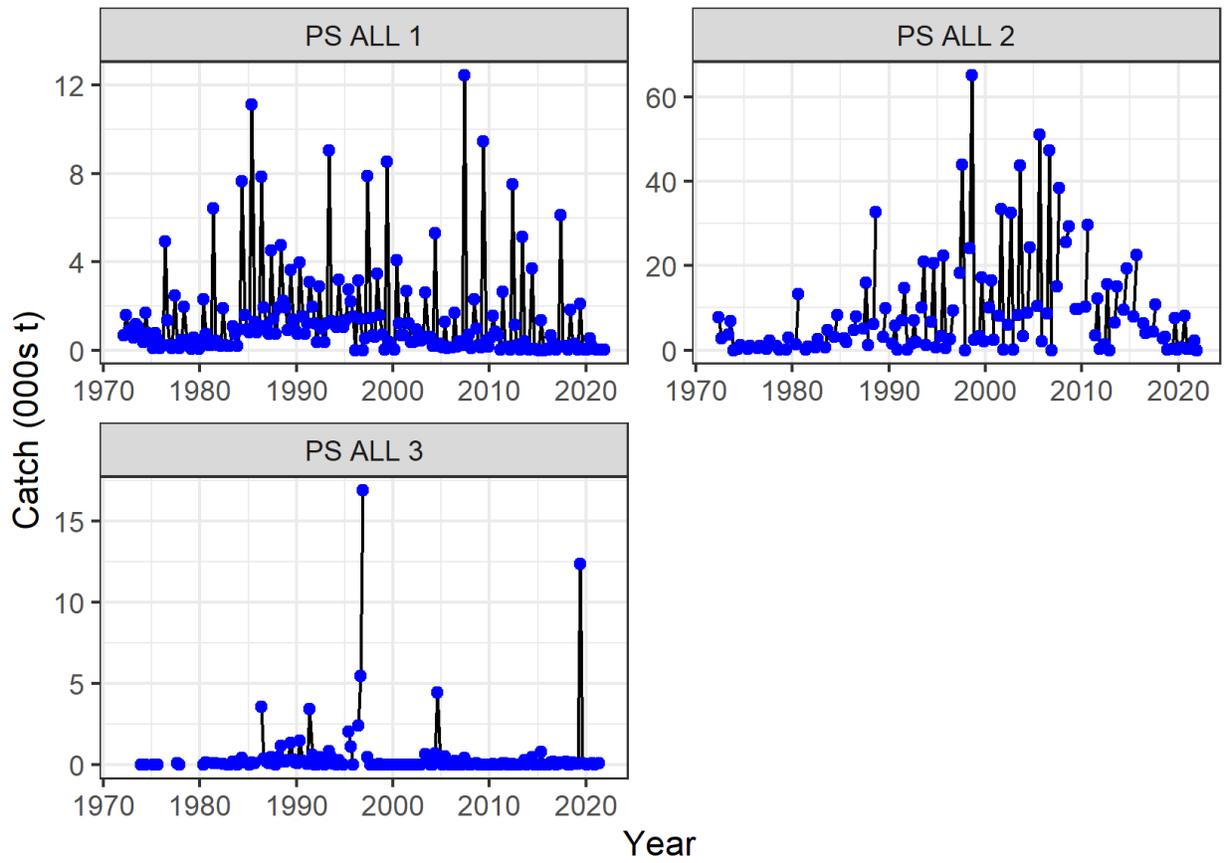


Figure 9: Time series of quarterly catches (mt) by fishery and regions: PS-all regions 1, 2, 3.

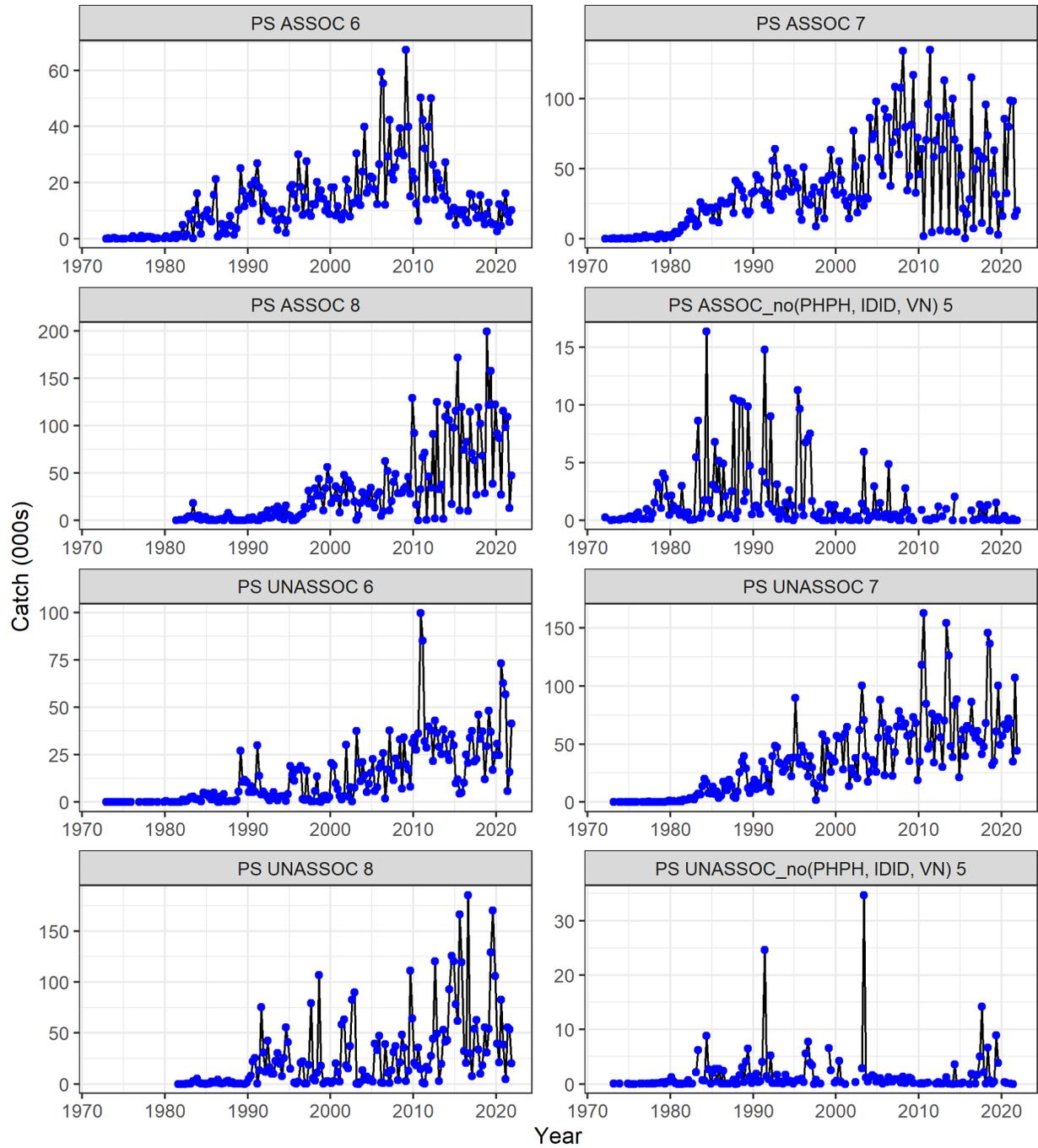


Figure 10: Time series of quarterly catches (mt) by fishery and regions: PS-associated and unassociated regions 5, 6, 7, 8

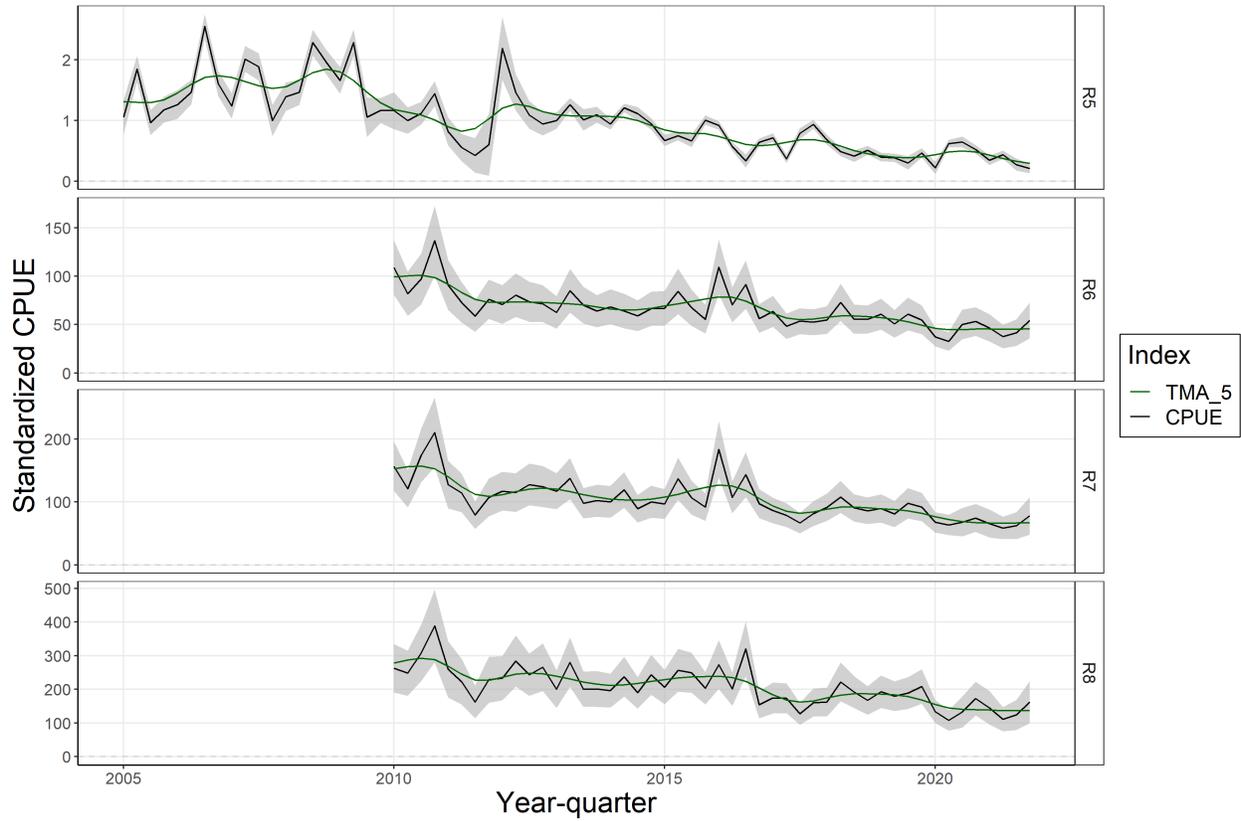


Figure 11: Standardised CPUE for the 'unassociated' purse seine survey fisheries in regions 6, 7 and 8, and the Philippines purse seine index in Region 5. Gray band is standard error (SE). Triangular moving average smoothing function applied to demonstrate overall trend (black line; smoothing window = 5)

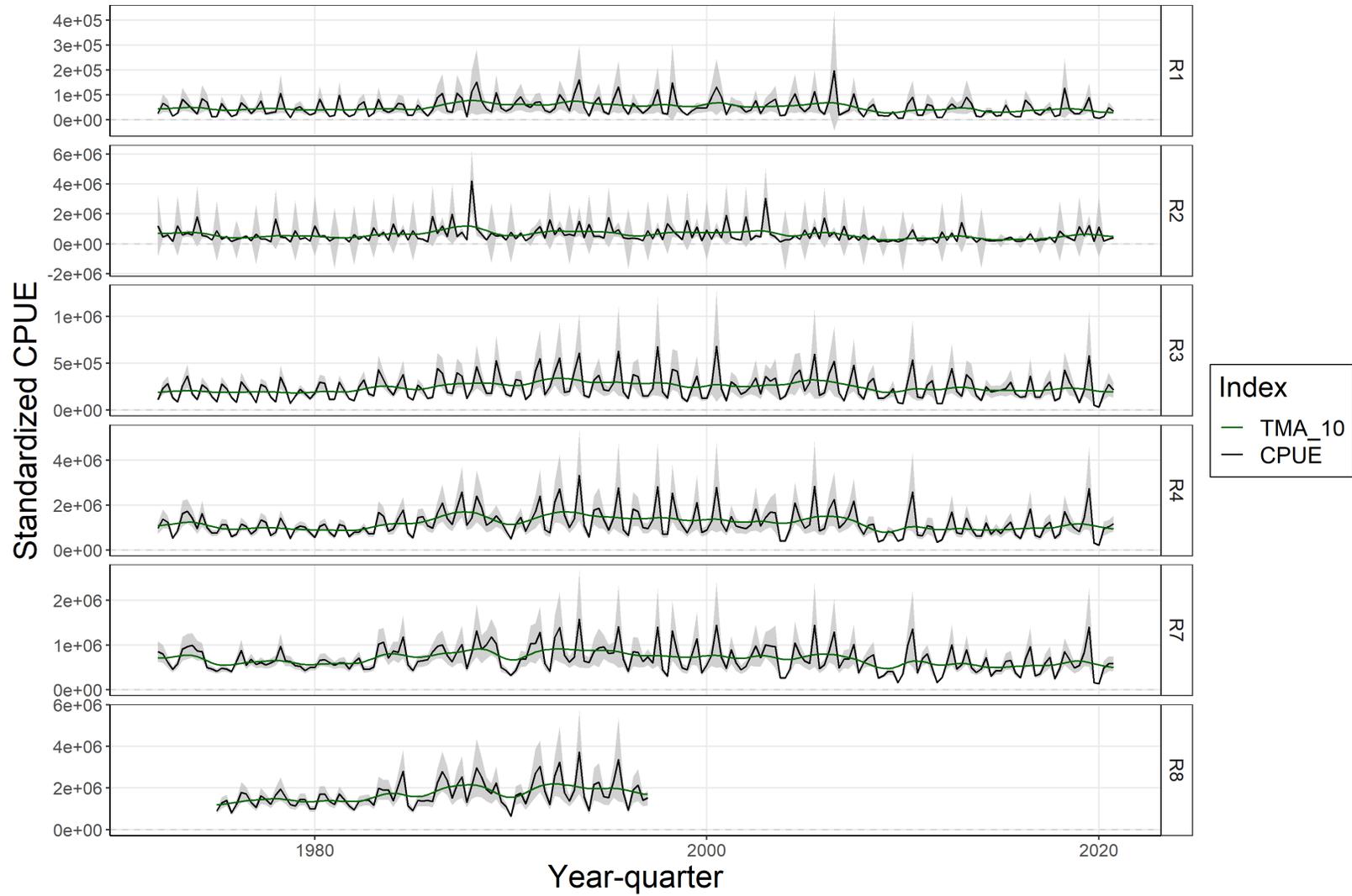


Figure 12: Standardised CPUE for the Japanese pole-and-line fisheries in regions 1, 2, 3, 4, 7, and 8. Gray band is standard error (SE). Triangular moving average smoothing function applied to demonstrate overall trend (black line; smoothing window = 5)

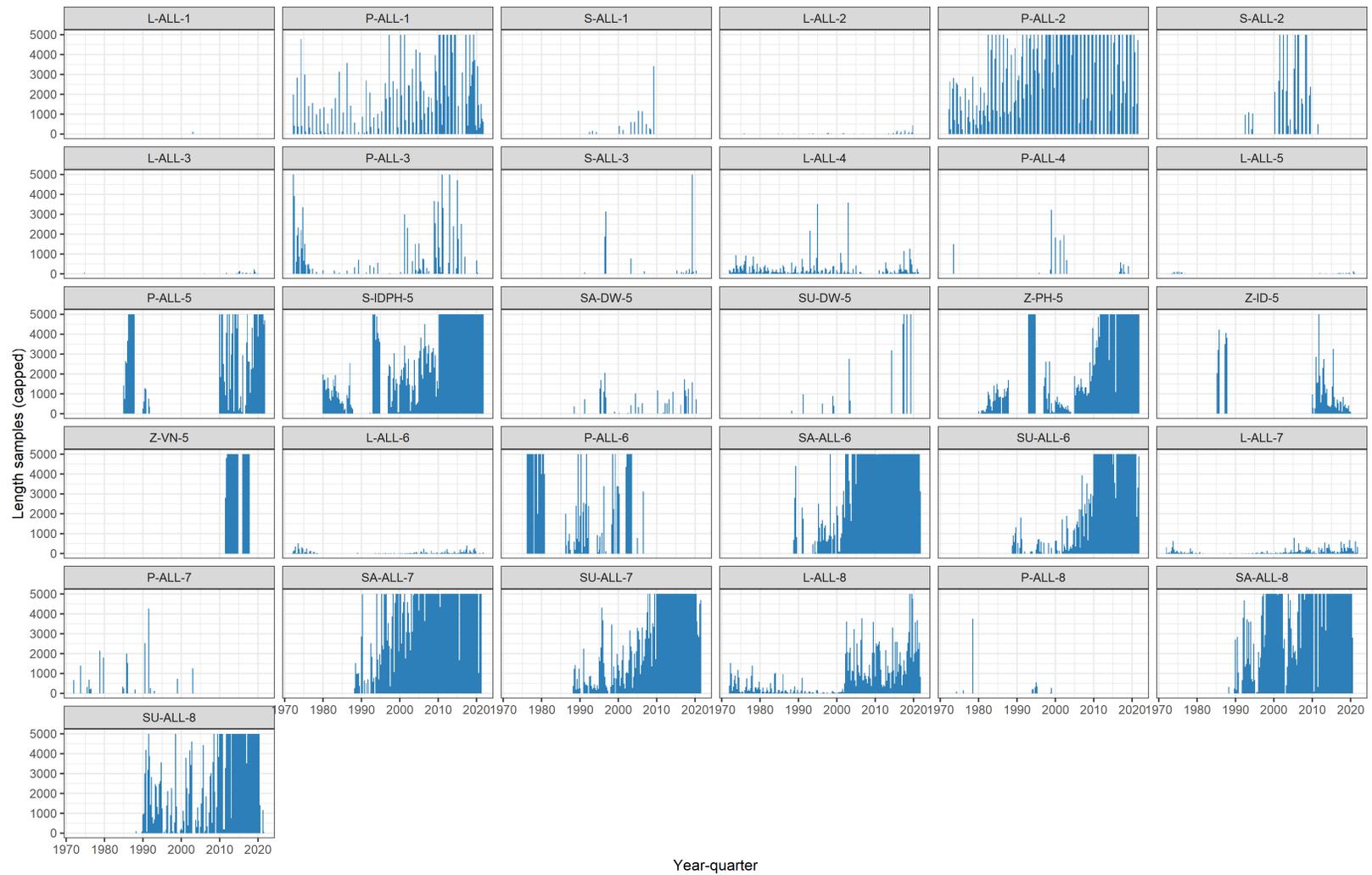


Figure 13: Plots of samples sizes (capped at 5,000) for length composition for each fishery in the model across the model period.

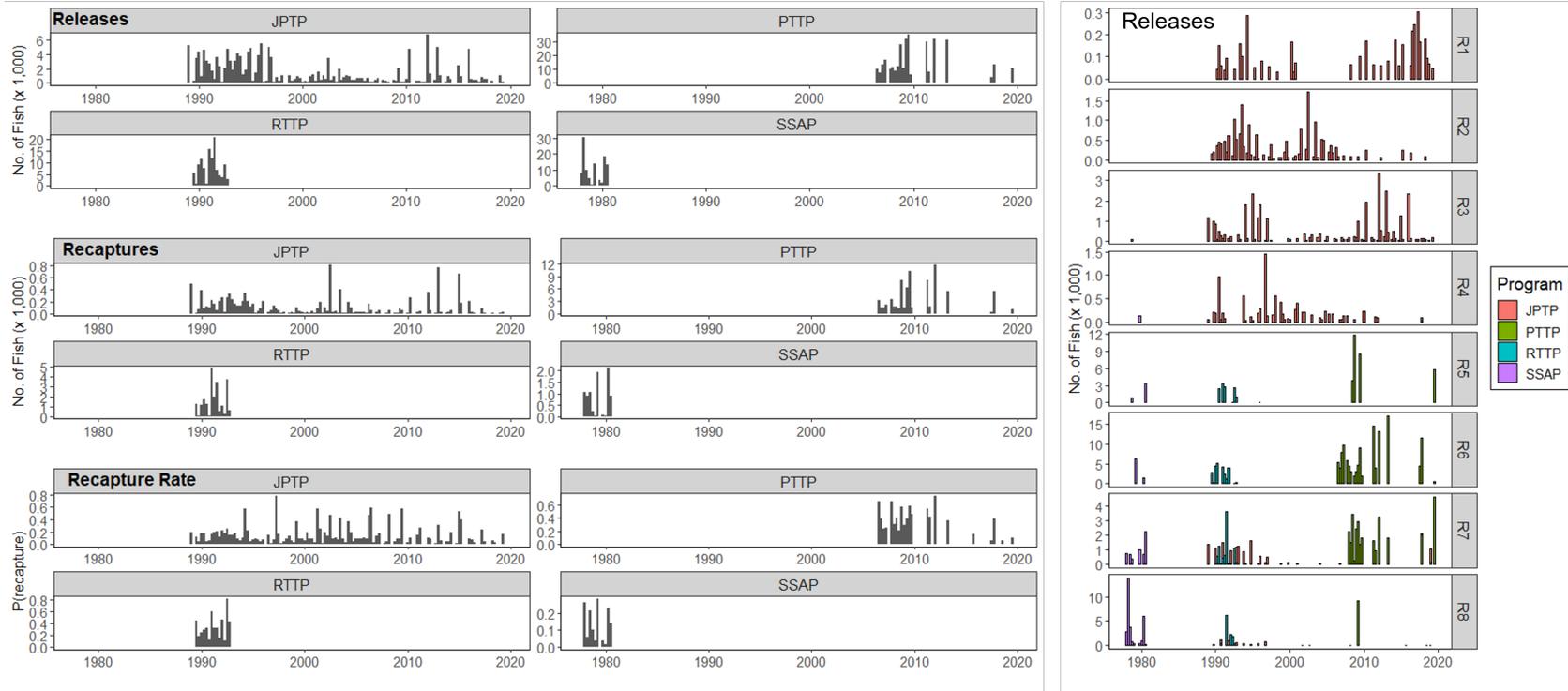


Figure 14: Summary plots of the number of releases, recaptures, and recapture rate of tags, by tagging program and region.

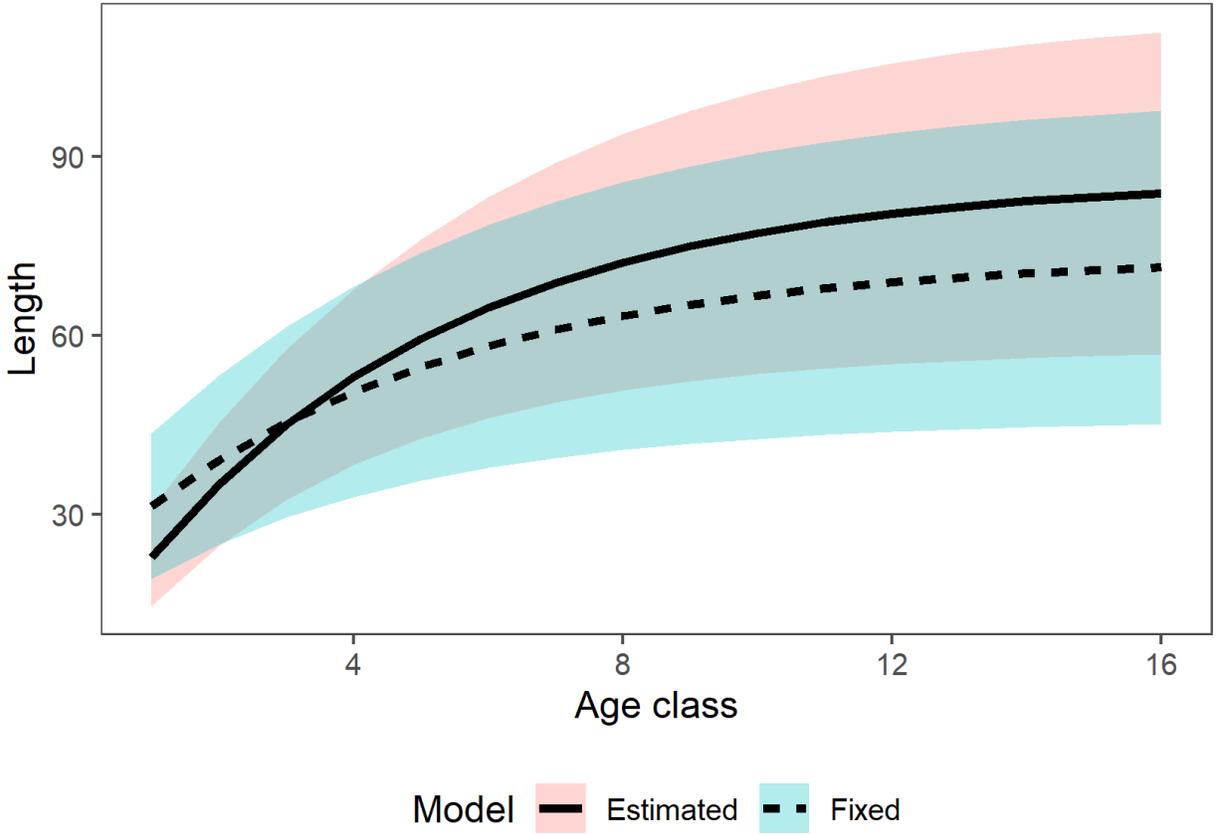


Figure 15: Growth curves estimated for skipjack. The diagnostic model curve was estimated internal (Estimated) to the model based on information on growth from the length frequency data and the use of the Dirichlet-multinomial to estimate effective sample size of length composition samples, the externally estimated curve (Fixed) was based on a combination of otolith daily aging and tag-recapture growth increment data. The parameters for the two growth curves are: Model Estimated $L_{inf} = 86.4$ cm FL, $K = 0.215$ and $t_0 = -0.422$, External Fixed $L_{inf} = 73.4$ cm FL, $K = 0.811$ and $t_0 = -0.607$

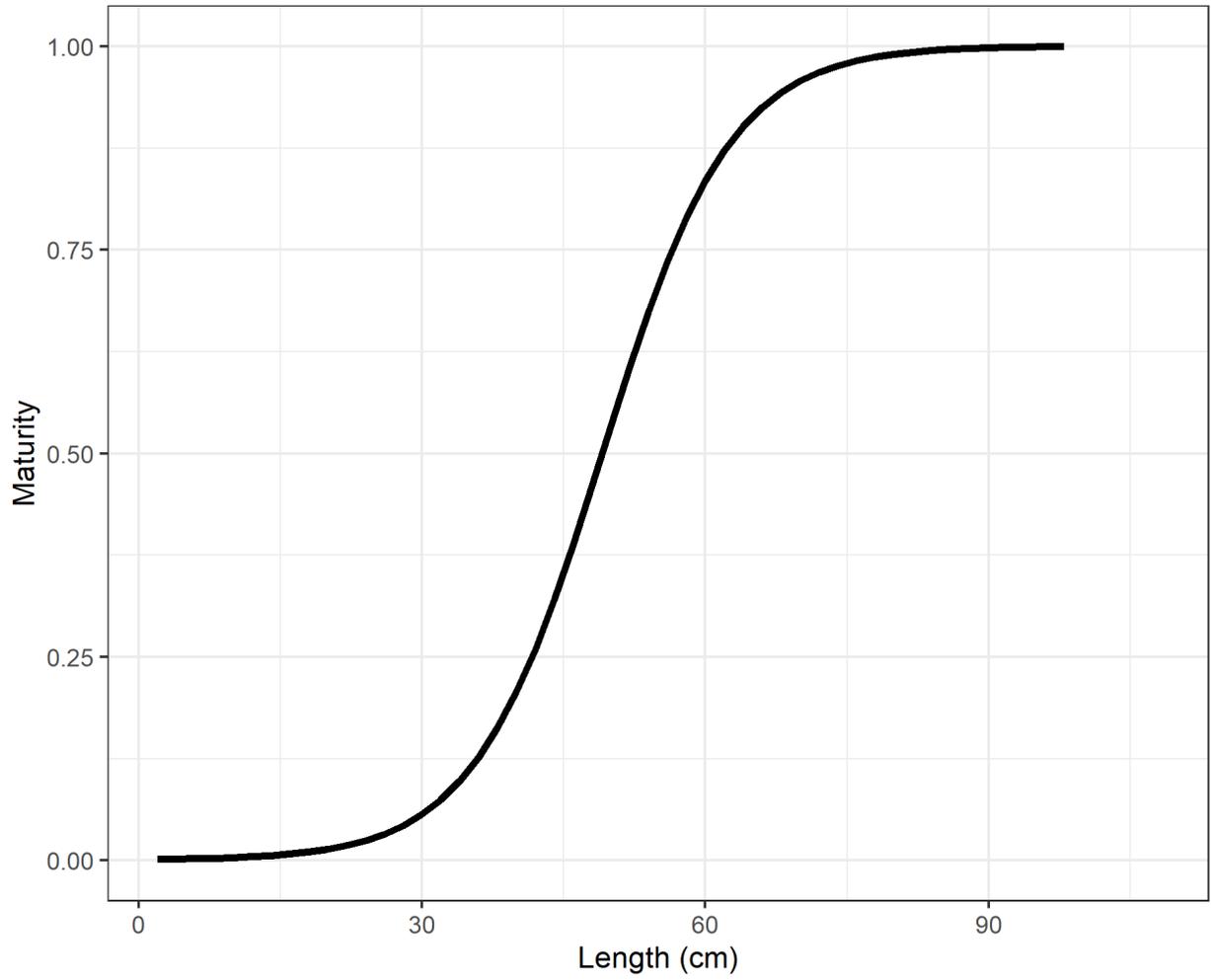


Figure 16: Maturity-at-length applied to all models in the assessment.

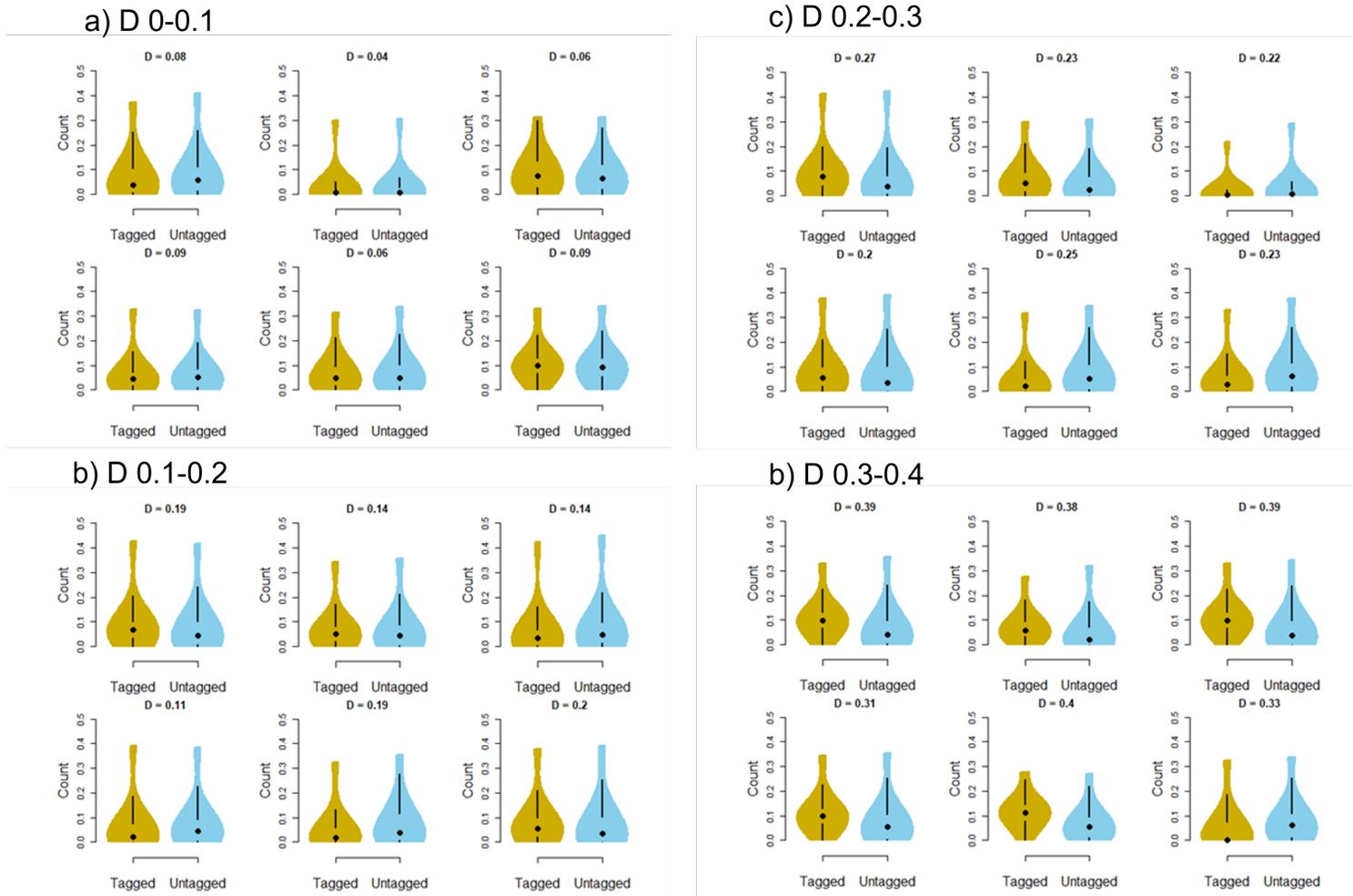


Figure 17: Example violin plots comparing distributions of estimated recapture probabilities between simulated tagged and untagged groups for different Komolgorov Smirnov (KS) D statistic values.

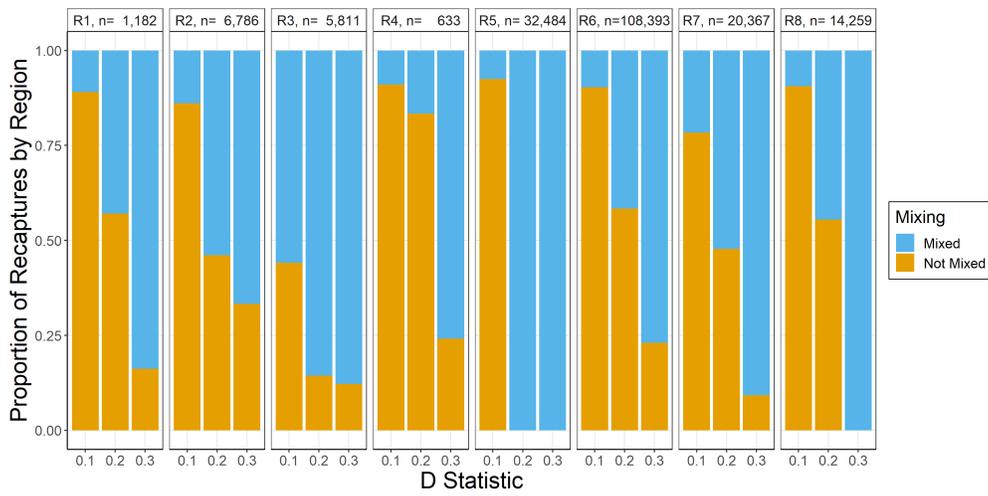
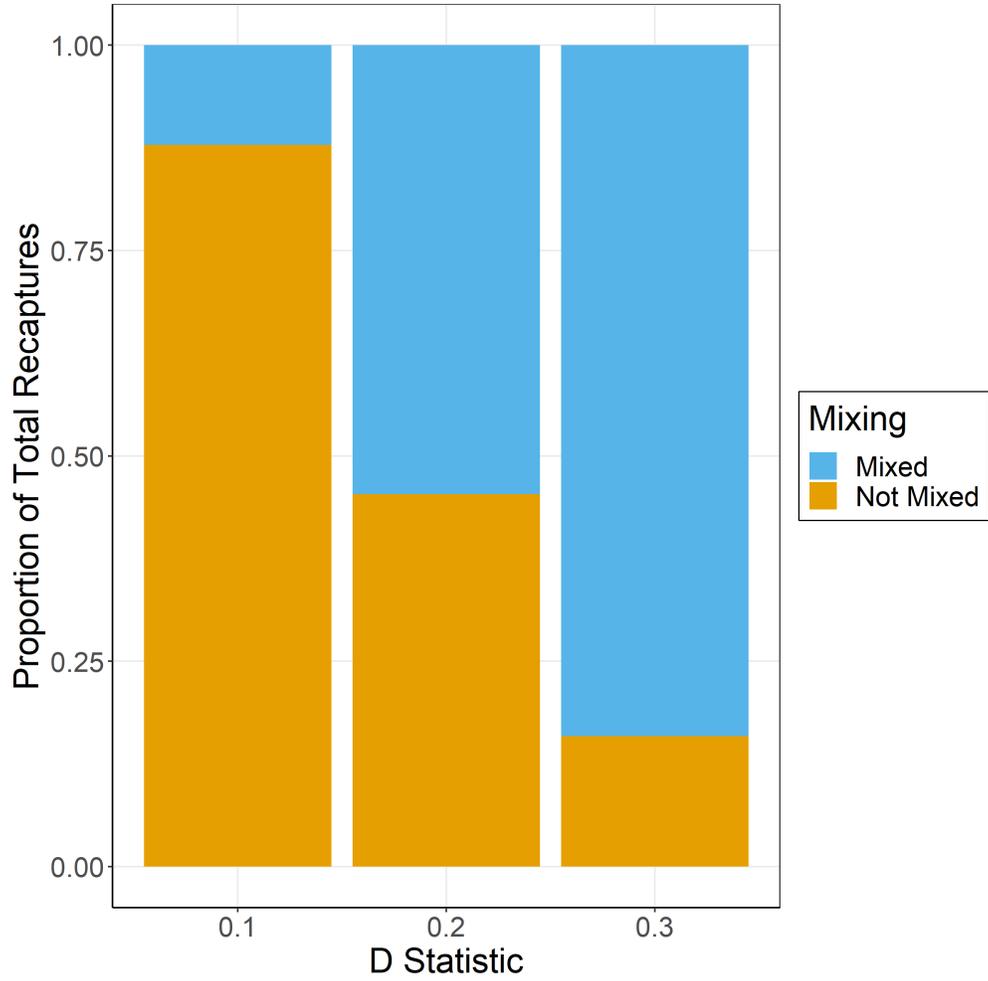


Figure 18: Proportions of tag recaptures considered as mixed or not mixed under the alternative KS D values. Top - all model regions combined, bottom - by model region with total numbers of recaptures indicated at top of each plot.

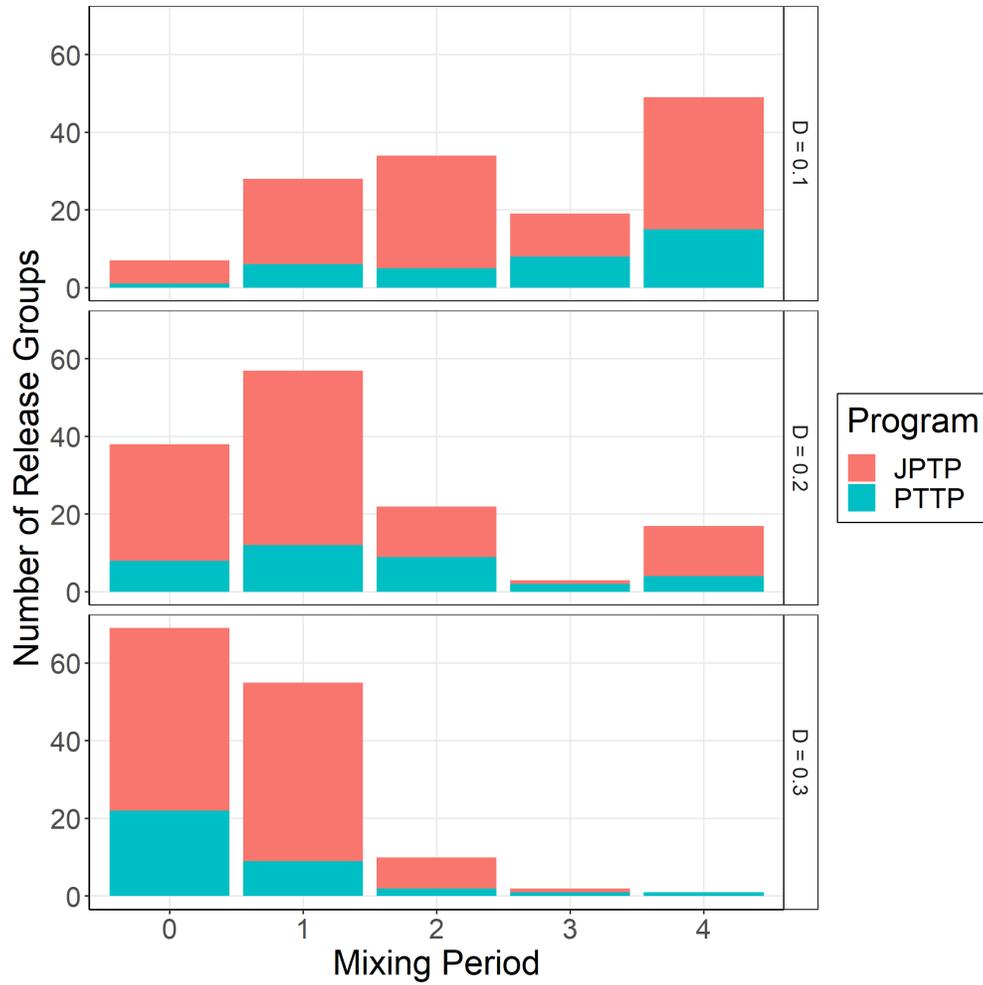


Figure 19: Distributions of mixing periods determined for MFCL tag release groups based on the tag mixing simulations for the different KS D statistic values used in the structural uncertainty grid.

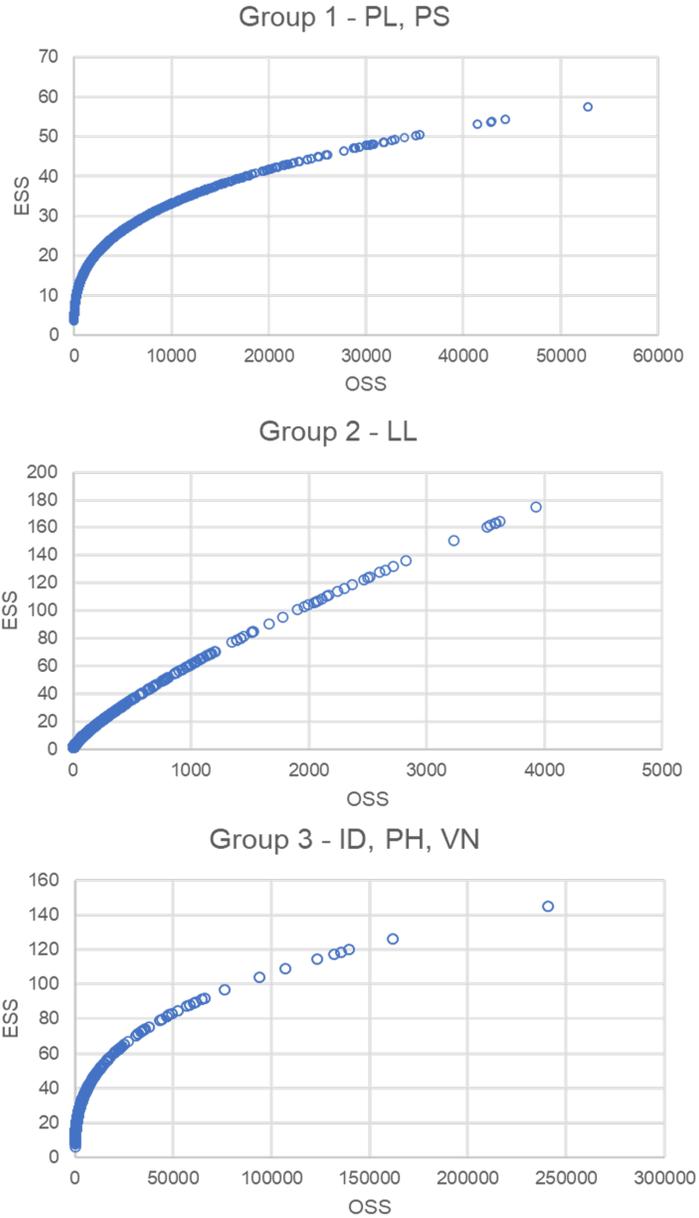


Figure 20: Plots of estimated effective sample (ESS) sizes versus observed sample sizes (OSS) from the MULTIFAN-CL implementation of Dirichlet-multinomial likelihood for the diagnostic model.

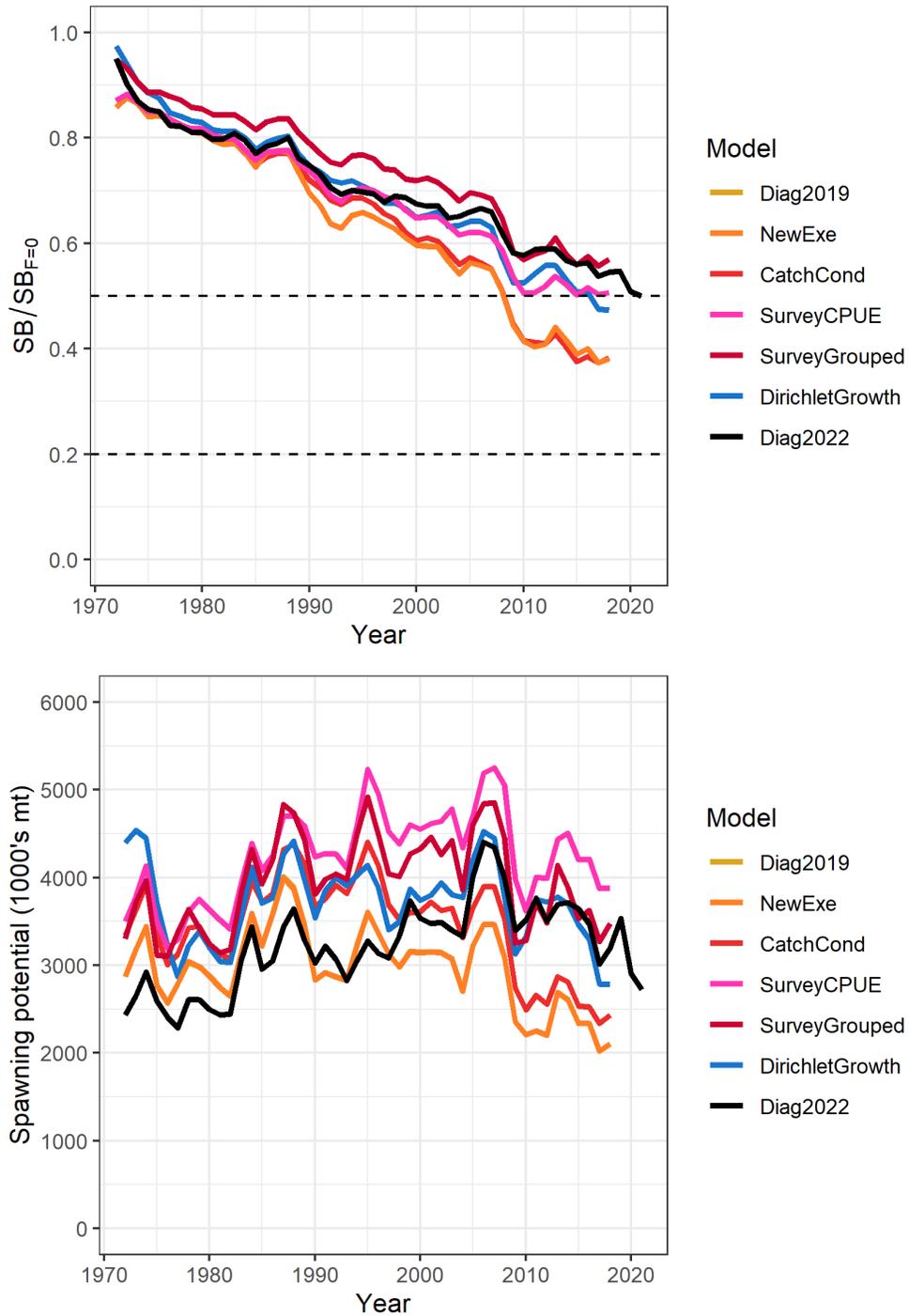


Figure 21: Estimated dynamic spawning depletion (Top) and spawning potential (Bottom) trajectories for each of the stepwise model runs, Diag2022 is the 2022 diagnostic model (refer to Section 8.1).

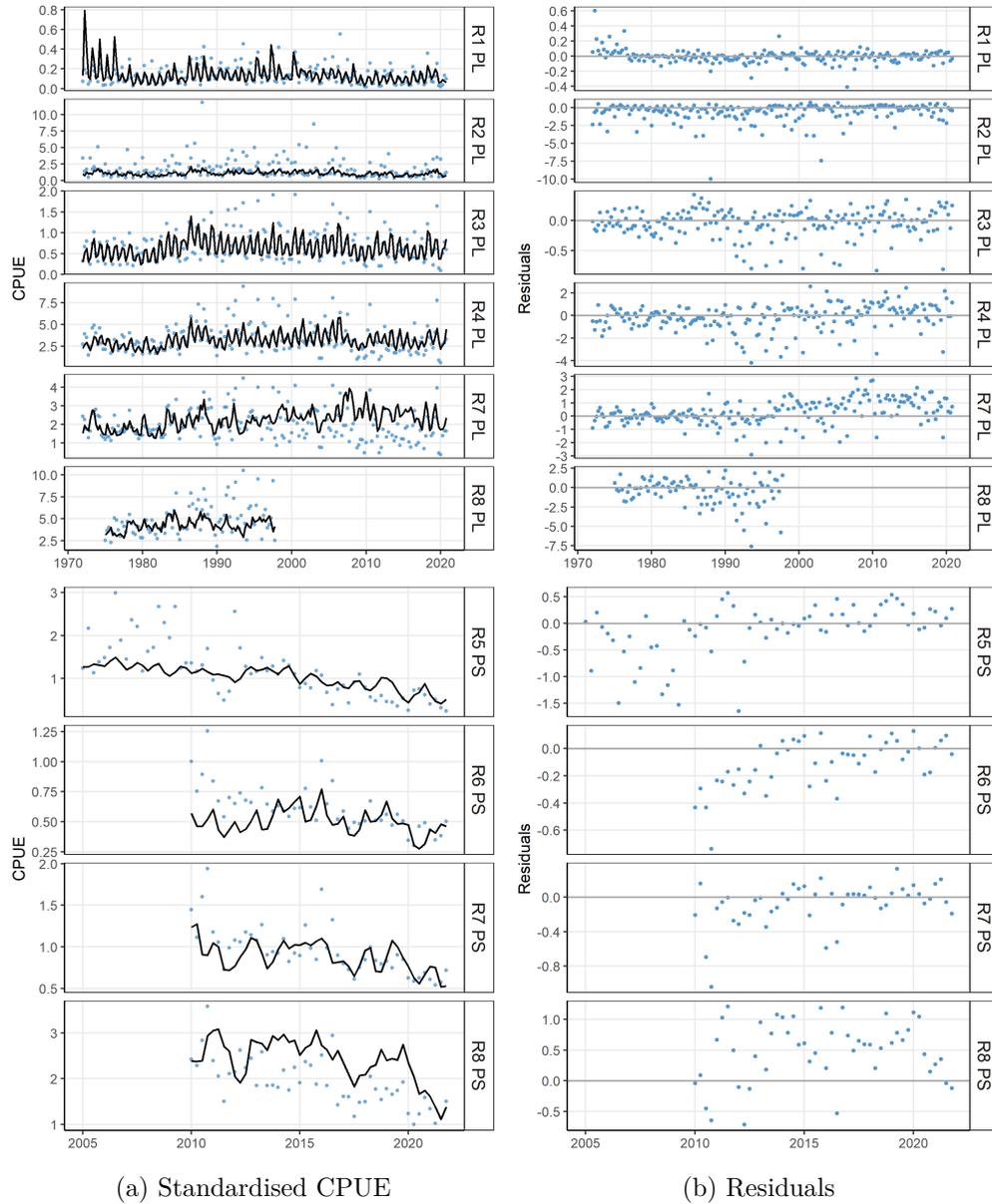


Figure 22: a) Comparison of quarterly model estimated (black line) and observed standardised CPUE (blue dots) for the pole-and-line (Top) and purse seine (Bottom) survey fisheries for the 2022 diagnostic model. b) Plots of residuals between estimated and observed standardised CPUE for the survey fisheries, (Top) pole-and-line, (Bottom) purse seine.

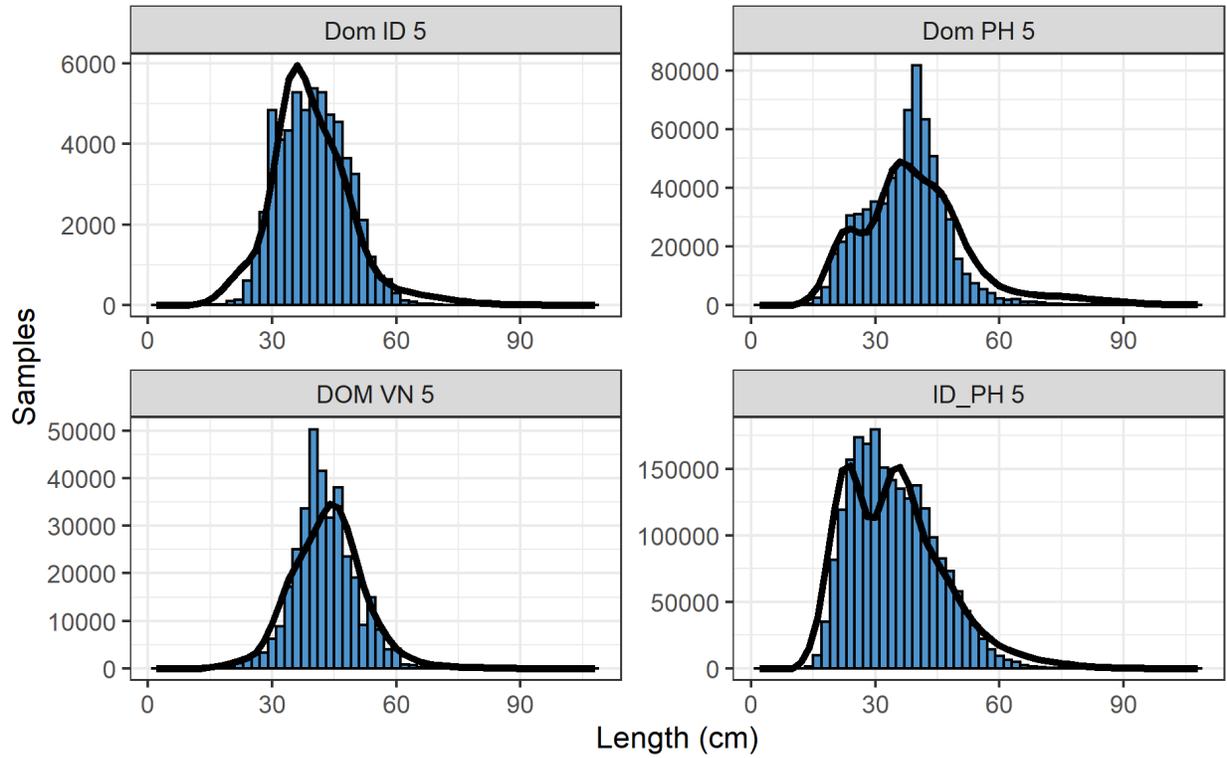


Figure 23: Composite (all time periods combined) observed (blue histograms) and predicted (black line) catch-at-length for ID, VN and PH domestic fisheries for the diagnostic model.

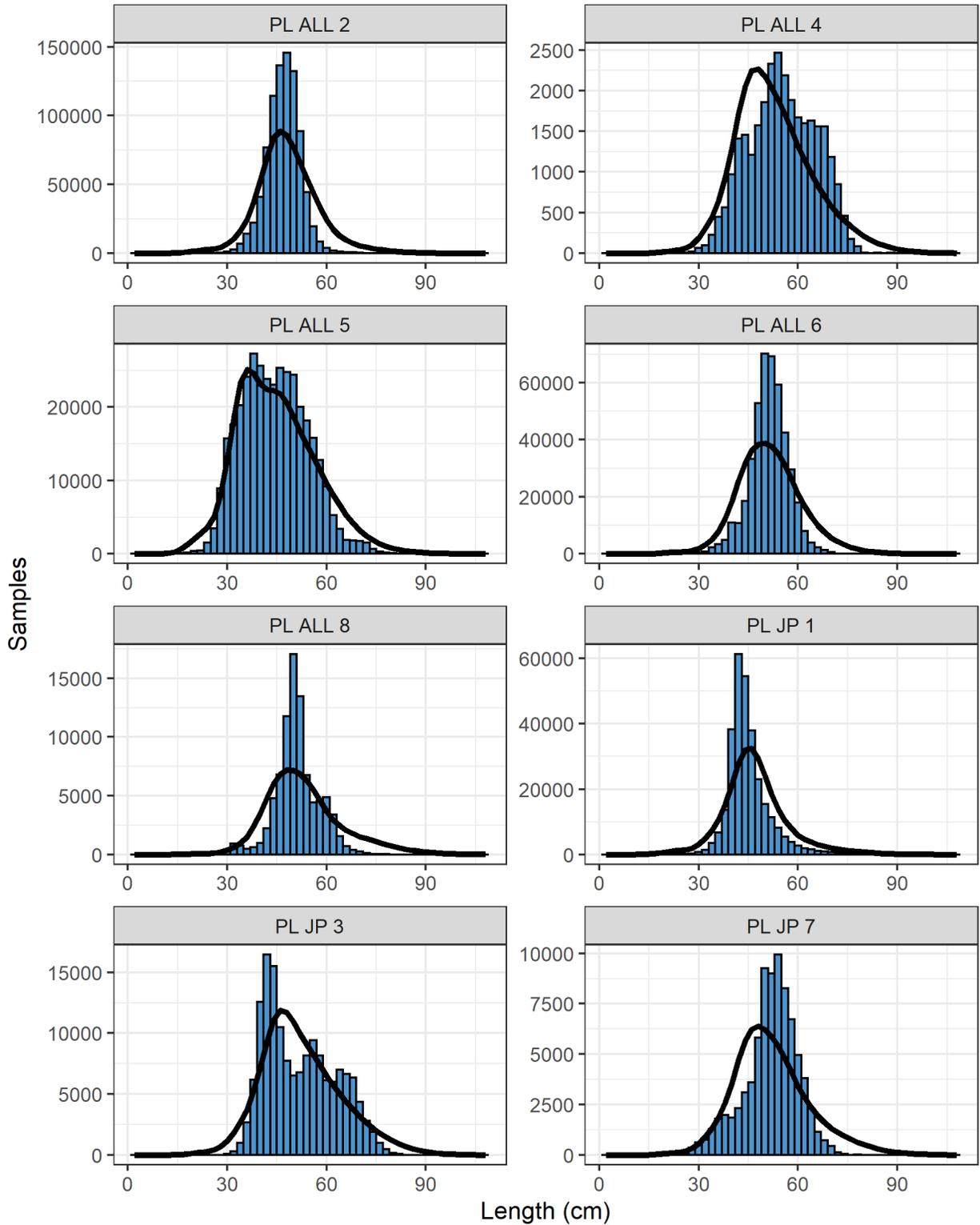


Figure 24: Composite (all time periods combined) observed (blue histograms) and predicted (black line) catch-at-length for pole-and-line fisheries for the diagnostic model.

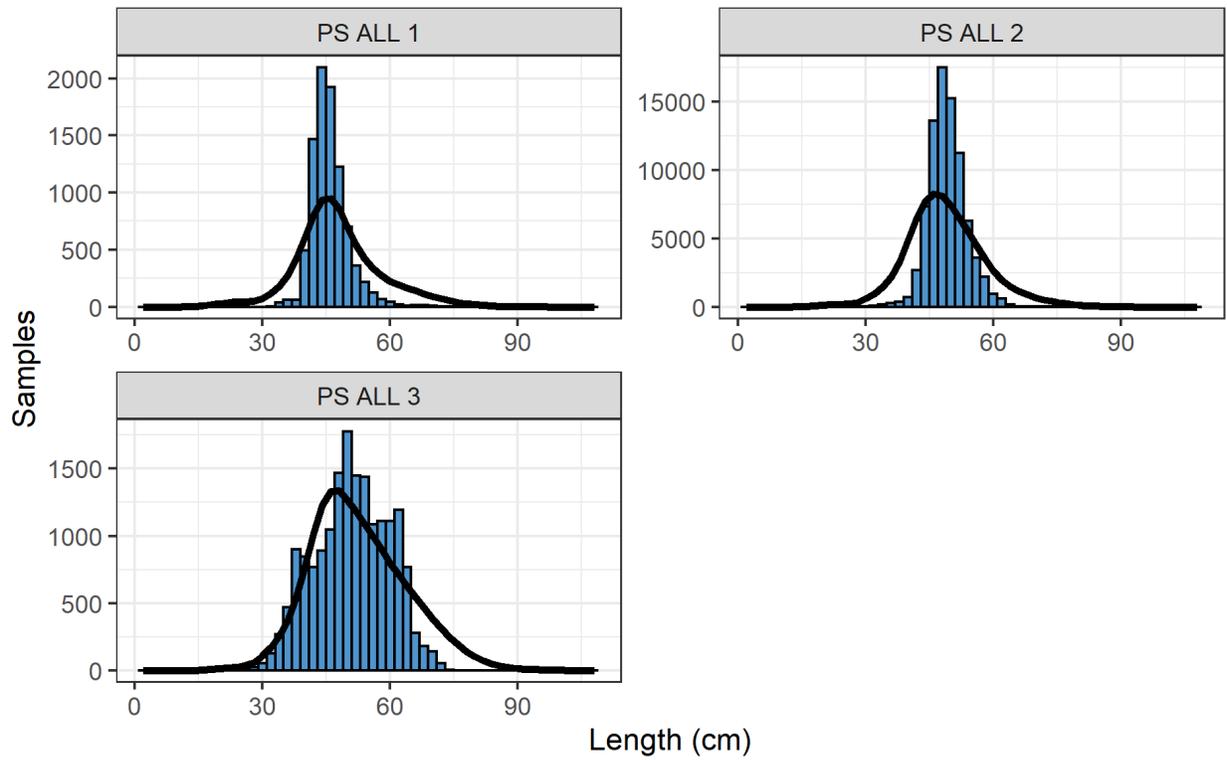


Figure 25: Composite (all time periods combined) observed (blue histograms) and predicted (black line) catch-at-length for purse seine fisheries in regions 1, 2, and 3 for the diagnostic model.

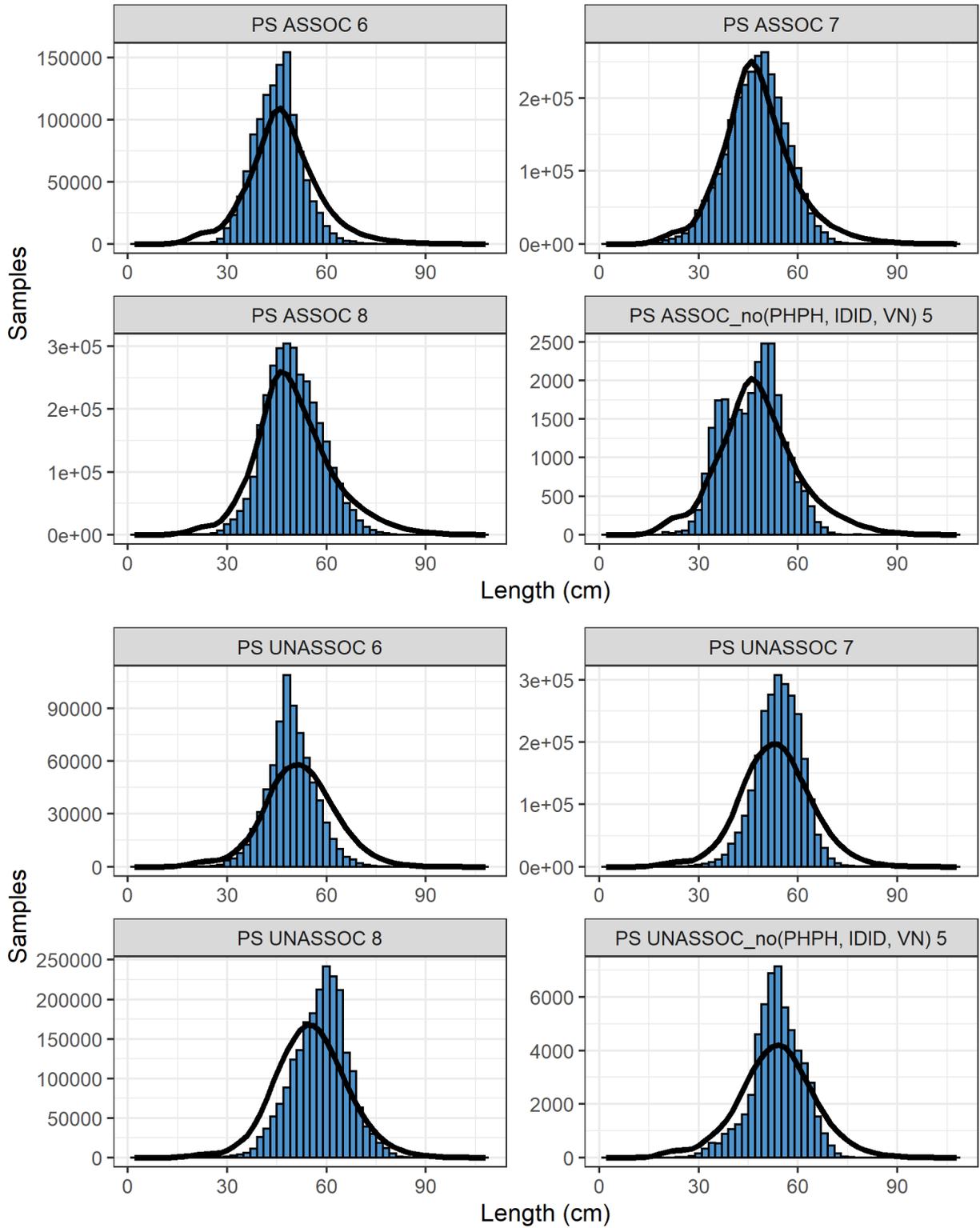


Figure 26: Composite (all time periods combined) observed (blue histograms) and predicted (black line) catch-at-length for purse seine fisheries, associated (top) and unassociated (bottom) in regions 5, 6, 7, and 8 for the diagnostic model.

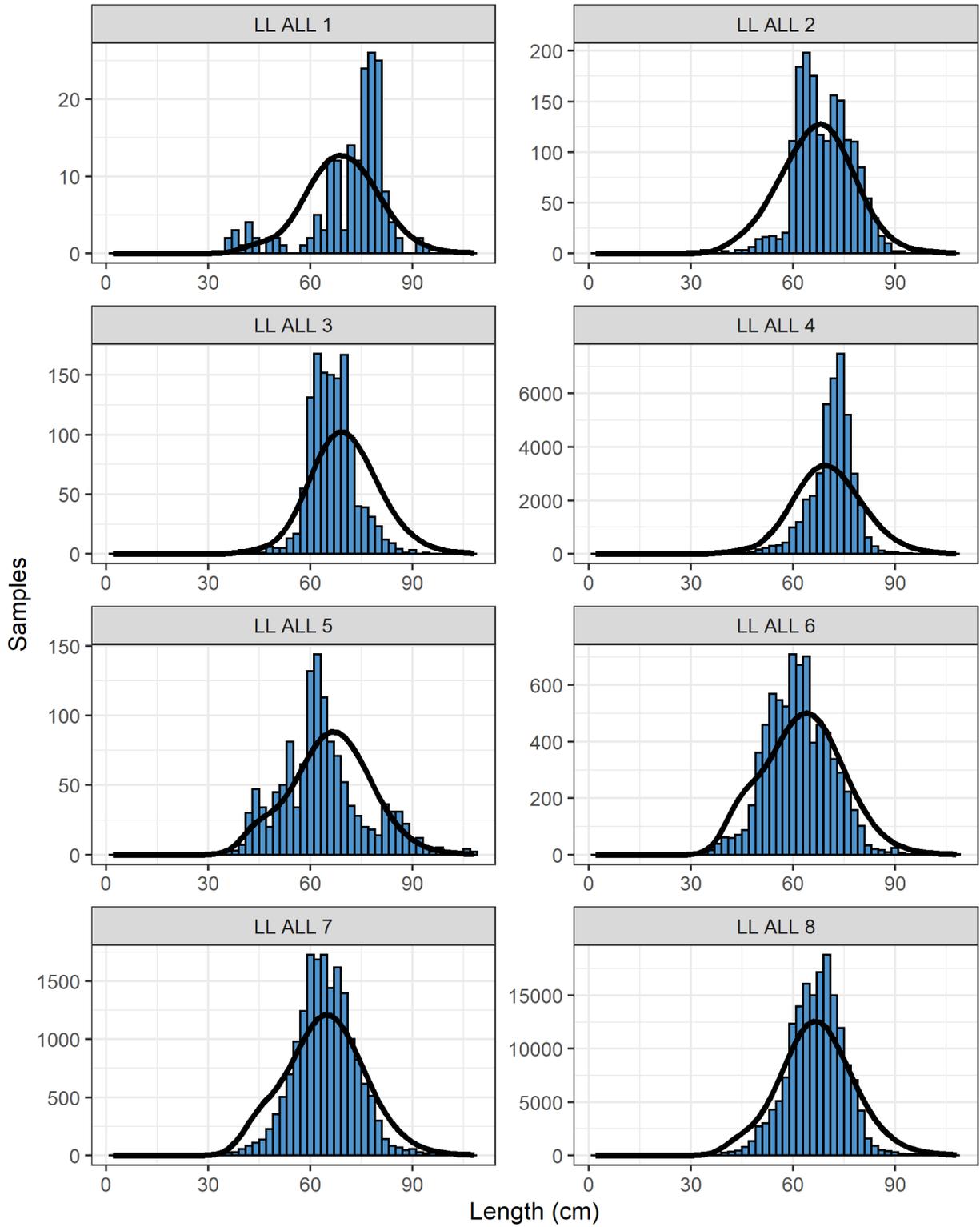


Figure 27: Composite (all time periods combined) observed (blue histograms) and predicted (black line) catch-at-length for longline fisheries for the diagnostic model.

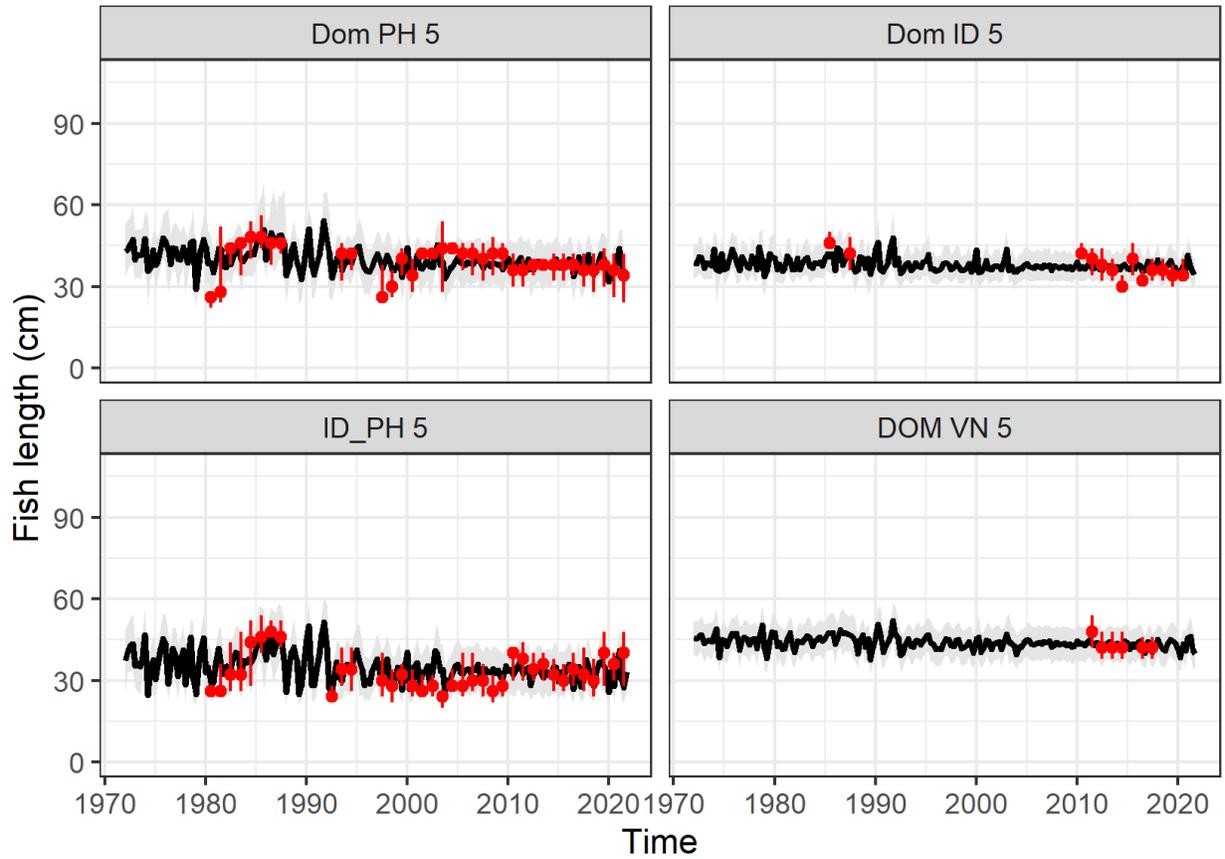


Figure 28: A comparison of the observed (red points) and predicted (black line) median fish length (FL, cm) for the ID, VN, and PH fisheries for the diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples more than 30 fish per quarter are plotted

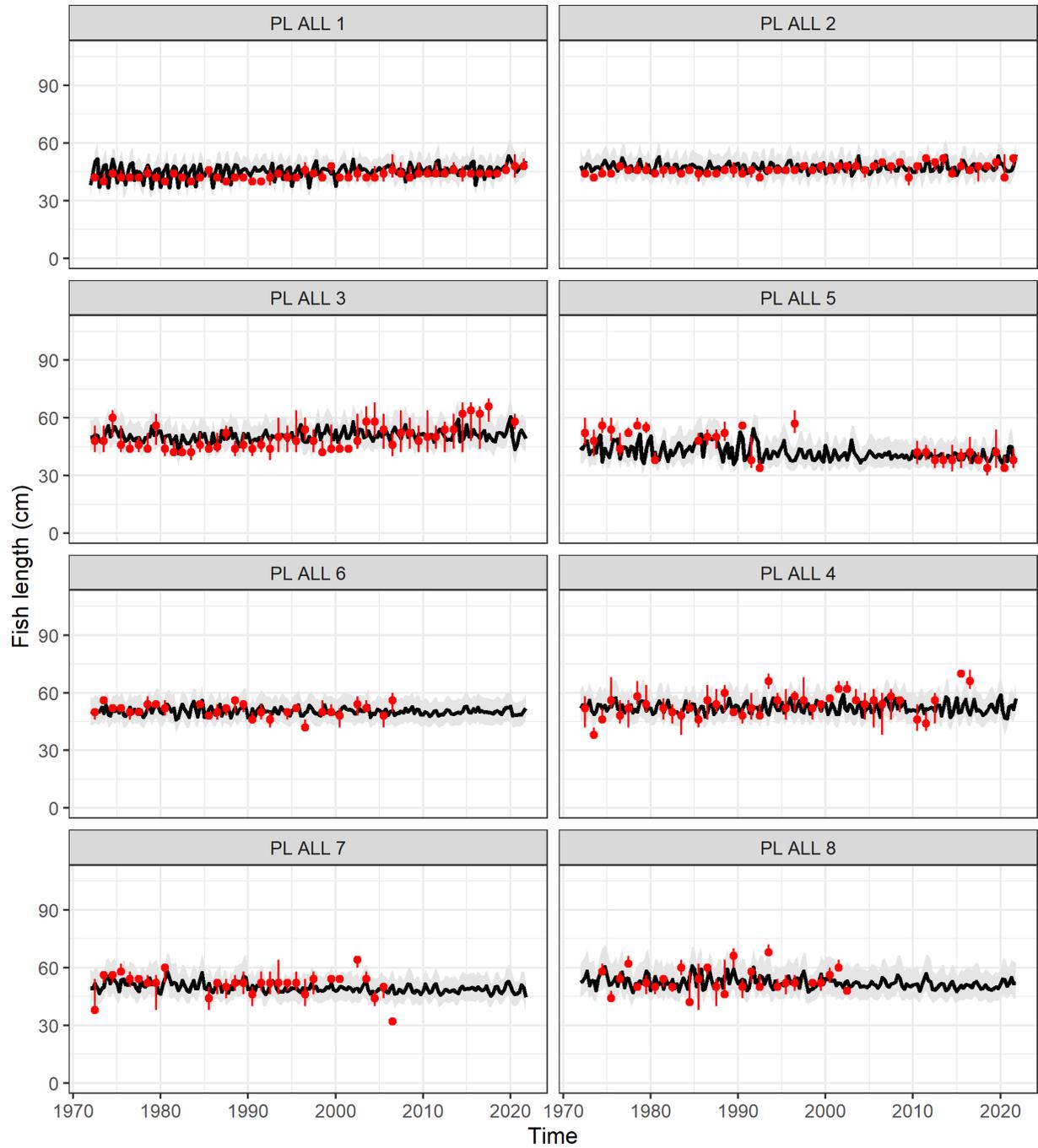


Figure 29: A comparison of the observed (red points) and predicted (black line) median fish length (FL, cm) for pole-and-line fisheries for the diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples more than 30 fish per quarter are plotted

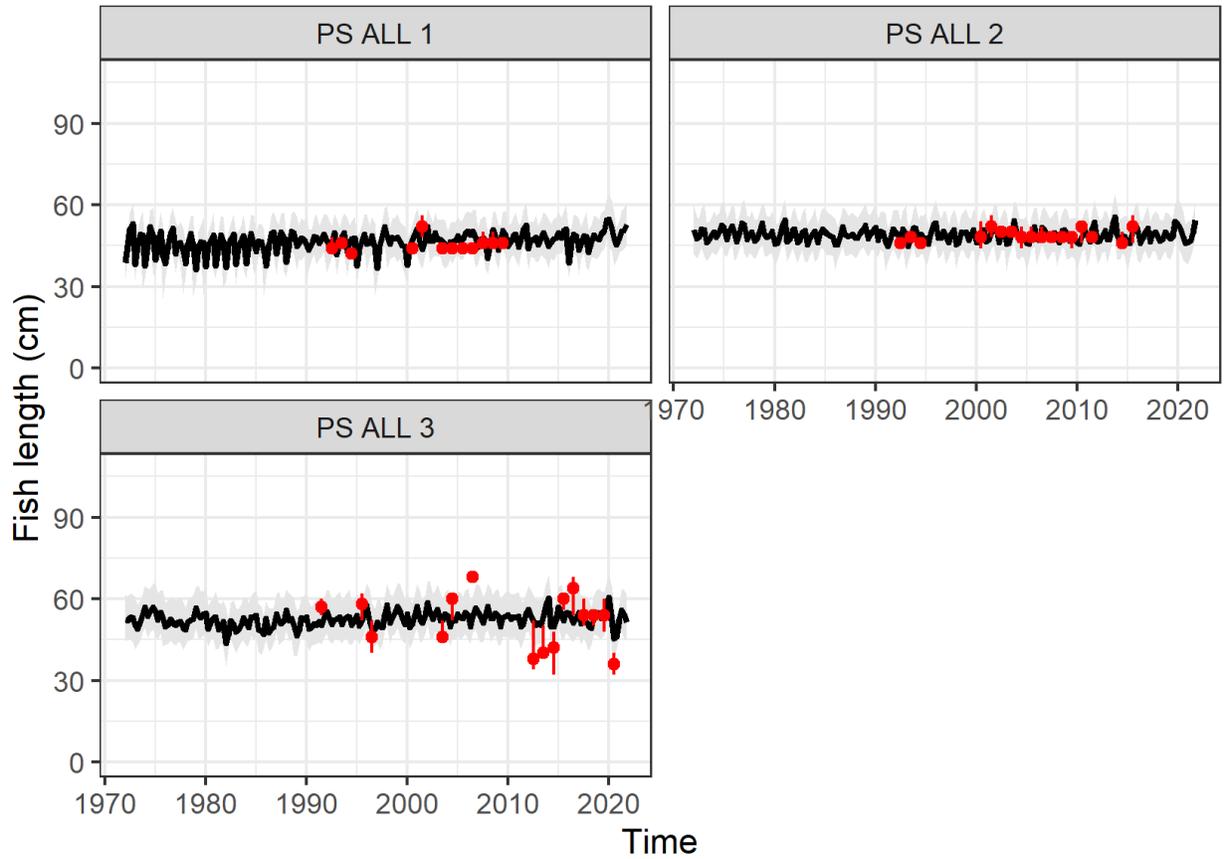


Figure 30: A comparison of the observed (red points) and predicted (black line) median fish length (FL, cm) for purse seine fisheries in regions 1,2,3 for the diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples more than 30 fish per quarter are plotted

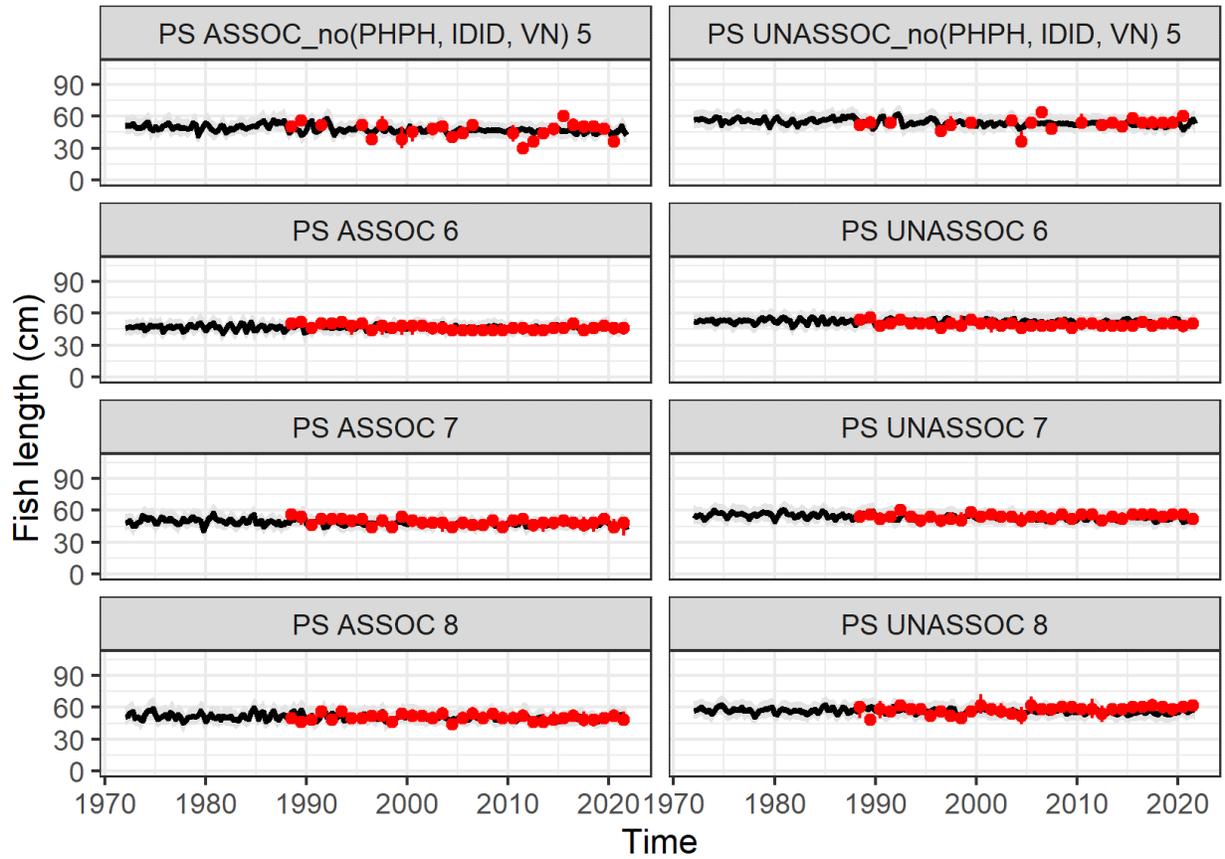


Figure 31: A comparison of the observed (red points) and predicted (black line) median fish length (FL, cm) for purse seine fisheries 'associated' and 'unassociated' regions 5,6,7,8 for the diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples more than 30 fish per quarter are plotted

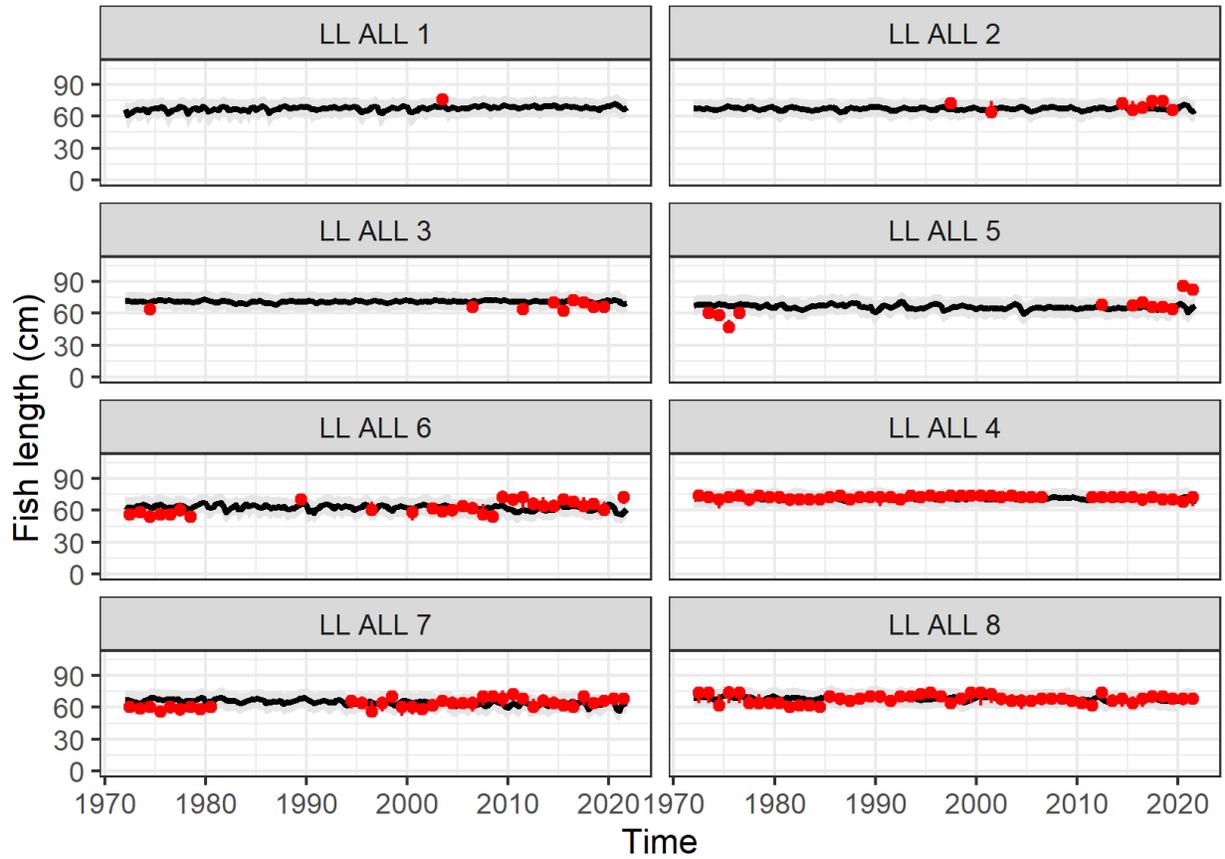


Figure 32: A comparison of the observed (red points) and predicted (black line) median fish length (FL, cm) for longline fisheries for the diagnostic model. The uncertainty intervals (gray shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are by quarter and only length samples more than 30 fish per quarter are plotted

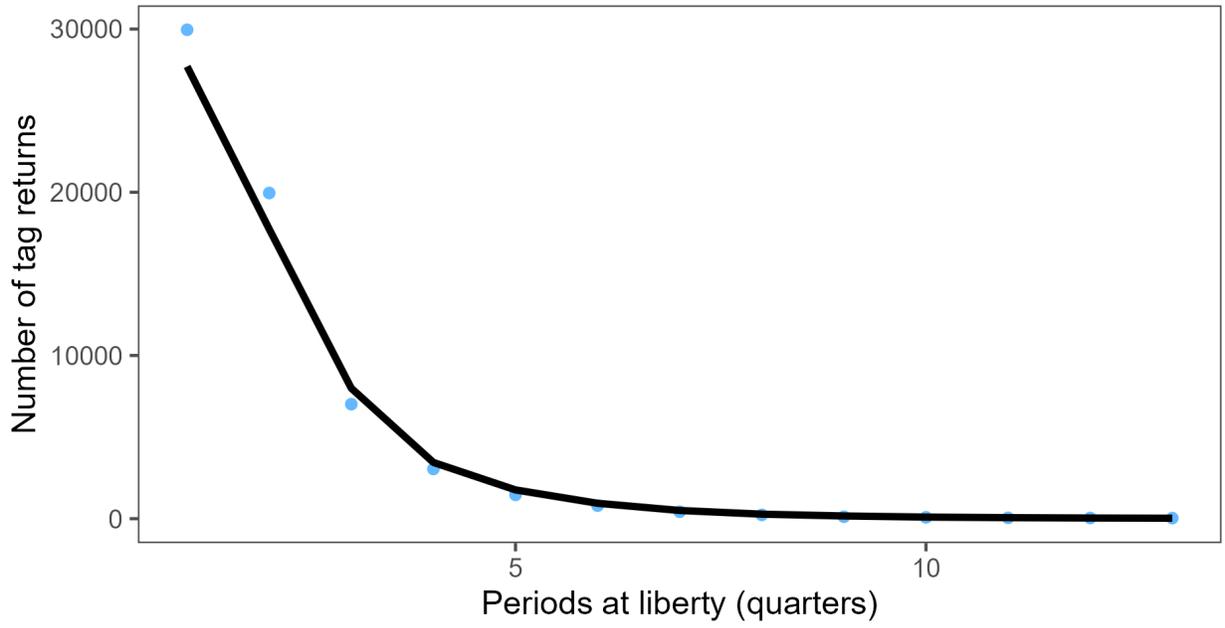


Figure 33: Observed (blue points) and model-predicted (black line) tag attrition across all tag release events for the diagnostic model.

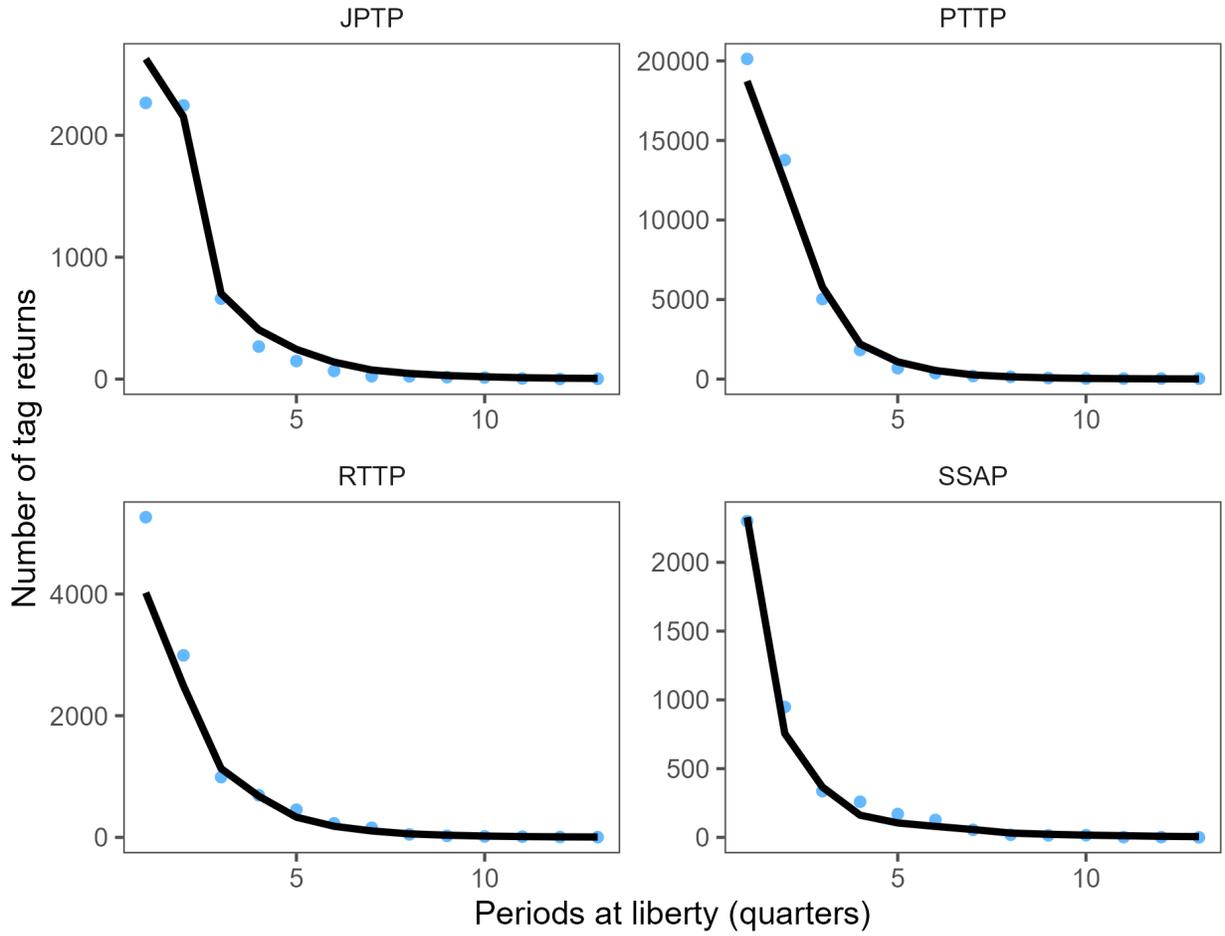


Figure 34: Observed (blue points) and model-predicted (black line) tag attrition by tagging programme for the diagnostic model.

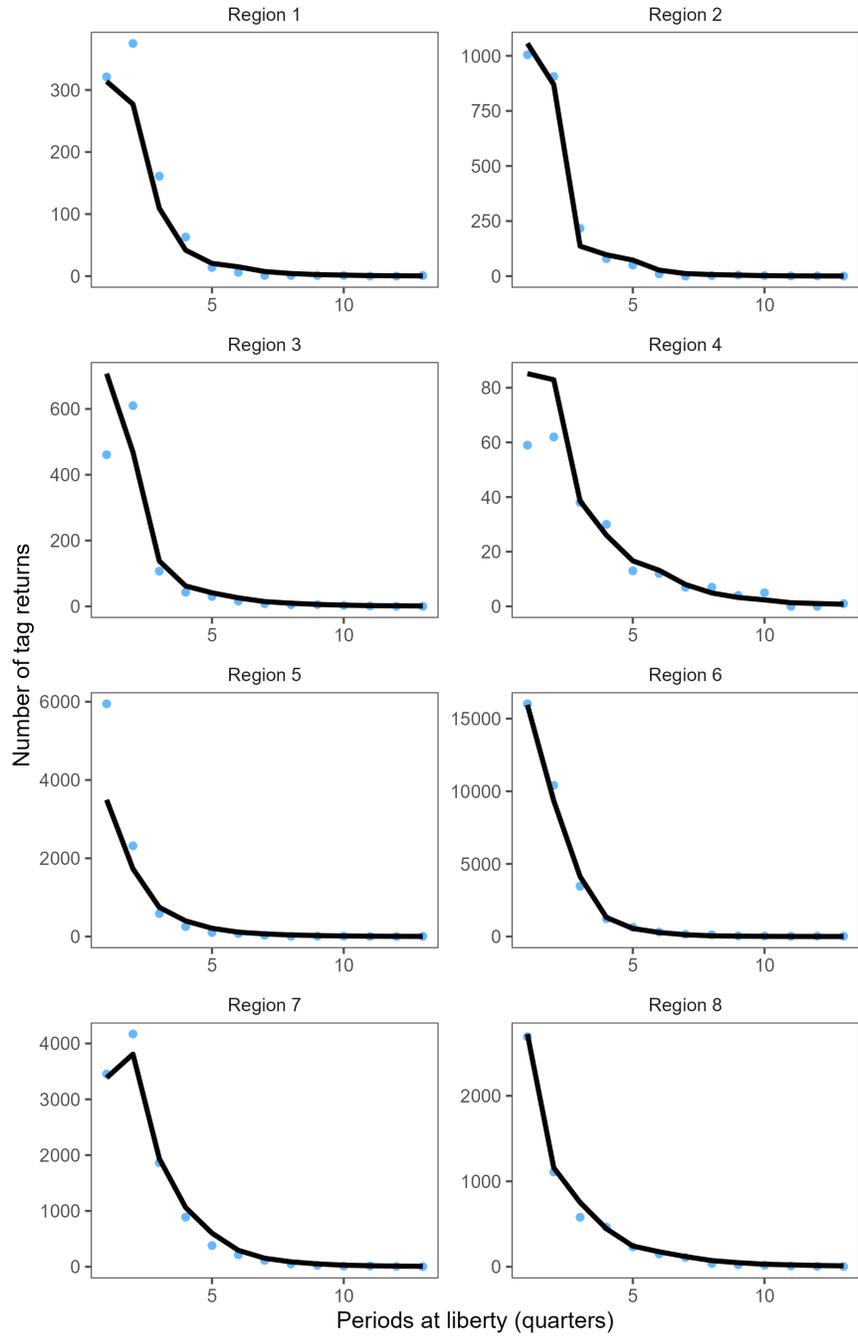


Figure 35: Observed (blue points) and model-predicted (black line) tag attrition by region for the diagnostic model.

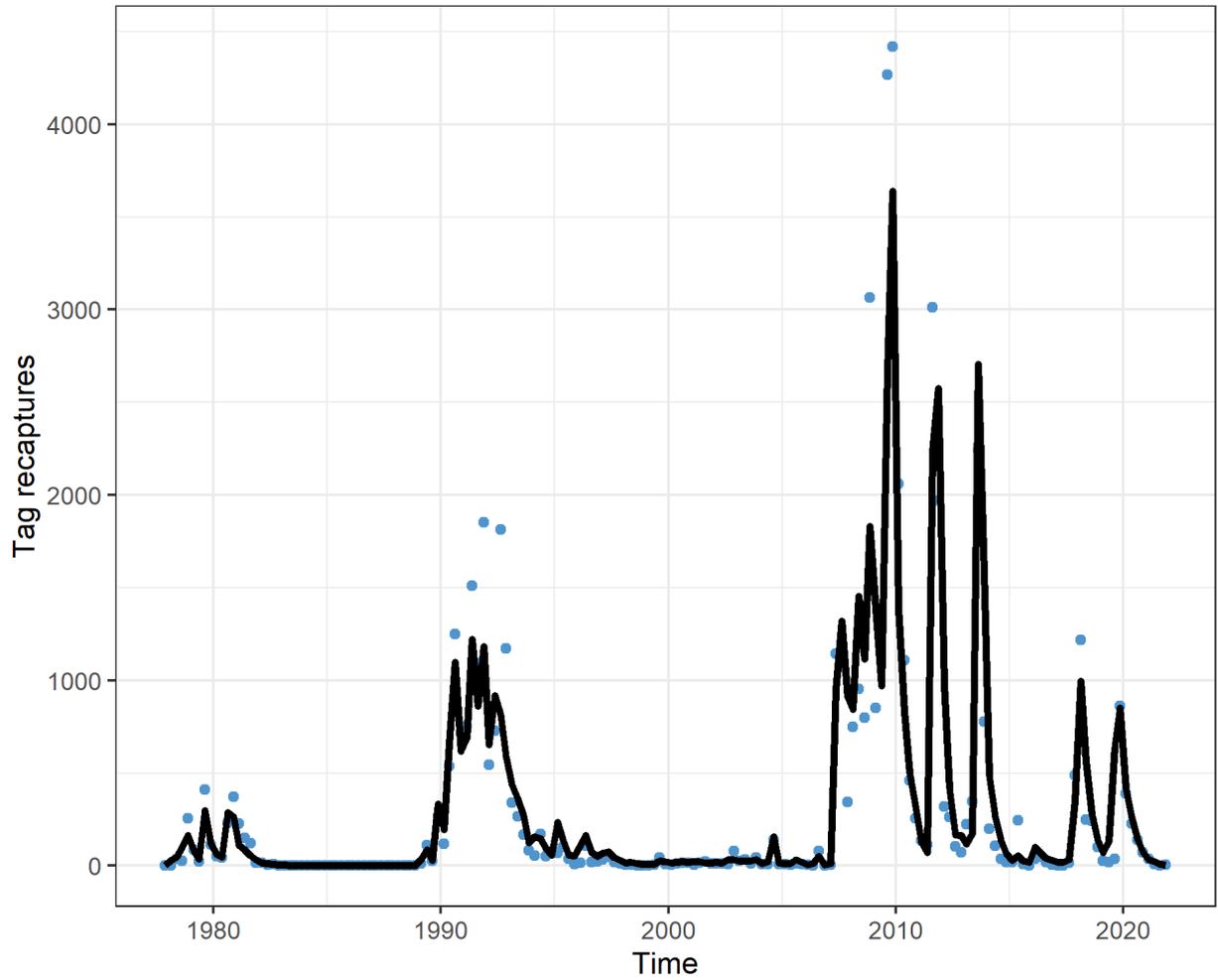


Figure 36: Observed (blue points) and model-predicted (black line) tag returns over time, with returns in the mixing period removed, for the diagnostic model across all tag release events with all tag recapture groupings aggregated.

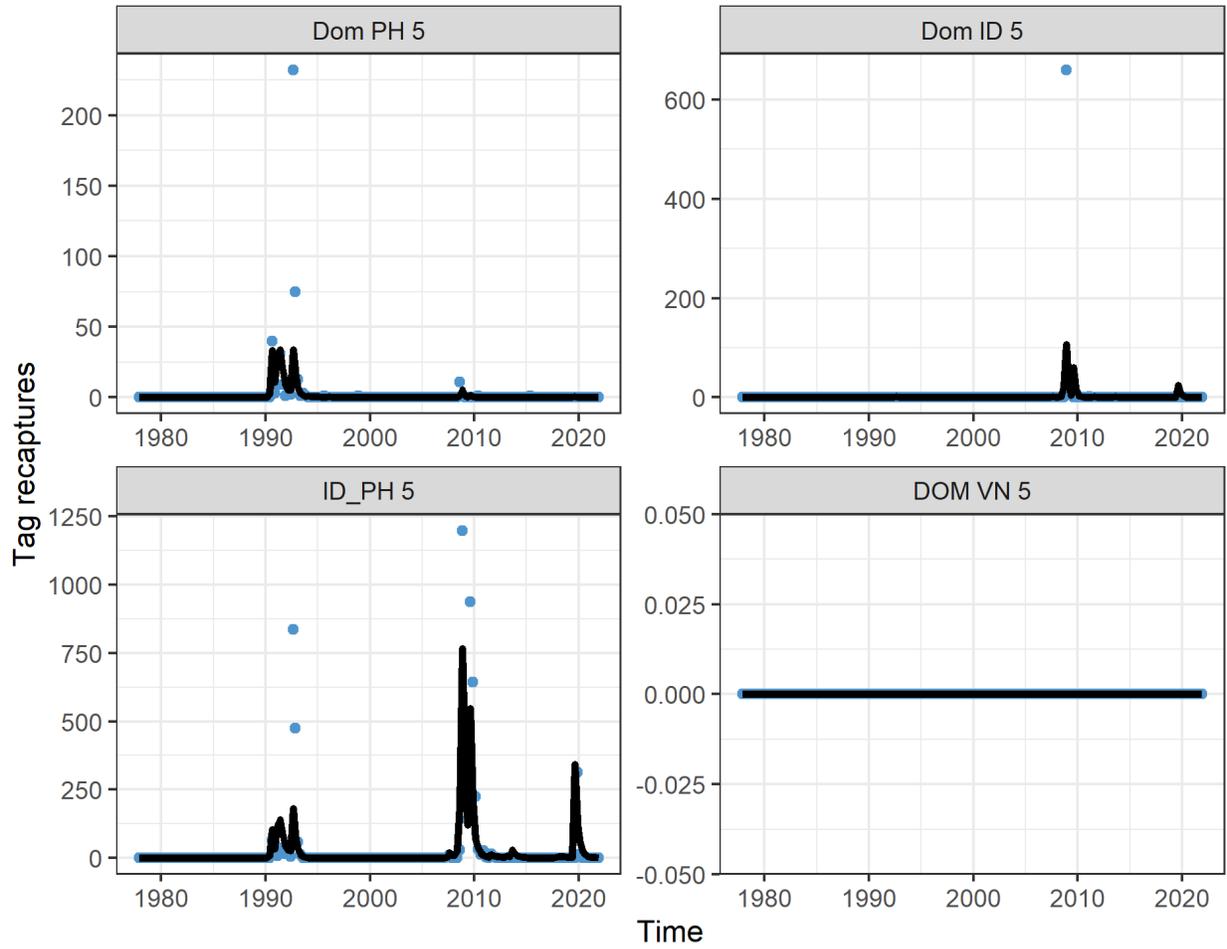


Figure 37: Observed (blue points) and model-predicted (black line) tag returns over time, with returns in the mixing period removed, for the diagnostic model for ID, VN and PH fisheries.

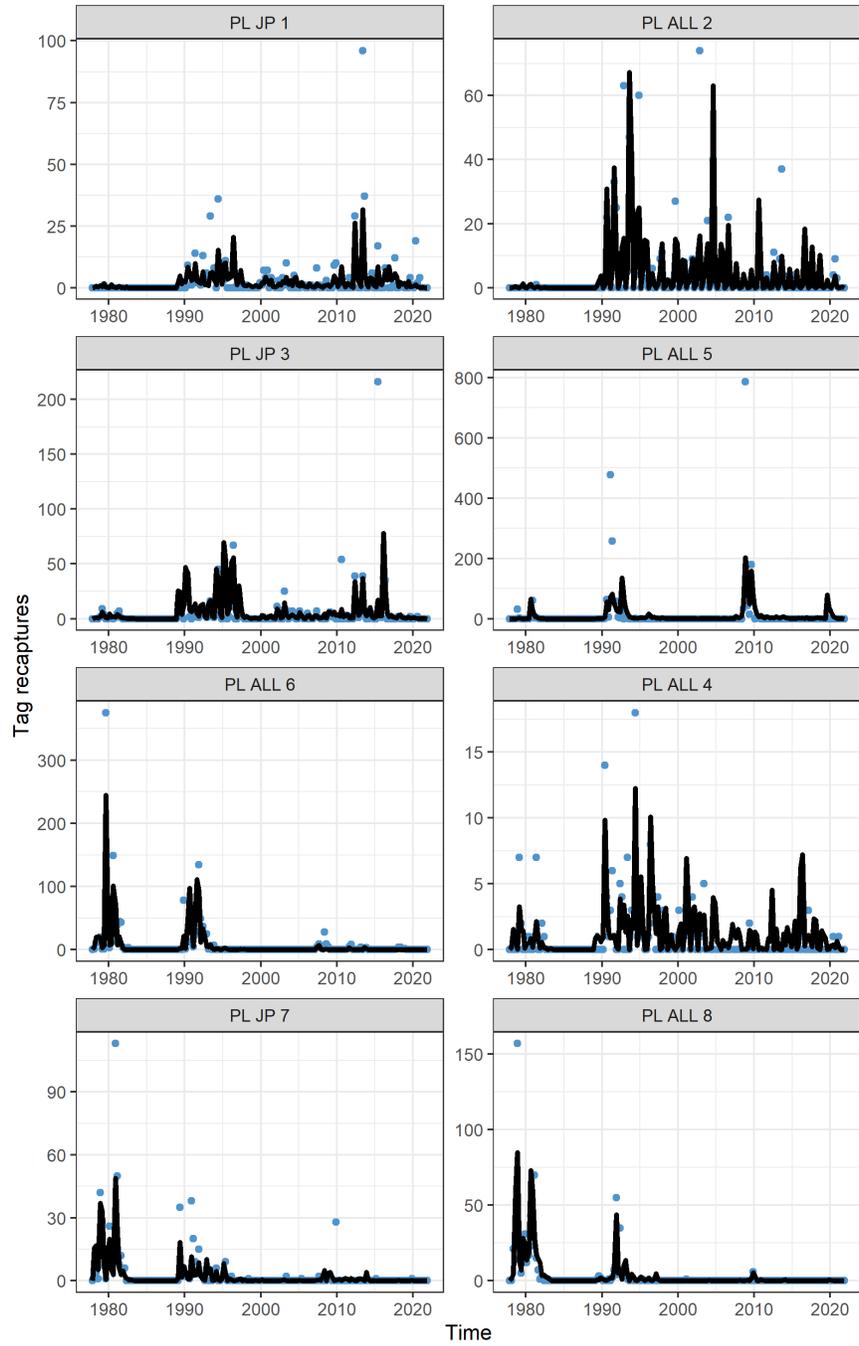


Figure 38: Observed (blue points) and model-predicted (black line) tag returns over time, with returns in the mixing period removed, for the diagnostic model for pole-and-line fisheries.

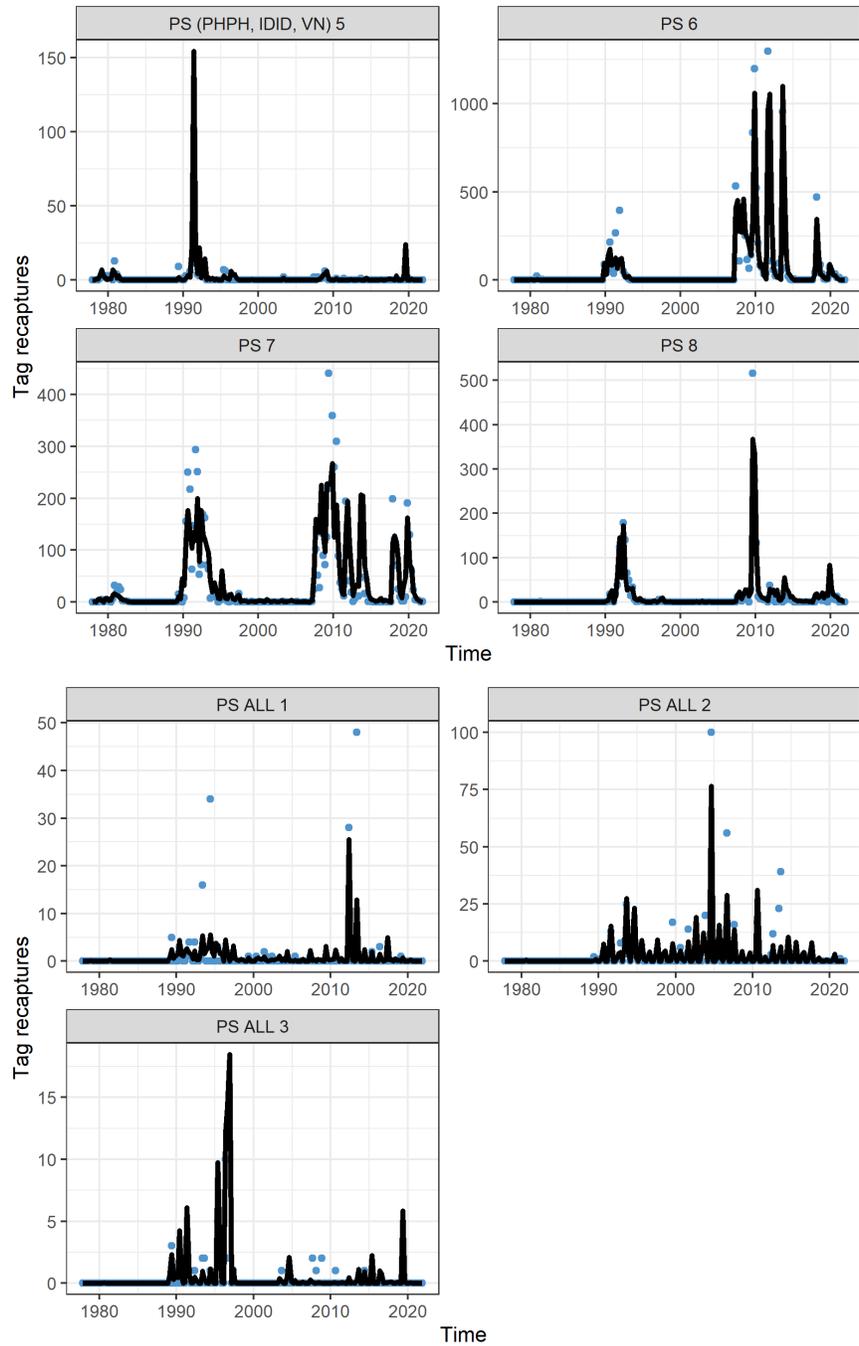


Figure 39: Observed (blue points) and model-predicted (black line) tag returns over time, with returns in the mixing period removed, for the diagnostic model for purse seine fisheries.

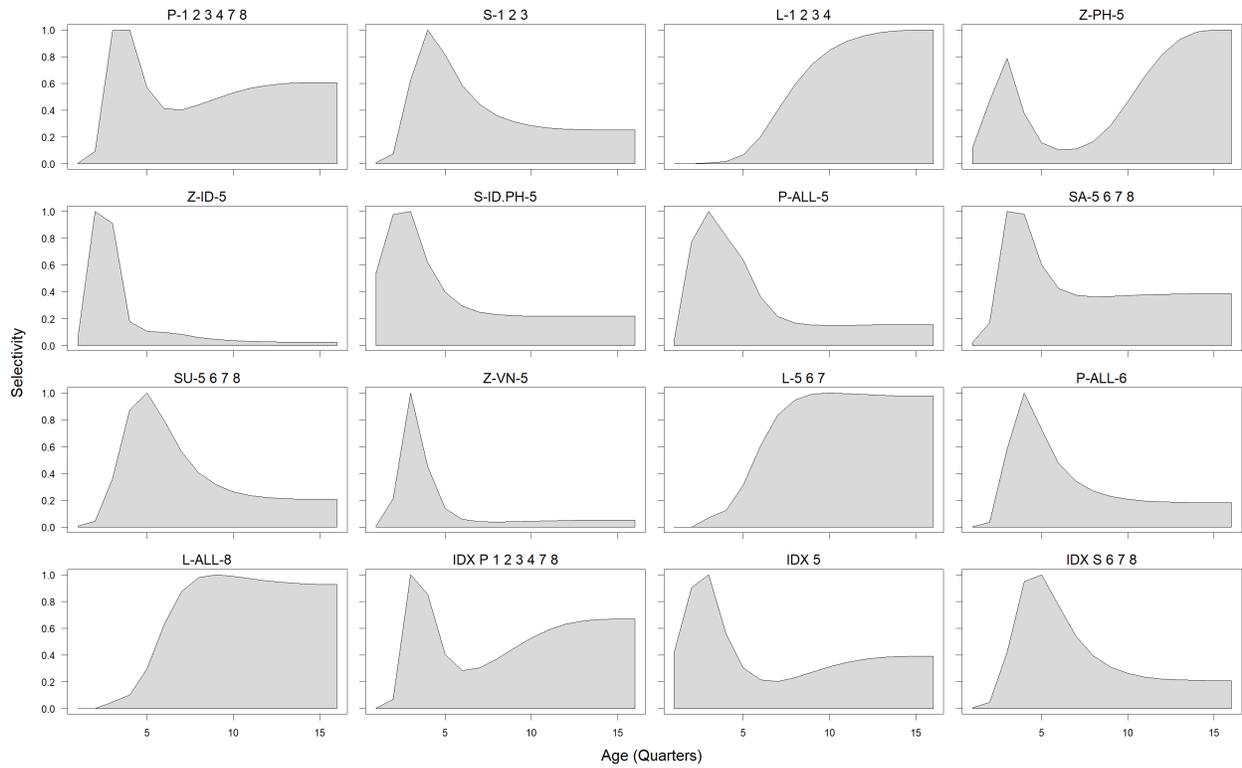


Figure 40: Age-specific (in number of quarters) selectivity coefficients by groups of fisheries with shared selectivities.

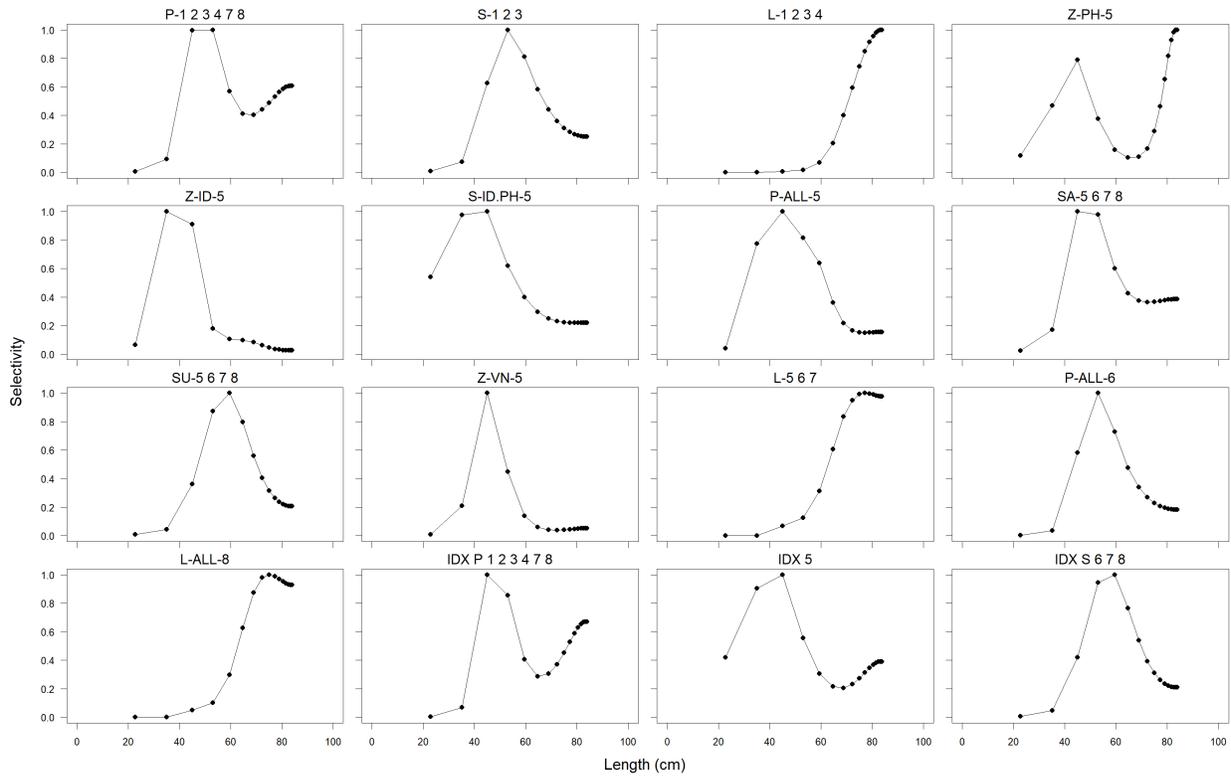


Figure 41: Length-specific selectivity coefficients by groups of fisheries with shared selectivities.

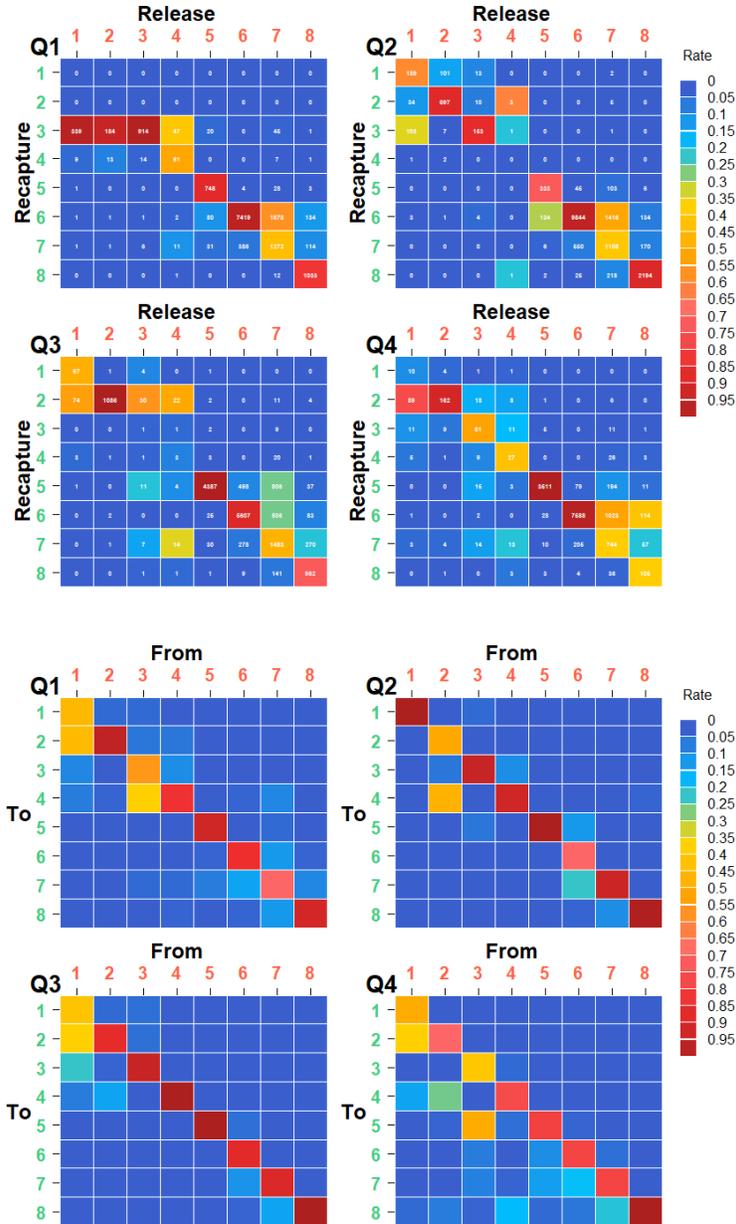


Figure 42: (Top) Observed proportion of tags returned by region of release (columns), region of recapture (rows), and quarter of recapture (panel) where the color of the tile indicates the proportion of tags returned from the region of releases, numbers in boxes indicate the actual numbers. (Bottom) Estimated movement coefficients by quarter for the diagnostic case model. The red numbers (horizontal axis) indicate the source model region (From); the green numbers (vertical axis) indicate the receiving (To) regions. The color of the tile shows the magnitude of the movement rate (proportion of individuals moving from region x to region y in that quarter), with each column adding up to 1.

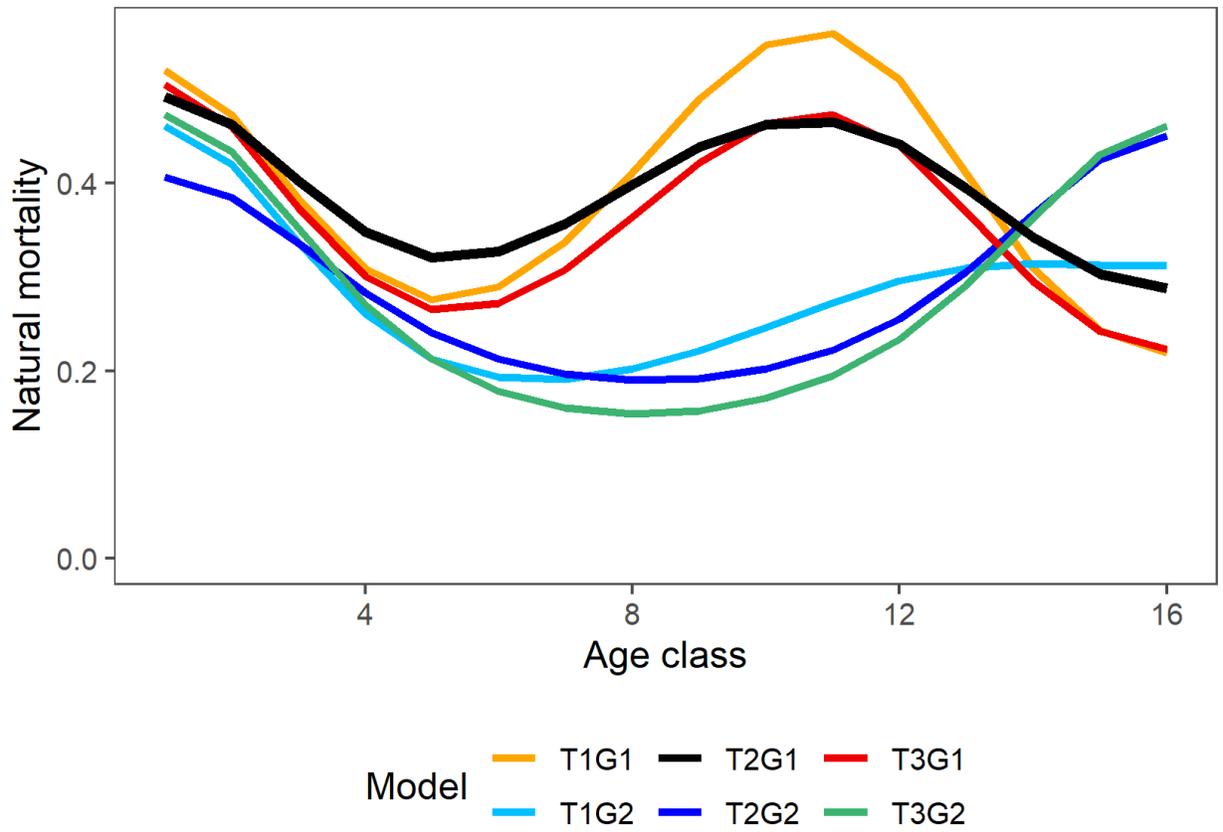


Figure 43: Quarterly natural mortality-at-age as estimated by the diagnostic model and five other models that vary in the growth curve applied and D value used to determine the tag mixing periods, all models in the figure have steepness 0.8, steepness having no impact on the model estimated natural mortality curves.

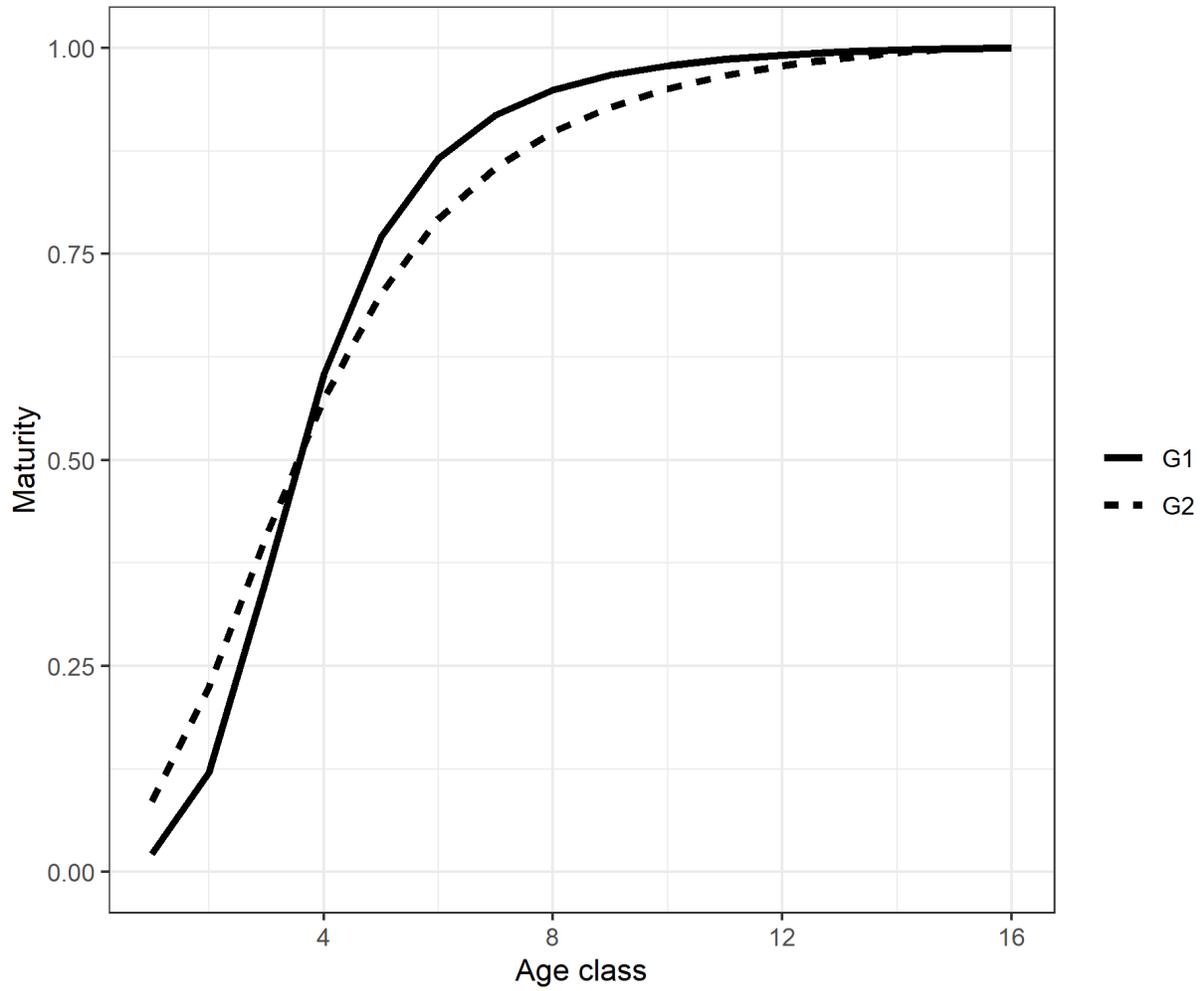


Figure 44: Maturity-at-age as estimated in the models that applied growth estimated internally in the model (G1) and externally (G2), the figures are for models with steepness of 0.8 and tag mixing based on D value of 0.2, these latter assumptions had negligible influence on Maturity-at-age estimates.

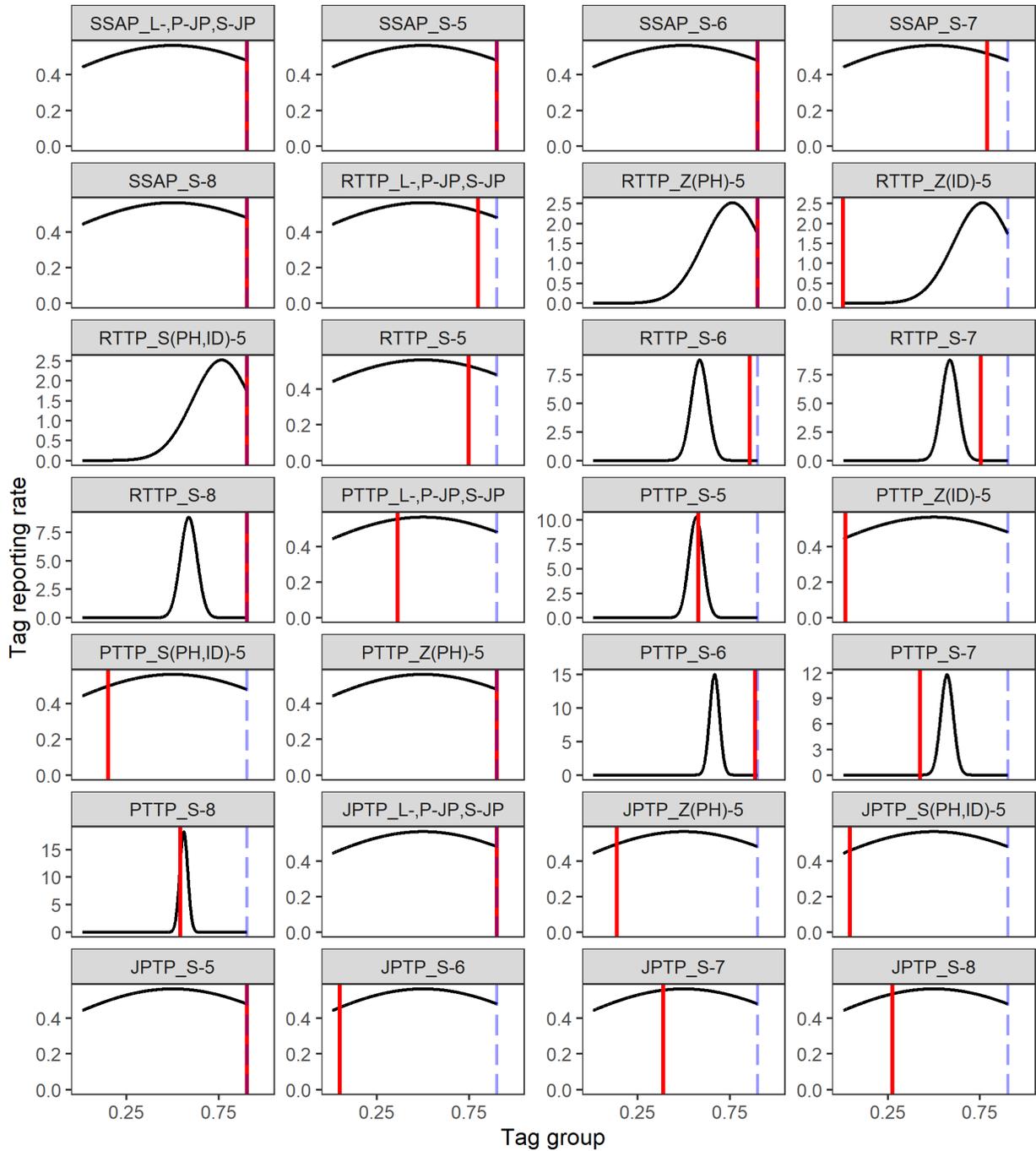


Figure 45: Estimated reporting rates for the diagnostic model (red lines) and the prior distribution for each reporting rate group (black lines). The imposed upper bound (0.99) on the reporting rate parameters is shown as a blue dashed line. Reporting rates can be estimated separately for each release program and recapture fishery group but in practice are aggregated over some recapture groups to reduce dimensionality. Fishery groups SSAP-S-8, RTP-S(PH,ID)-5, RTP-S-8, PTP-S-5 were below the upper bound sensitivity of 0.99

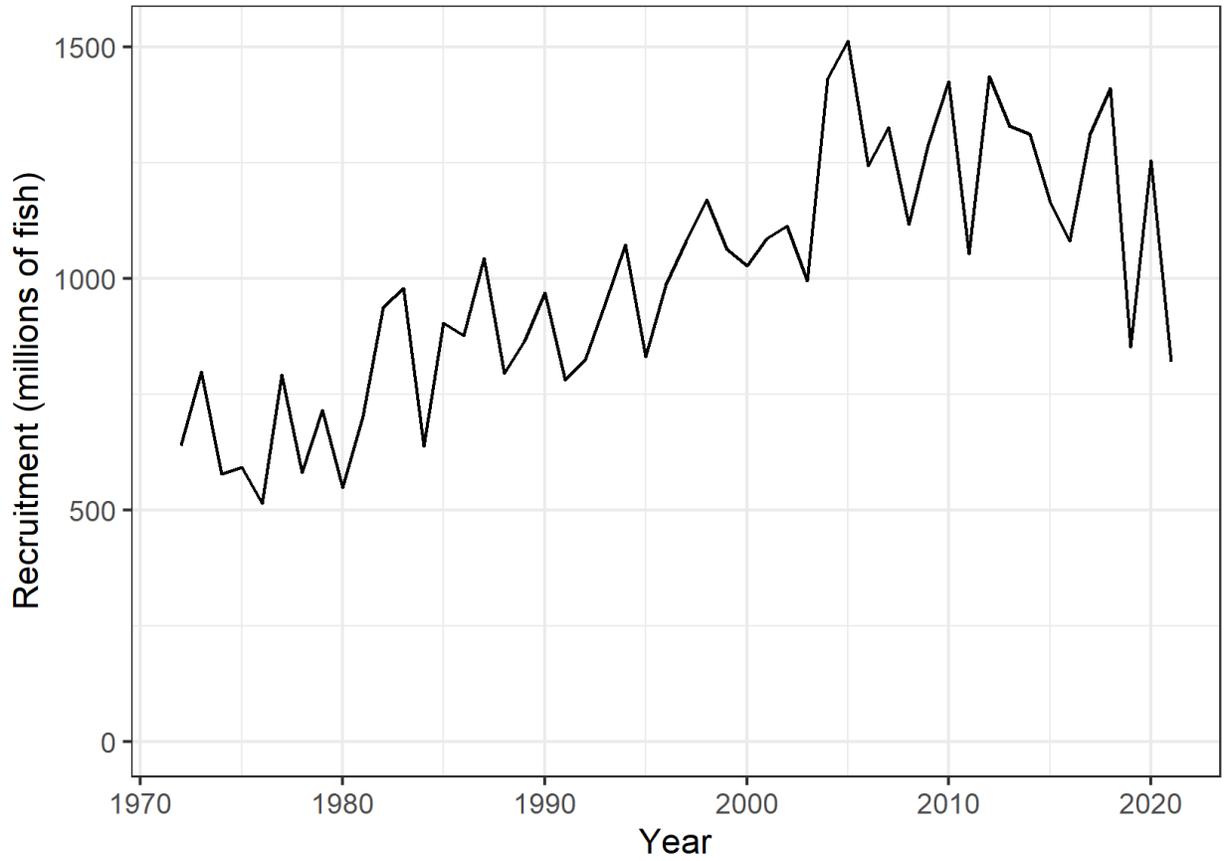


Figure 46: Annual time series of estimated average quarterly recruitment summed across regions for the diagnostic model.

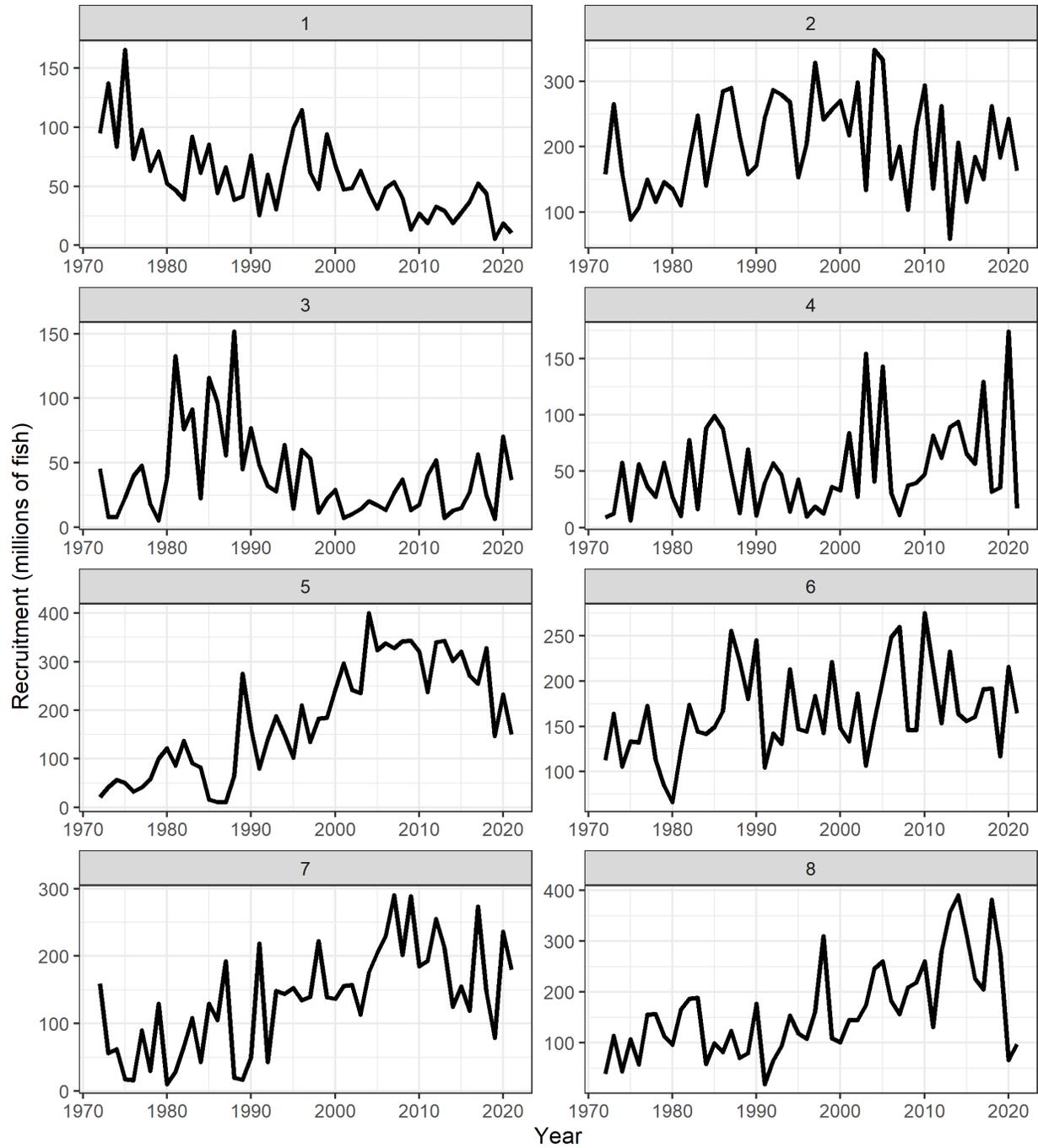


Figure 47: Annual time series of estimated average quarterly recruitment by model region for the diagnostic model. Note that the scale of the y-axis is not constant across regions.

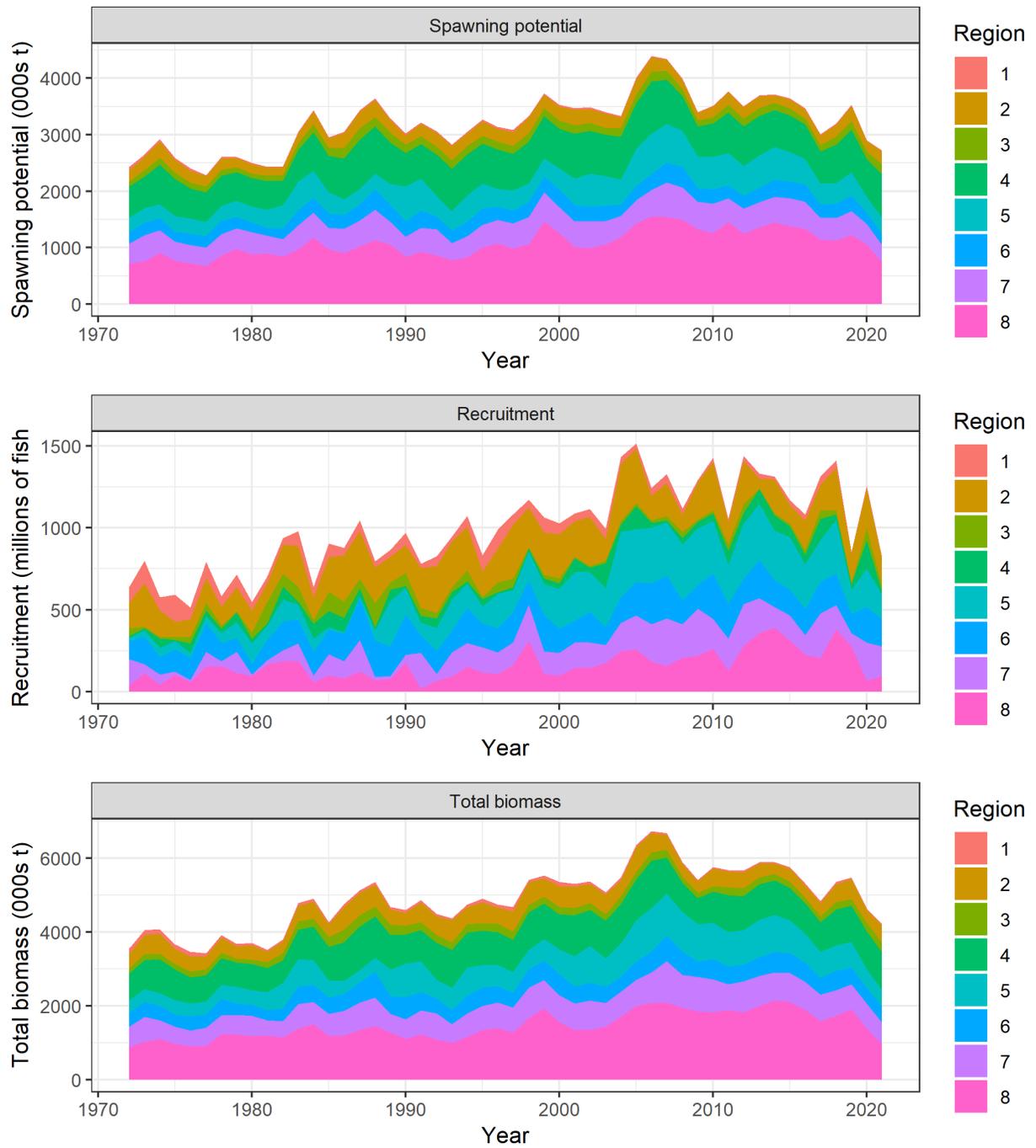


Figure 48: Annual time series of estimated average quarterly recruitment, spawning potential and total biomass by model region for the diagnostic model, showing the relative proportions among regions.

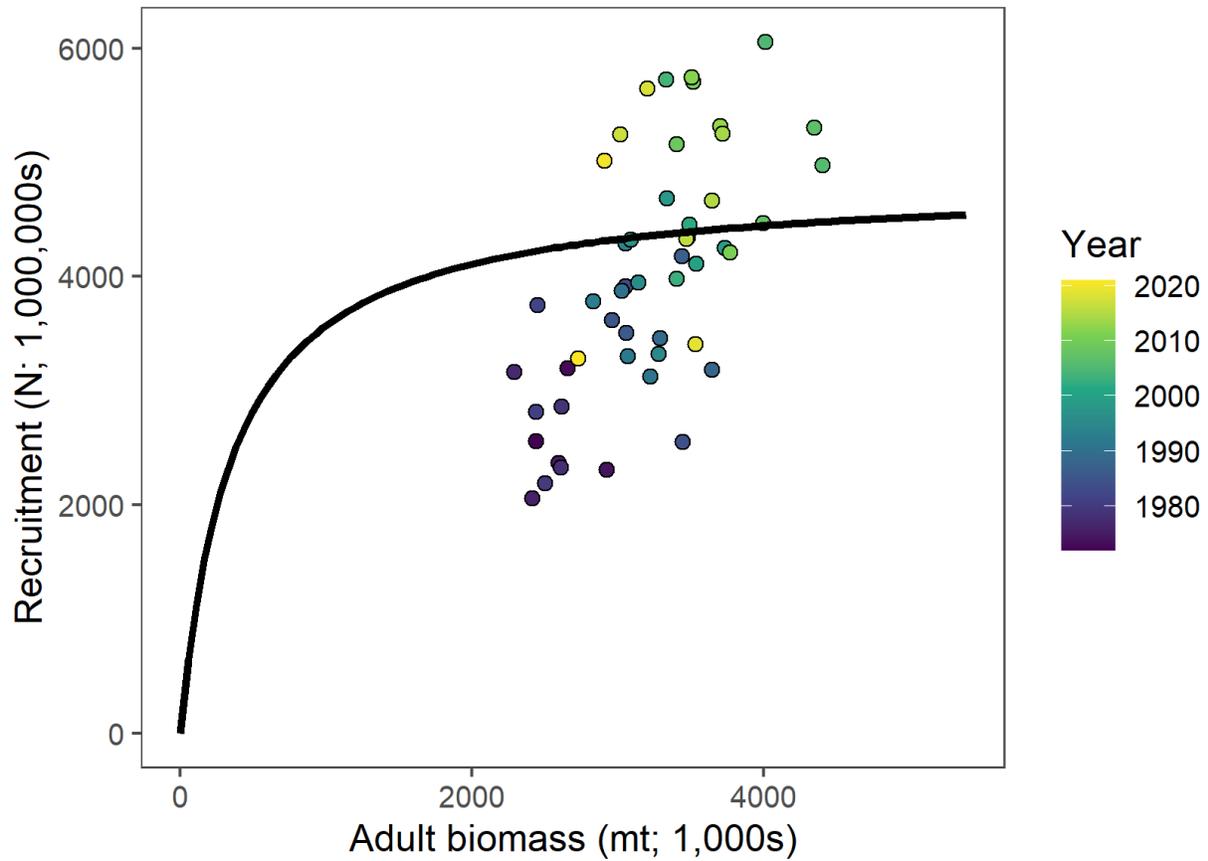


Figure 49: Estimated relationship between recruitment and spawning potential based on annual values for the diagnostic model. The darkness of the circles changes from light (more recent) to dark (earlier) through time.

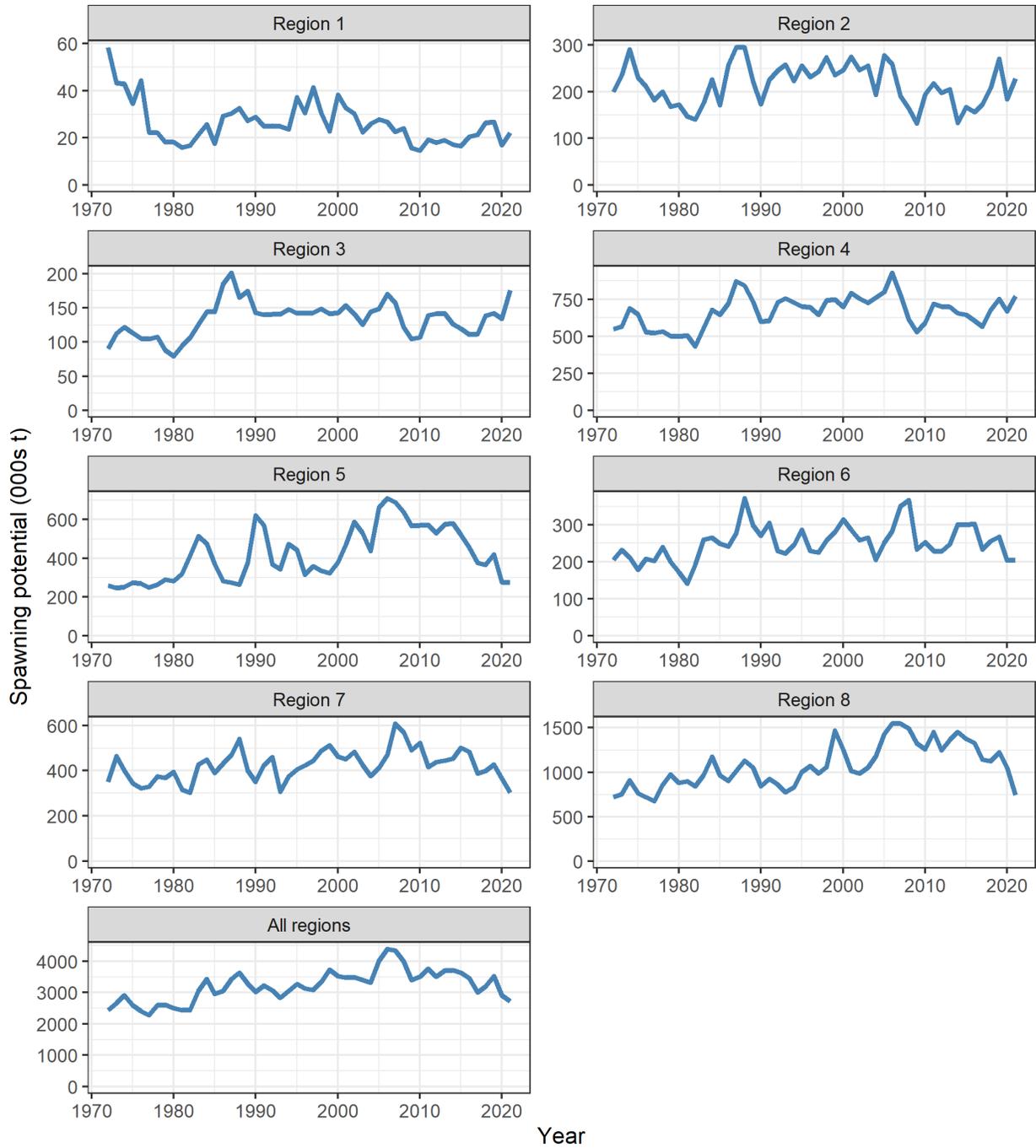


Figure 50: Estimated temporal spawning potential by model region, and for all model regions summed for the diagnostic model. Note that the scale of the y-axis is not constant across regions.

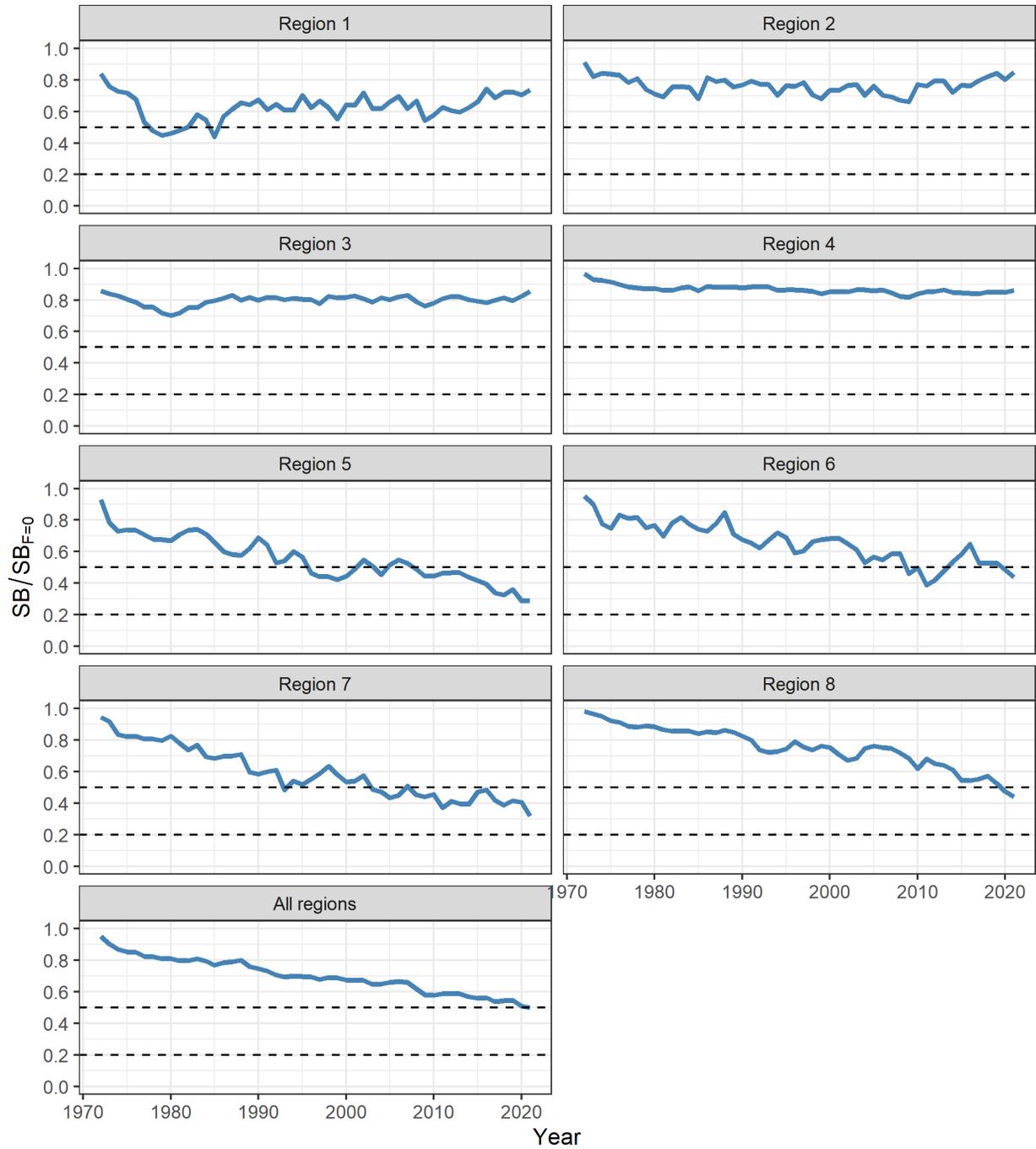


Figure 51: Estimated annual spawning potential depletion by model region, and for all model regions summed for the diagnostic model.

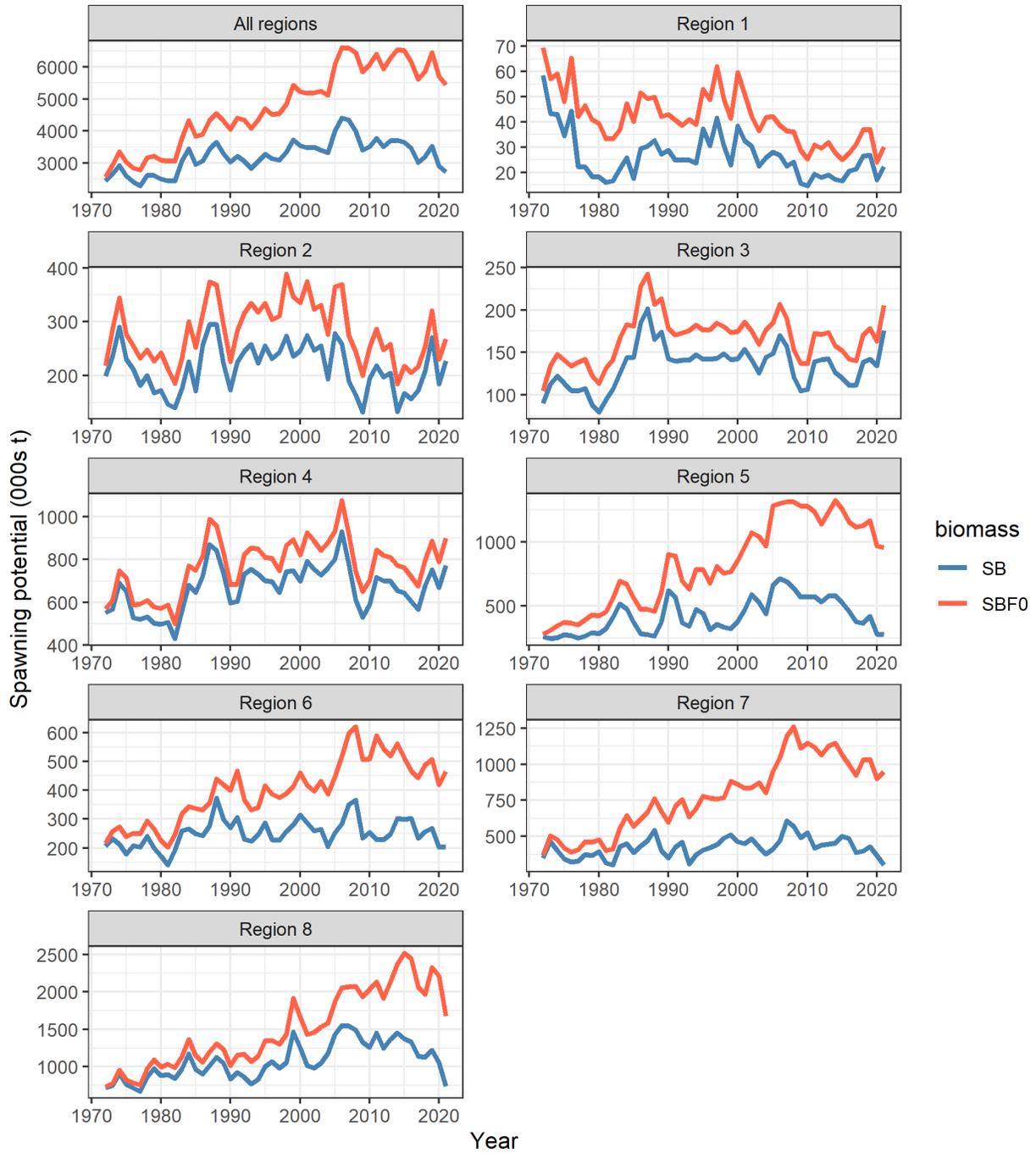


Figure 52: Comparison of the estimated annual spawning potential trajectories (lower blue lines) with the spawning potential trajectories predicted to have occurred in the absence of fishing (upper red lines) for each region and overall, for the diagnostic model. Note the scales of the Y-axes vary.

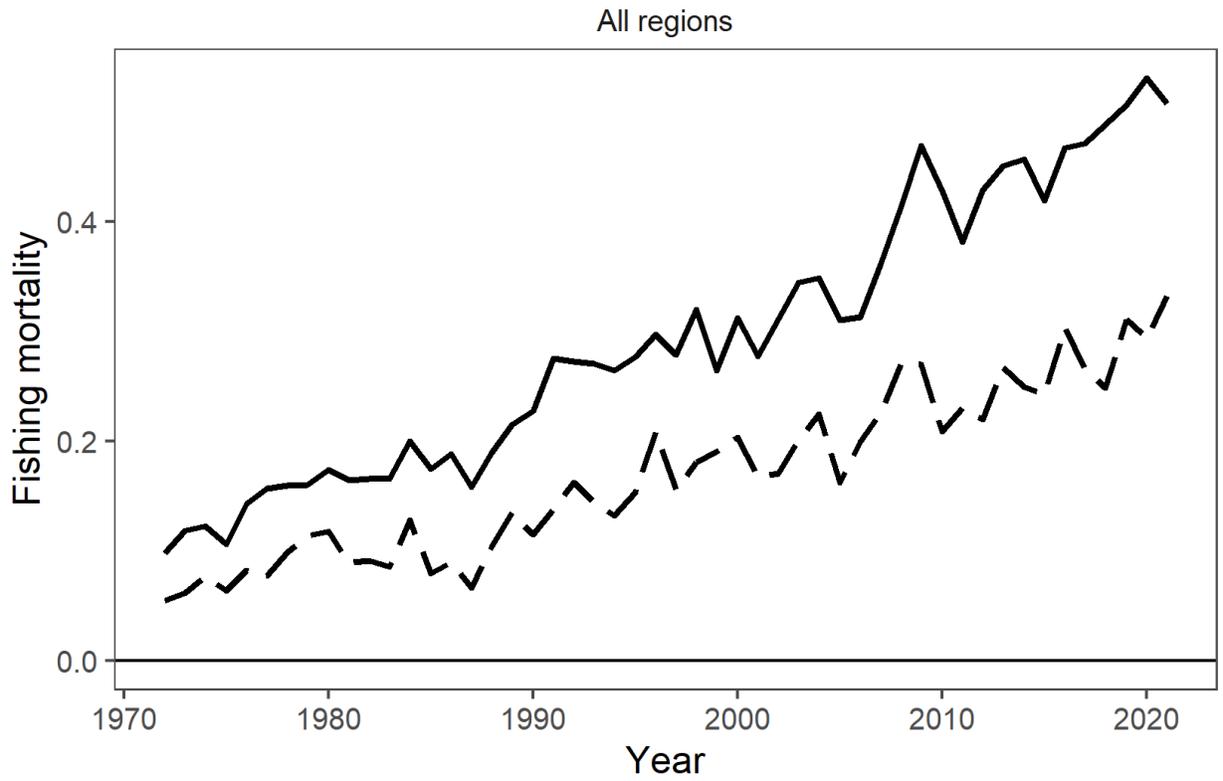


Figure 53: Estimated annual average adult (solid line) and juvenile (dashed line) fishing mortality for the diagnostic model.

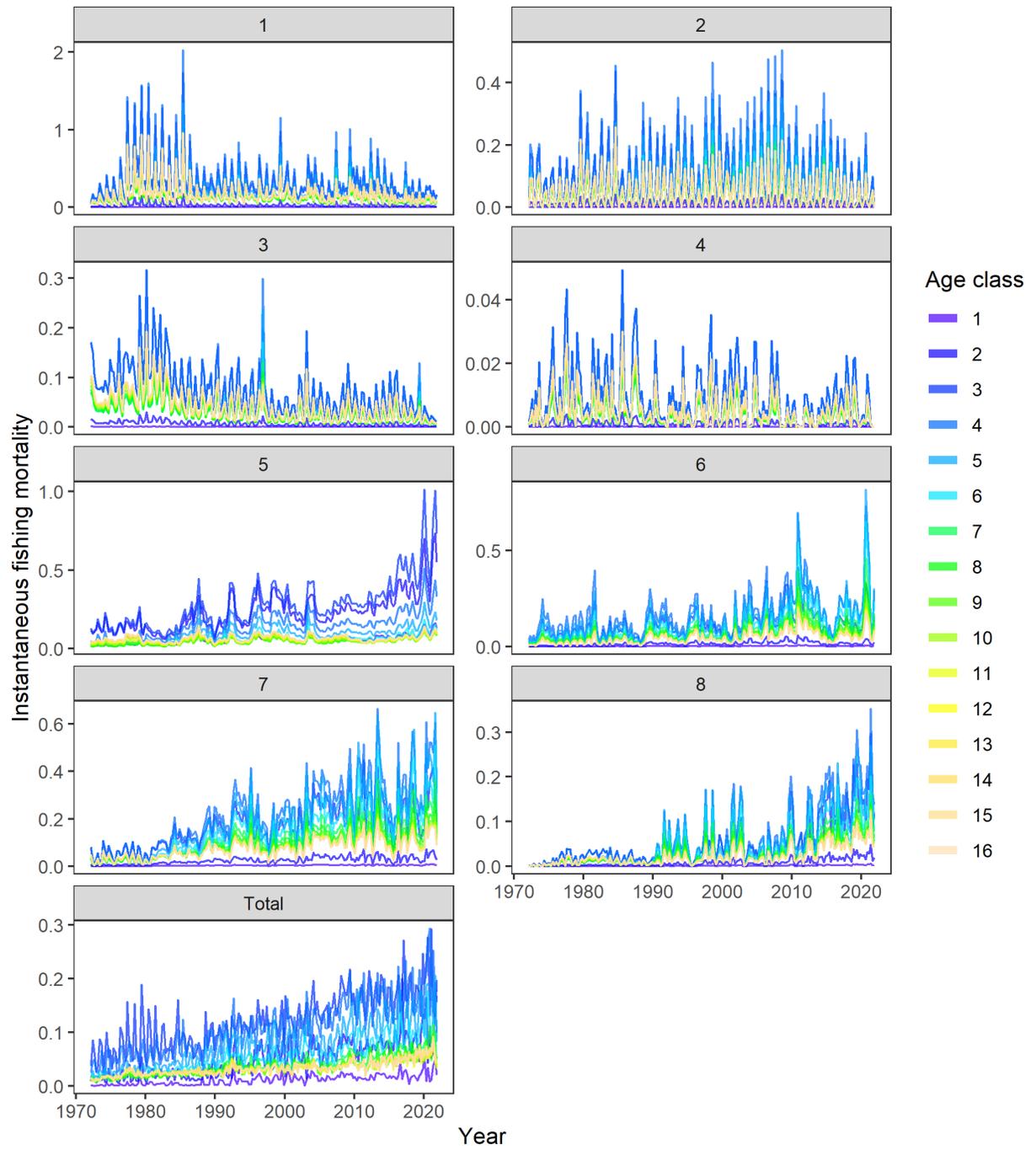


Figure 54: Estimated age-specific fishing mortality for the diagnostic model, by region and overall.

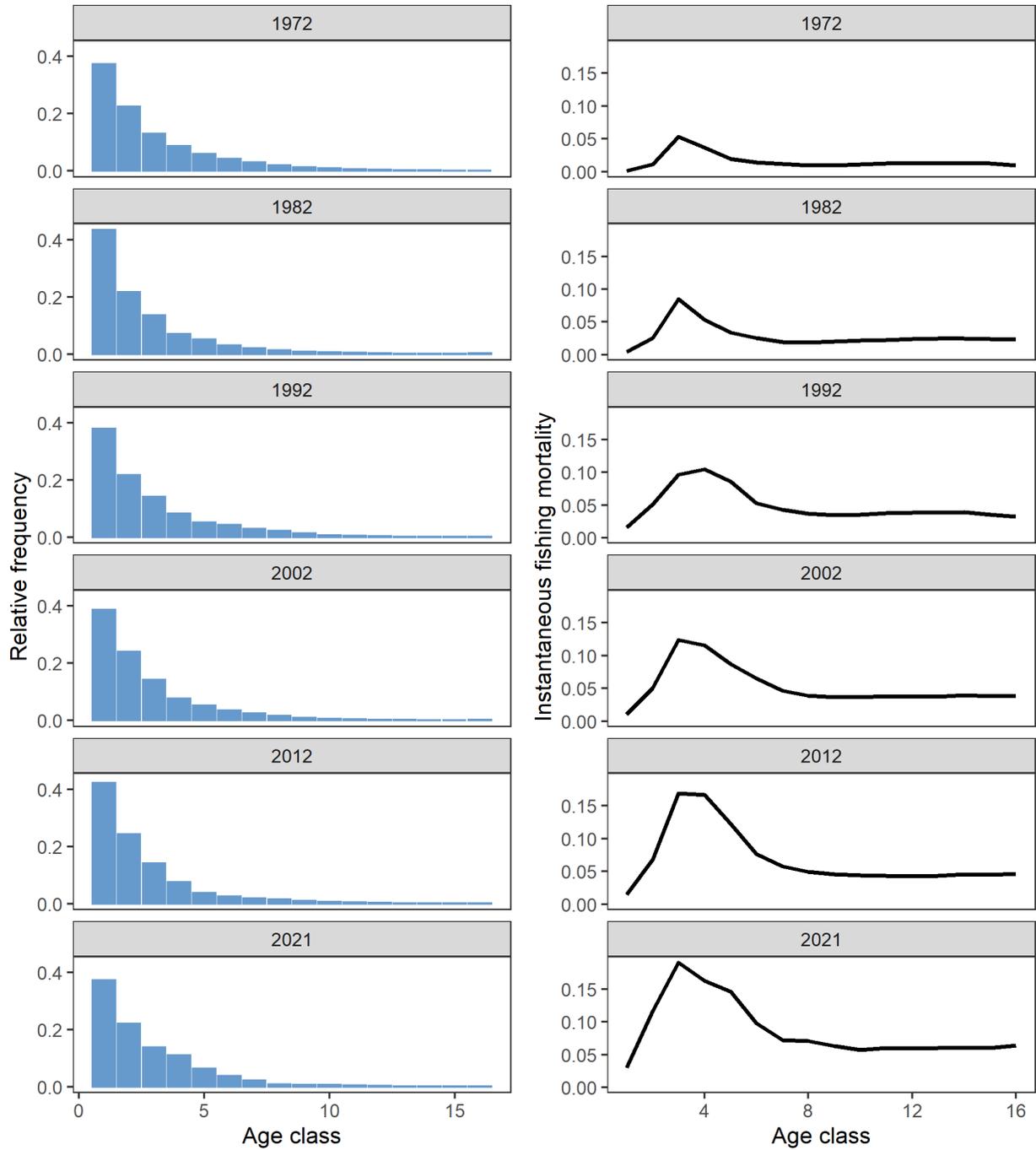


Figure 55: Estimated proportion at age (quarters) and fishing mortality at age (right), by year, at decade intervals, for the diagnostic model.

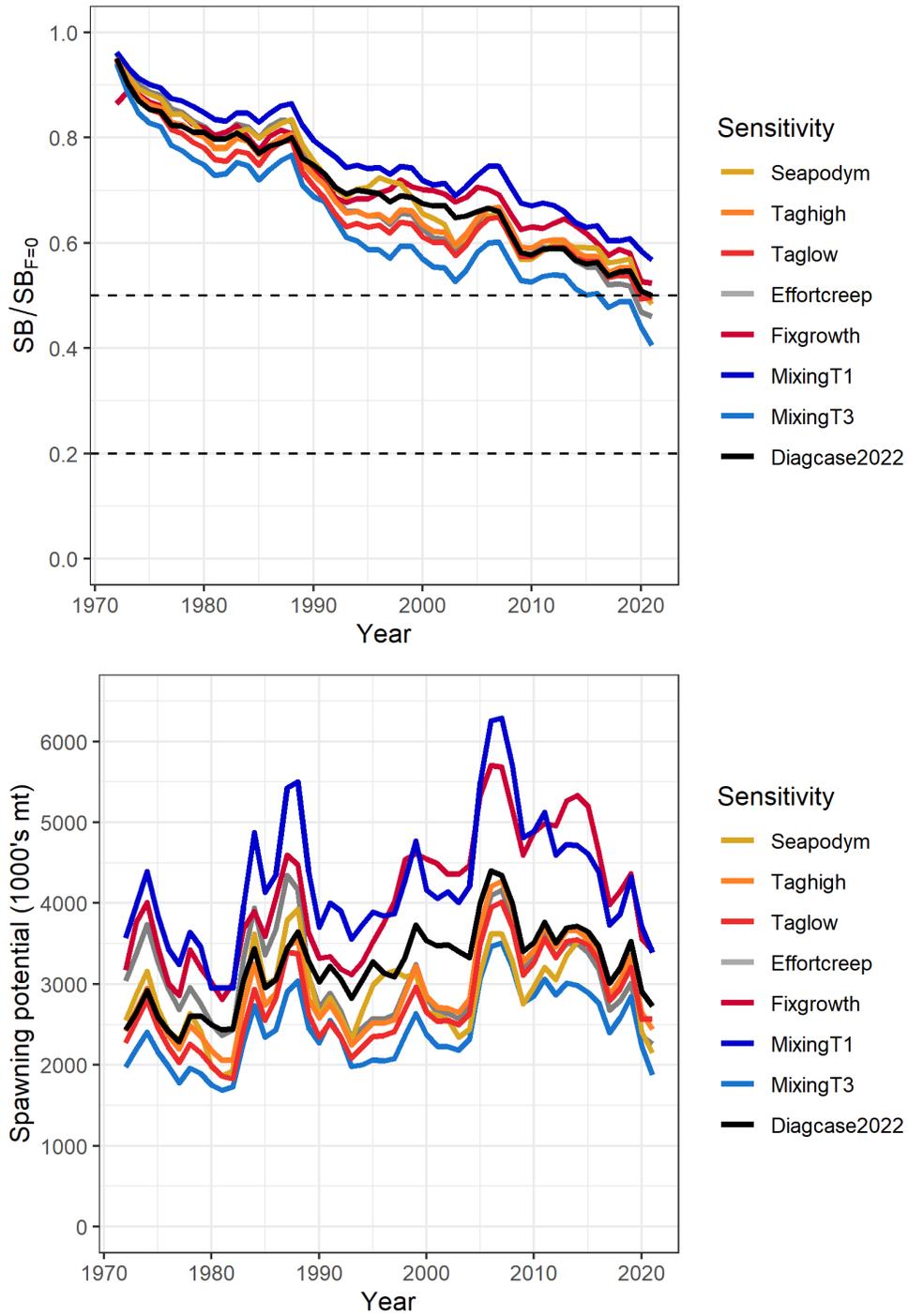


Figure 56: Estimated dynamic spawning depletion (Top) and spawning potential (Bottom) for the one-off sensitivities from the 2022 diagnostic case.

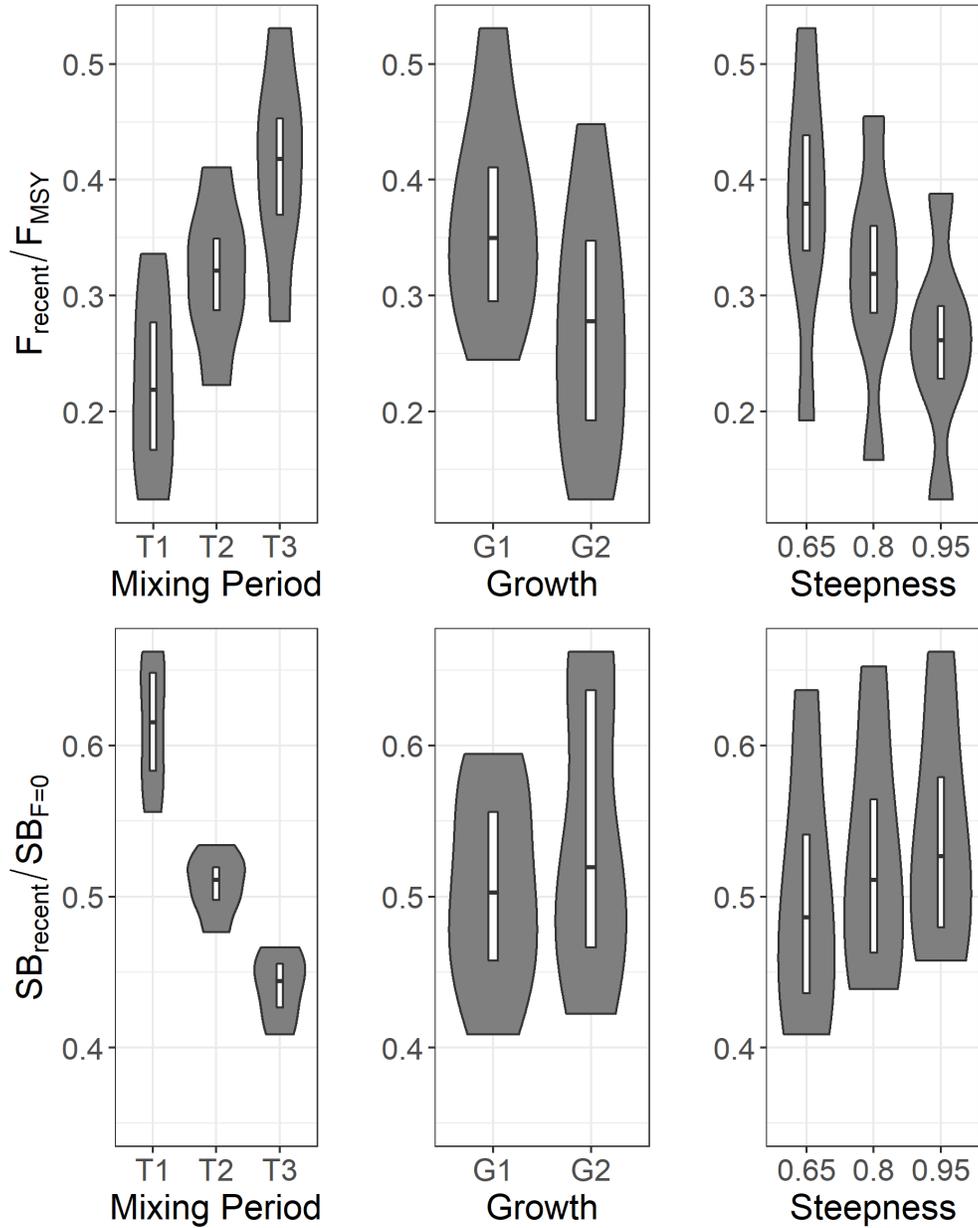


Figure 57: Box and violin plots summarizing (Top) the estimated F_{recent}/F_{MSY} and (Bottom) $SB_{recent}/SB_{F=0}$ for each of the models in the structural uncertainty grid grouped by uncertainty axes (growth, tag mixing and steepness). The line in the box is the median of the estimates, while the box shows the 50th percentile. The shaded area shows the probability distribution (or density) of the estimates of all models of the structural uncertainty grid.

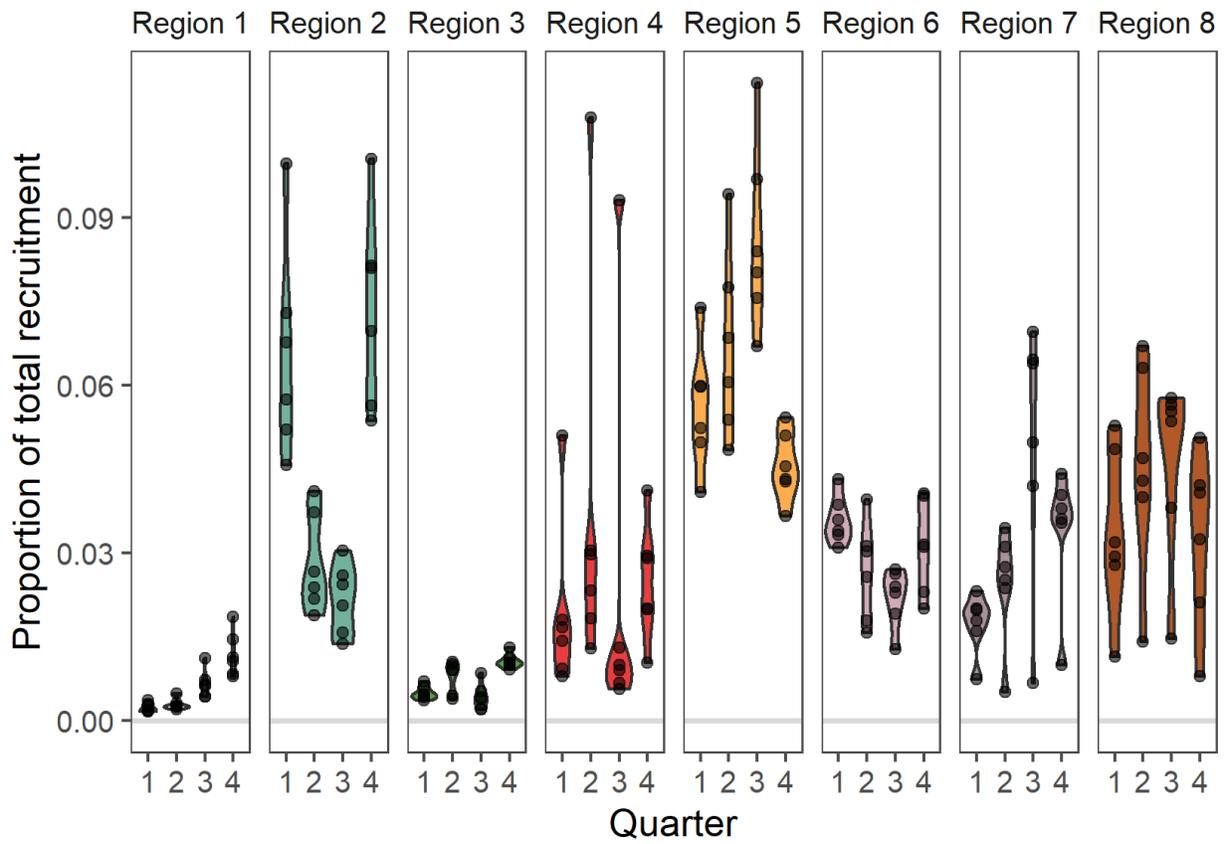


Figure 58: Distribution of annual proportions of recruitment by region (panels) and quarter (within panel group) for all years and all models in the structural uncertainty grid.

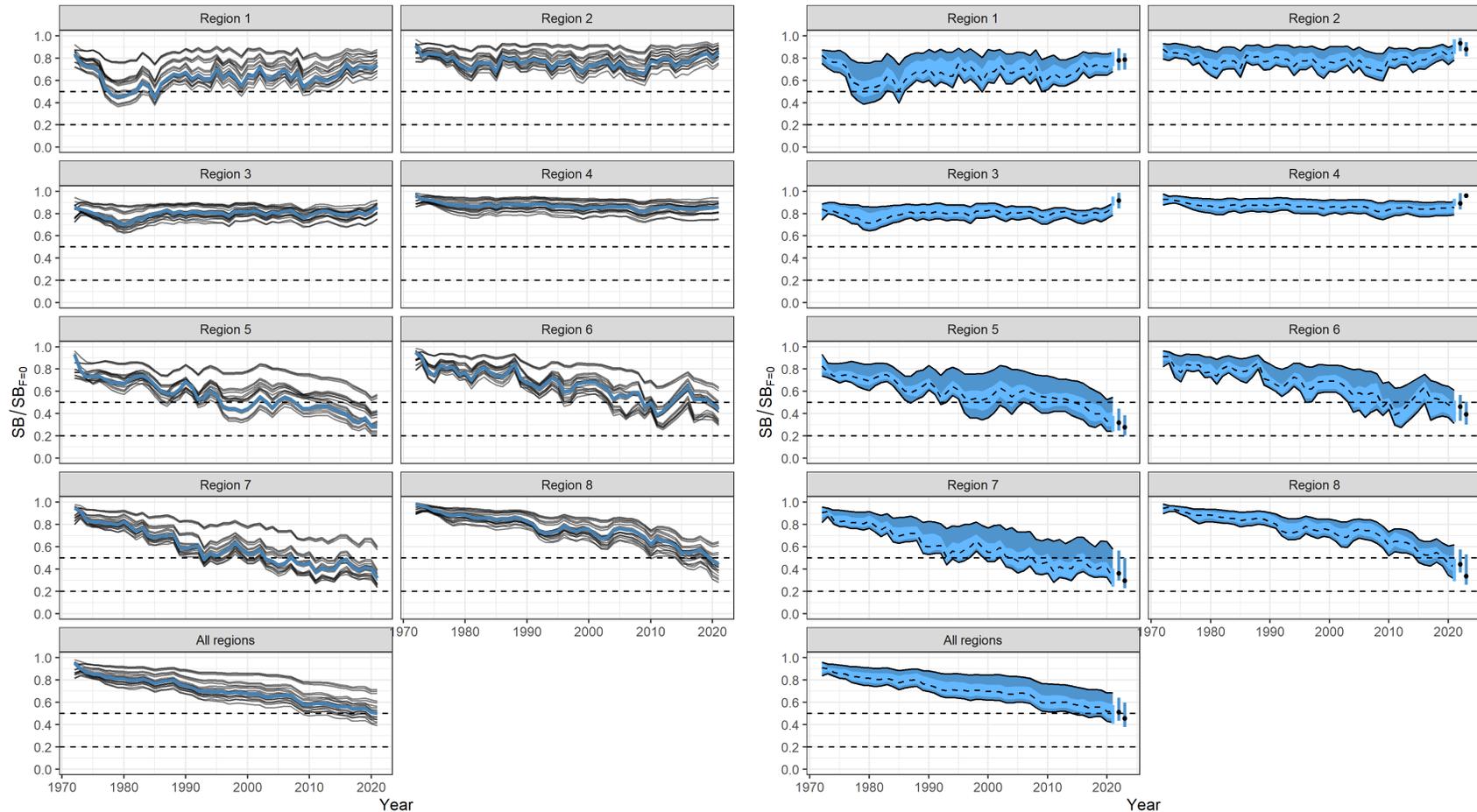


Figure 59: (Left) Trajectories of spawning depletion for the individual model runs included in the structural uncertainty grid over the period 1972-2021. (Right) Estimated spawning depletion across all models in the structural uncertainty grid over the period 1972-2021. The dashed line represents the median. The lighter band shows the 50th percentile, and the dark band shows the 80th percentile of the model estimates. The bars at the right of each ribbon indicate the median (black dots) and 80th percentile range for (left bar) $SB_{recent}/SB_{F=0}$ and (right bar) $SB_{latest}/SB_{F=0}$.

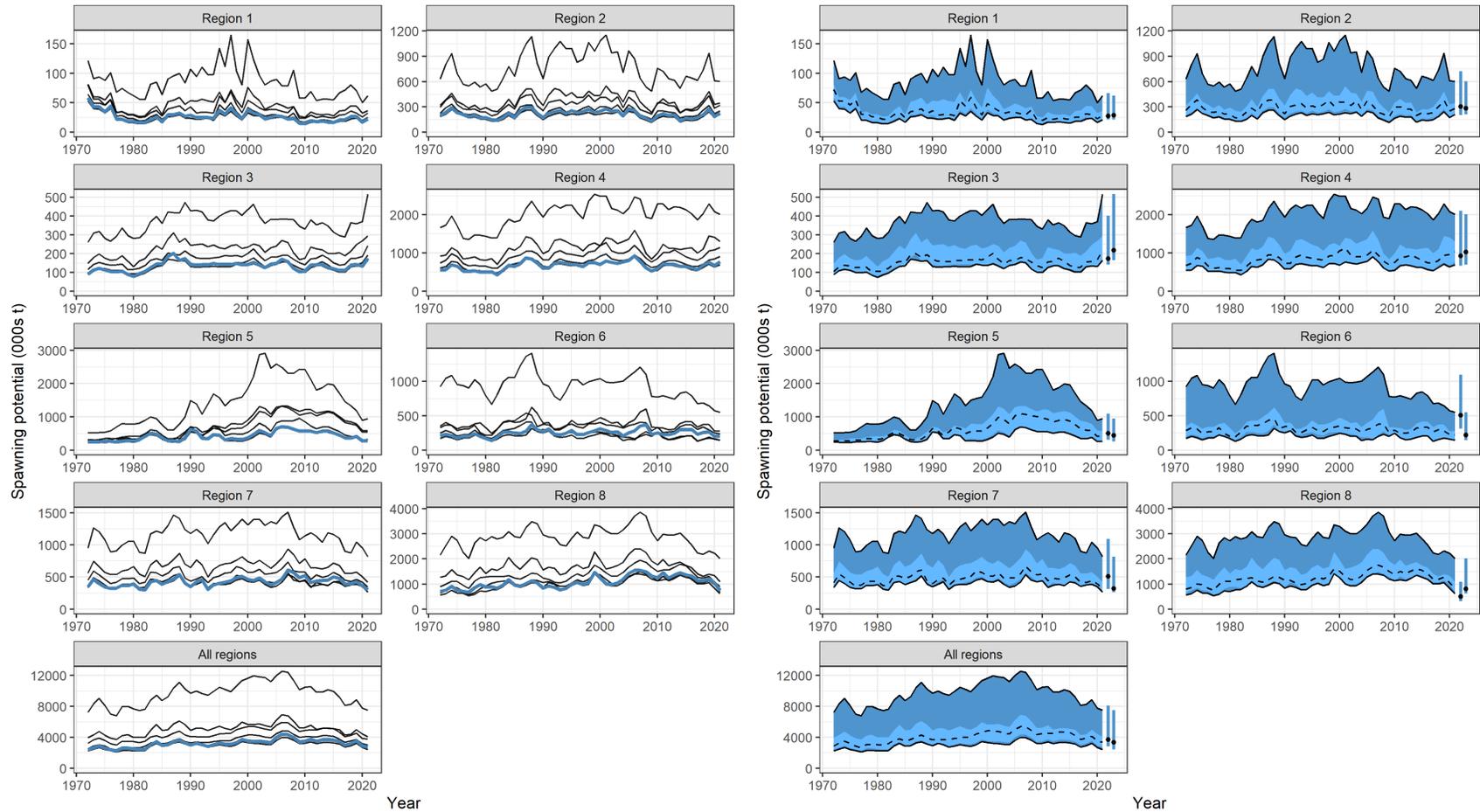


Figure 60: (Left) Trajectories of spawning potential for the individual model runs included in the structural uncertainty grid over the period 1972-2021. (Right) Estimated spawning potential across all models in the structural uncertainty grid over the period 1972-2021. The dashed line represents the median. The lighter band shows the 50th percentile, and the dark band shows the 80th percentile of the model estimates. The bars at the right of each ribbon indicate the median (black dots) and 80th percentile range for (left bar) SB_{recent} and (right bar) SB_{latest} .

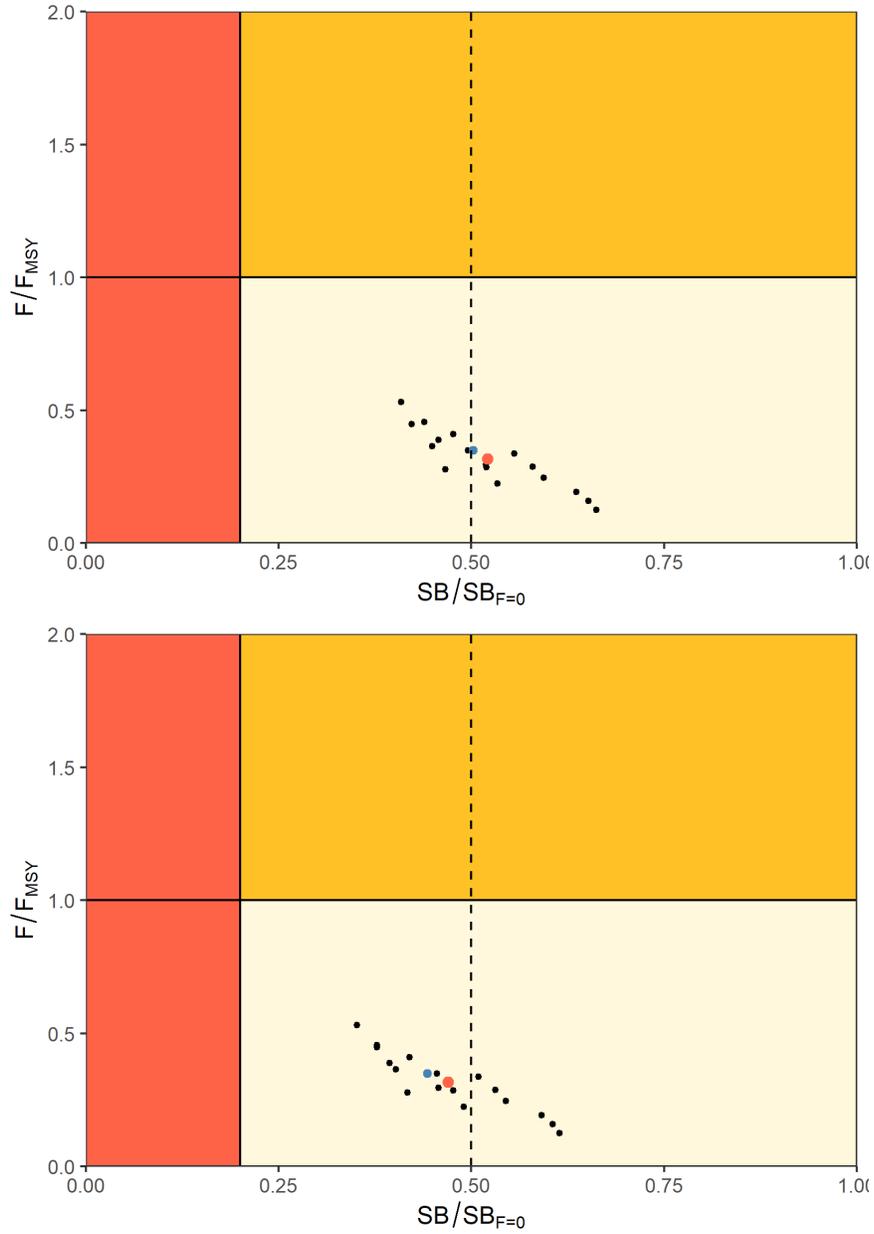


Figure 61: Majuro plots summarising the results for each of the models in the structural uncertainty grid for (Top) recent (2018-2021) and (Bottom) latest (2021) periods. The vertical dotted line on the Majuro plots is included to indicate the interim TRP of 0.50 $SB_{F=0}$ for the WCPFC-CA skipjack stock as specified in CMM 2021-01. The blue point is the diagnostic case model and the red point is the median.

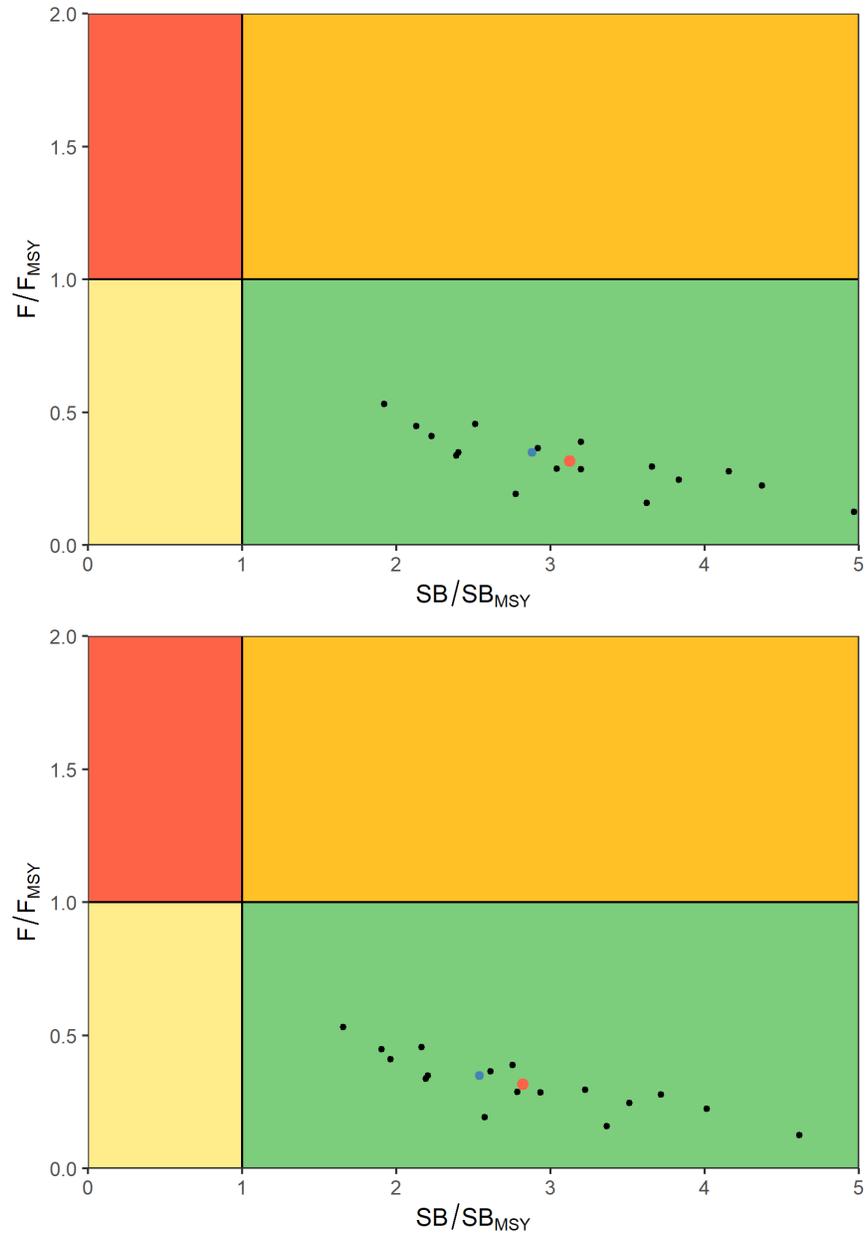


Figure 62: Kobe plots summarising the results for each of the models in the structural uncertainty grid for (Top) recent (2018-2021) and (Bottom) latest (2021) periods. The blue point is the diagnostic case model and red point is the median.

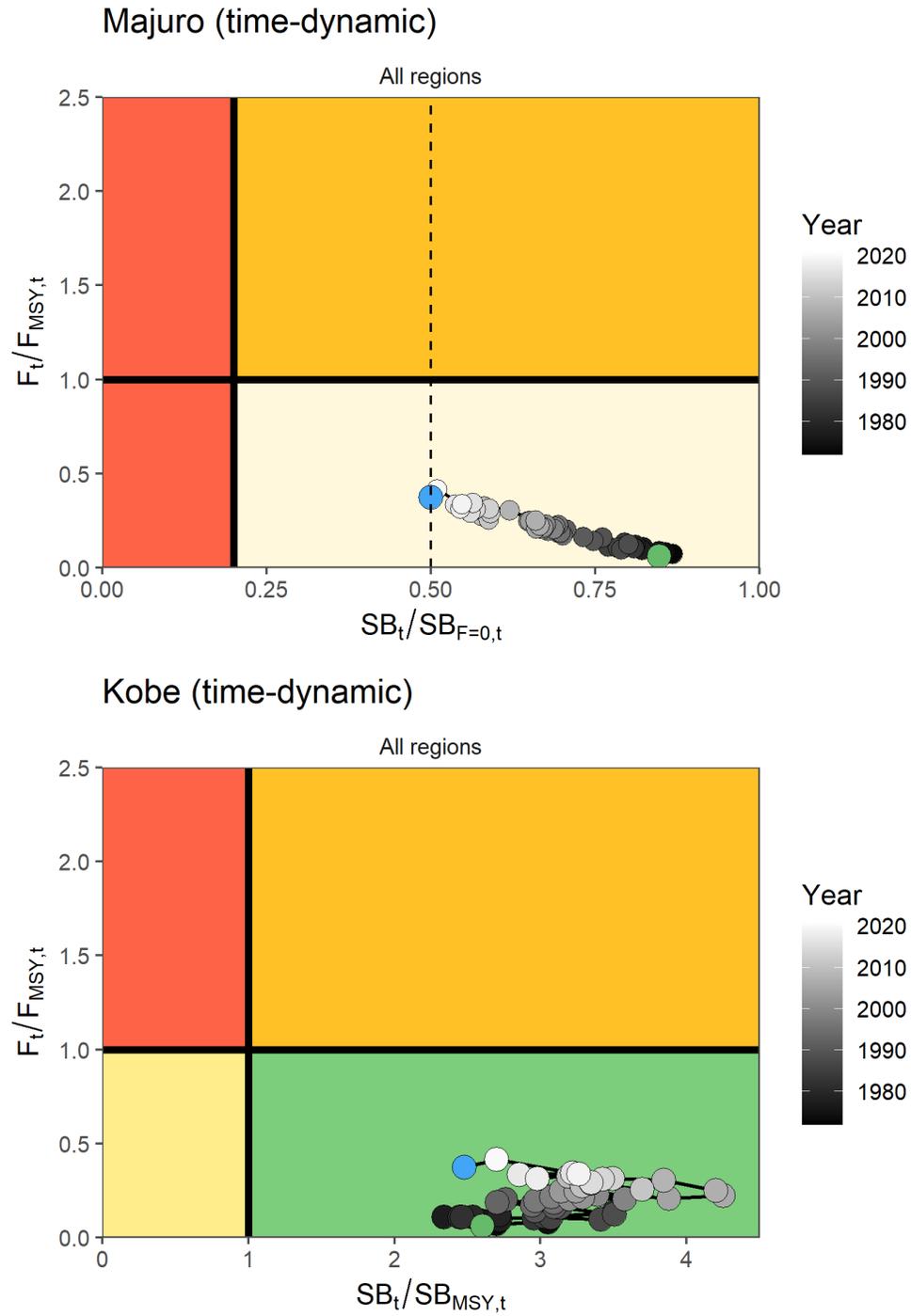


Figure 63: Time dynamic Majuro (Top) and Kobe (Bottom) plots summarising the results for the diagnostic case model over the model period. The vertical dotted line on the Majuro plot is included to indicate interim TRP of $0.50 SB_{F=0}$ for the WCPFC-CA skipjack stock. The blue point is the estimated 2021 status.

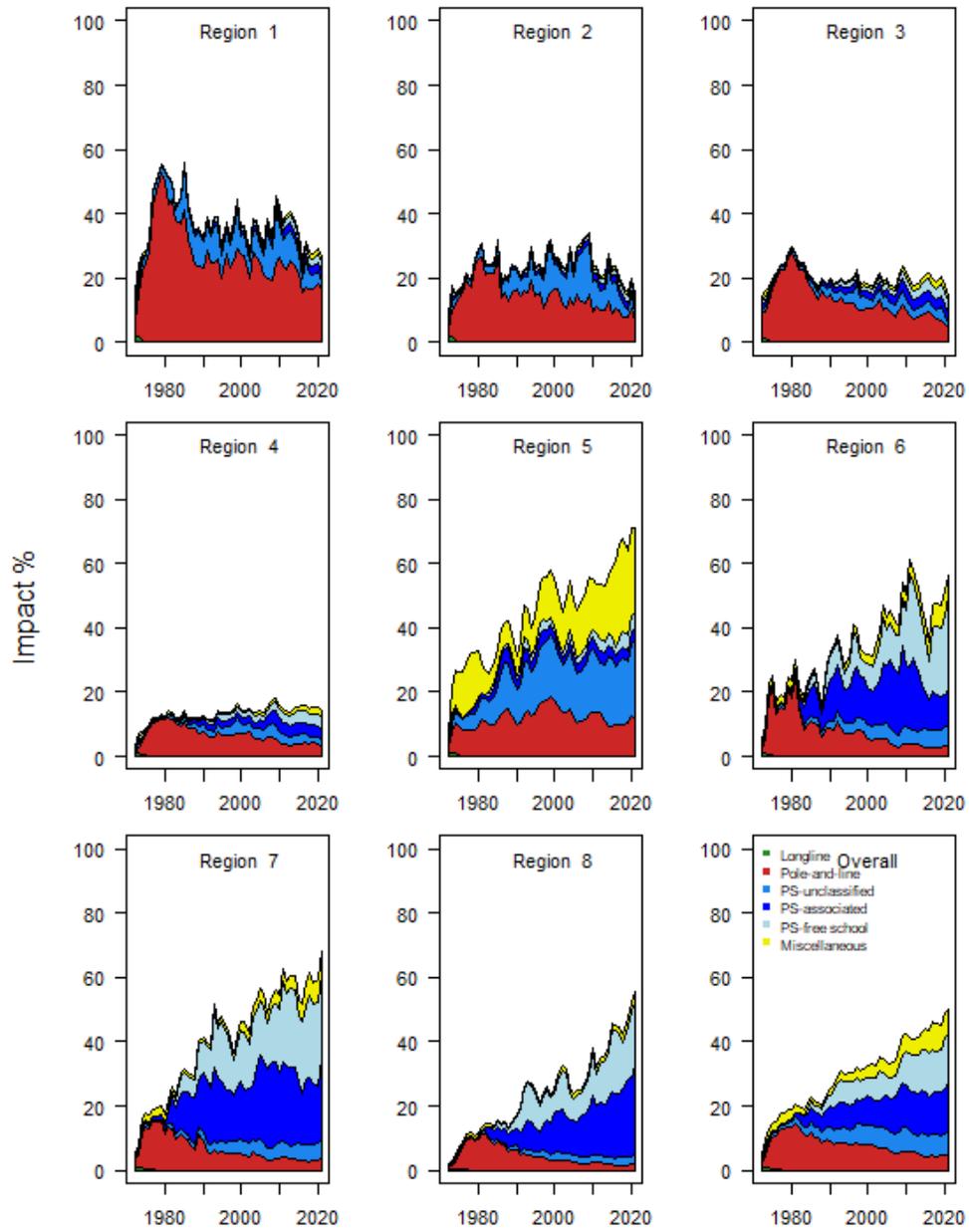


Figure 64: Estimates of reduction in spawning potential due to fishing (Fishery Impact = $1 - SB_{latest}/SB_{F=0}$) by region, and over all regions (lower right panel), attributed to various fishery groups for the diagnostic model.

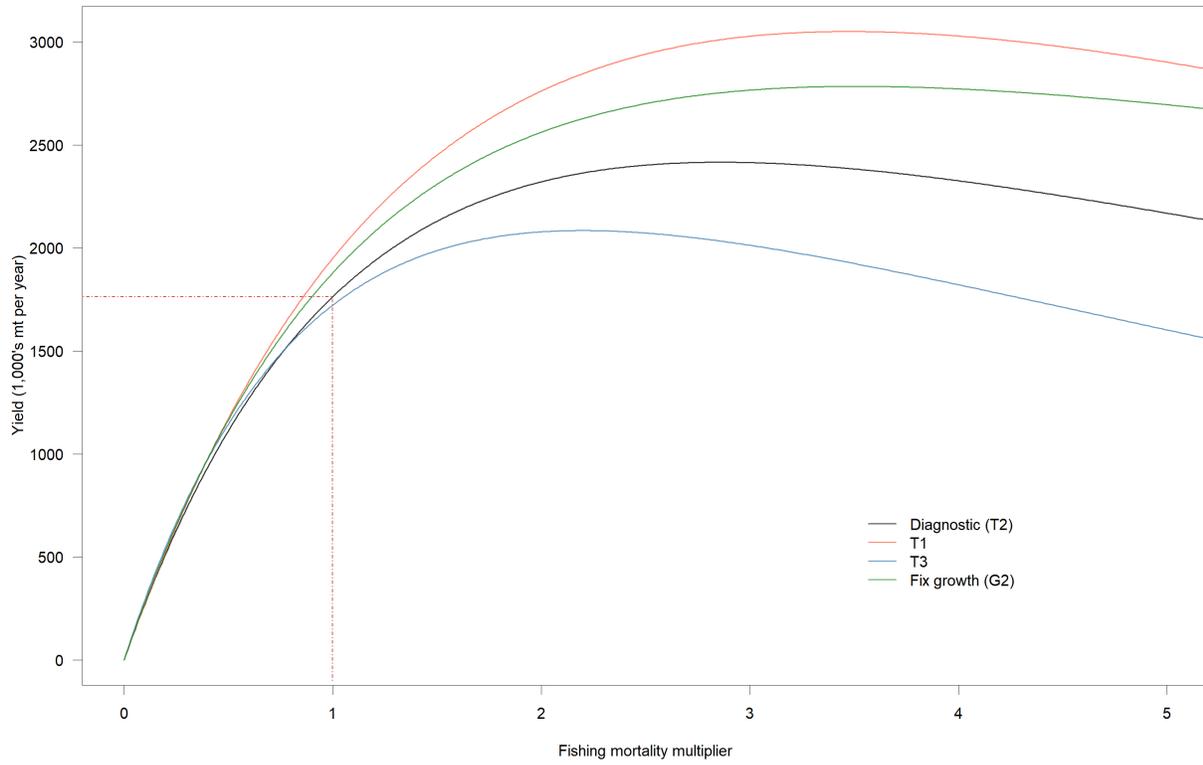


Figure 65: Estimated yield as a function of fishing mortality multiplier for the diagnostic model and a few of the one-off sensitivity models. The red dashed line indicates the equilibrium yield at current fishing mortality.

15 Appendices

15.1 Likelihood profiles

The approach for calculating a likelihood profile of the total population scaling parameter is outlined in [Section 7.6](#). The profile reflects the loss of fit over all the data, i.e., the overall objective function value, caused by changing the population scale from that of the maximum likelihood estimated value. A range of fixed values were used until the best fit for each data source was found. The likelihood profile for the diagnostic model is shown in [Figure 15.1](#) for both the phase 8 (i.e., the second to last phase without the estimation of the effort-fishing mortality regressions required for projections using the catch conditioned method) and the phase 9 (i.e., including the extra parameters for the effort-fishing mortality regressions).

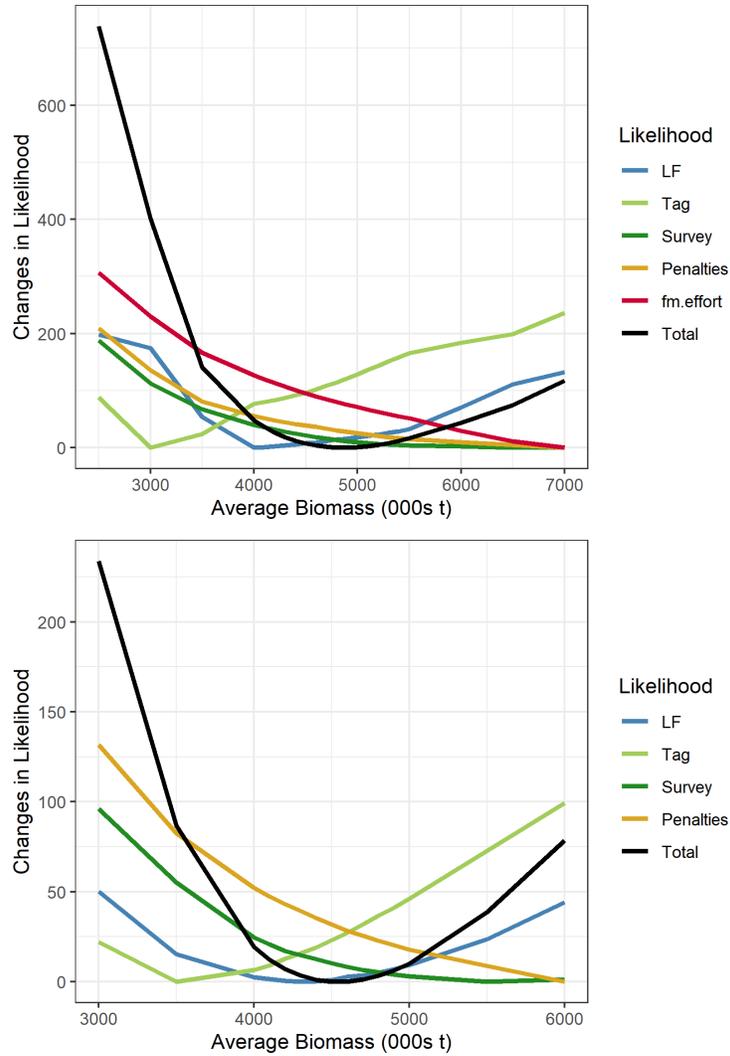


Figure 15.1: Profile of the total log-likelihood with respect to average total biomass in million mt across a range of fixed values for the model data sources, the black line indicates the total likelihood. Top plot is for phase 9 (including the extra parameters for the effort-fishing mortality regressions, fm.effort), and bottom plot is for phase 8 (excluding the extra parameters for the effort-fishing mortality regressions), survey refers to CPUE indices.

15.2 Convergence status: Jitter analyses

Jittering of model parameters' from starting values and re-running the model minimisation can be used to assess whether the model has converged to a global solution rather than a local minimum (Carvalho et al., 2021). The jitter analyses for the diagnostic model involved three components: 1. Jittering of parameters from the initial estimation phase 1 (i.e., the initial starting parameter values). 2. Jittering of parameters from phase 8 (i.e., the second to last phase without the estimation of the effort-fishing mortality regressions required for projections using the catch conditioned method). 3. Jittering of parameters from phase 9 (i.e., including the extra parameters for the effort-fishing mortality regressions). For each jitter model, a zero-centered random deviate was added to the original parameter (or a one-centered random deviate was multiplied against the original parameter depending on the link function used) to perturb the model from its initial solution. Four different standard deviations were used: 0.1, 0.01, 0.05, 0.005. All major estimated parameter groups (e.g., movement, selectivity, recruitment, tag reporting rates, selectivity, growth, M etc.) were perturbed. Each run perturbs the initial values used for minimization so the search for a minimum explores a broader region of the likelihood surface (Methot and Wetzel, 2013). For each of the 48 jitter runs 20,000 function evaluations were conducted.

The results of the jitter analyses are shown in Figure 15.2, Figure 15.3. Thirteen of the 16 models jittered in phase 1 achieved the same total likelihood as the original diagnostic case, while several other models within the 48 runs achieved substantially better likelihoods. Overall, the jitter analyses suggested that the converged model fits are somewhat sensitive to initial values and the likelihood surface has considerable complexity with numerous local minima in the region of best fit. However, estimates of $SB_{recent}/SB_{F=0}$ and $SB_{latest}/SB_{F=0}$ were robust to these alternative minima with minimal changes between each other and from the 2022 diagnostic model, regardless of the differences in total likelihood. The alternative solutions found from jittering showed more variation in the estimated levels of spawning potential, but the temporal dynamics were consistent among the models. The $SB_{recent}/SB_{F=0}$ and $SB_{latest}/SB_{F=0}$ are robust to scaling up or down the level of estimated spawning potential as the unfisher spawning potential is similarly rescaled up or down. As the management reference points are based on spawning depletion ($SB_{recent}/SB_{F=0}$), the sensitivity of the spawning potential estimates to jittering is probably of less concern than in other situations where management advice is based on estimates of spawning potential or biomass.

The best likelihood was obtained from one of the phase 9 jitters, however that model also had several negative eigenvalues. The alternative solutions from the jittering will produce different Hessians and variance covariance matrices often used to estimate model specific confidence intervals for the important dependent variables such as $SB_{recent}/SB_{F=0}$. Due to some negative eigenvalues, for this assessment we elected not to use the variance covariance matrices to estimate model specific estimation uncertainty (i.e. CV 0.02 for the diagnostic case). However, based on the example in Section 15.3 this could potentially be reconsidered in future. The structural uncertainty grid alone is used to characterise uncertainty in the key management reference points.

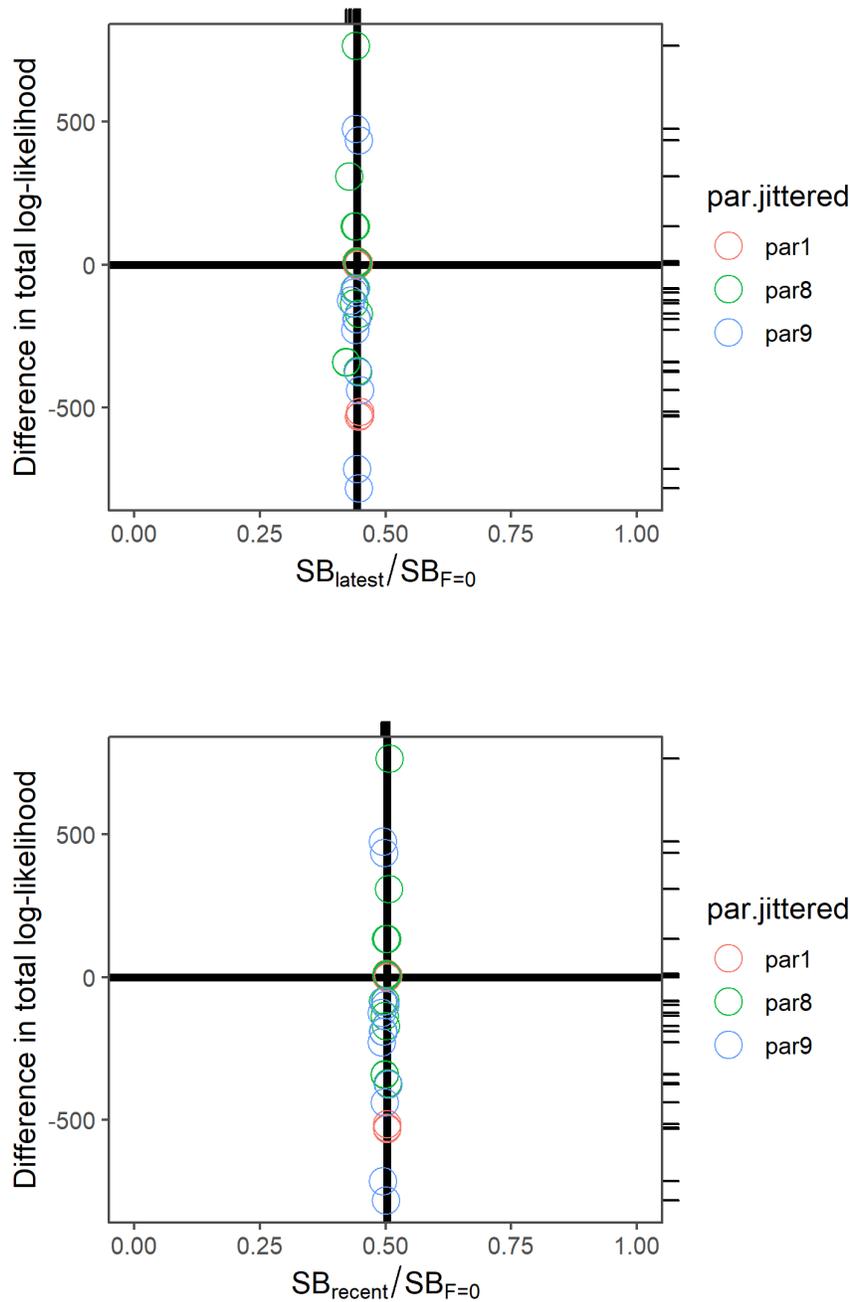


Figure 15.2: Results from the jitter analyses of the diagnostic model. (Top) $SB_{latest}/SB_{F=0}$ (Bottom) $SB_{recent}/SB_{F=0}$. Points below the horizontal line indicate better relative likelihood values (i.e., diff likelihood = diag case - (new fit)). The vertical line indicates the estimate of $SB_{latest}/SB_{F=0}$ and $SB_{recent}/SB_{F=0}$ from the diagnostic model. The horizontal line represents the diagnostic case likelihood. A rug plot is included in the margins to show the positions of points that are plotted on top of one-another. par 1 = jitters from phase 1 parameters, par 8 = jitters from phase 8 parameters, par 9 = jitters from phase 9 parameters.

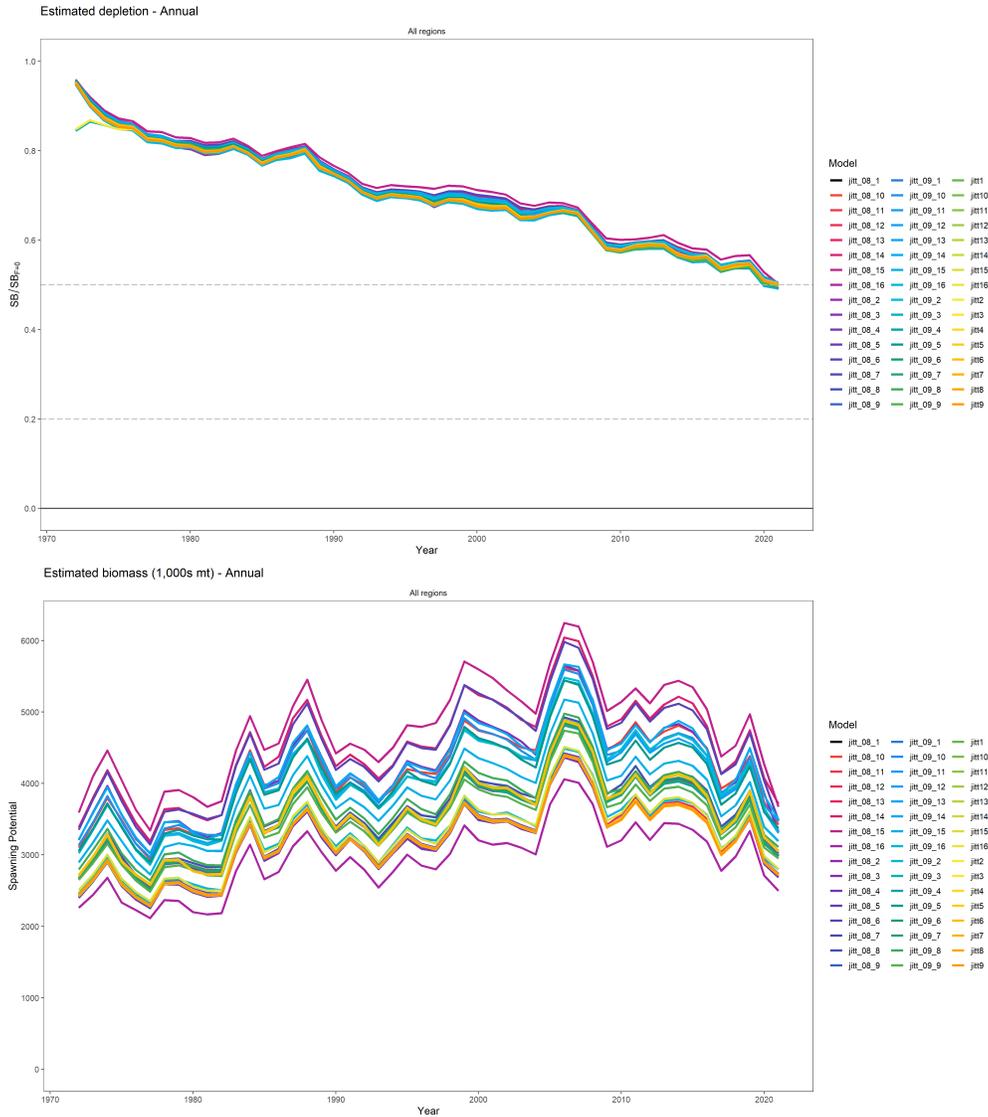


Figure 15.3: Estimated trajectories of $SB/SB_{F=0}$ (Top), and spawning potential (Bottom) for each of the 48 Jitter models. jitt-08 = jitters from phase 8 parameter, jitt-09 = jitters from phase 9 parameters, jitt = jitters from phase 1 parameters.

15.3 Hessian diagnostic

As noted in Section 7.6 and discussed further in Section 10.3.1, we were unable to obtain a positive definite Hessian (PDH) matrix for the estimated model parameters, as would be indicated by the absence of negative eigenvalues, for any of our grid models. This indicates that it is likely that the models have not converged at a global minima of the penalised negative log-likelihoods. This may not be particularly unusual for complex models such as this but indicates that the minimisation algorithm may not have fully converged to a local minimum in this vicinity of the parameter space. To further explore this issue and the implications, we undertook; 1. ‘jitter’ analysis of the diagnostic case model described in Section 15.2, to see if better solutions could be obtained by varying the starting values of parameters; and, 2. explored previous comparisons of PDH and non-PDH solutions for a bigeye tuna model (2020 assessment) to see if any substantive difference occurs in the estimate of the key stock status reference point $SB_{recent}/SB_{F=0}$ and its standard deviation. In this section we focus on the likely implications of the non-PDH solutions for estimates of $SB_{recent}/SB_{F=0}$ and its standard deviation (SD). We use 2020 bigeye tuna model because despite considerable effort we have been unable to obtain an example of a PDH solution for the current skipjack assessment to make a direct comparison between estimates from equivalent models with and without a PDH. We continue efforts to develop a model with a PDH for skipjack, however, this does not seem feasible before the SC18 meeting. For the 2020 bigeye model example, the estimated values and variance estimates of $SB_{recent}/SB_{F=0}$ for the PDH and non-PDH solutions are shown in Table 15.5, and expressed in normal space with the confidence intervals of ± 2 SD.

Table 15.5: Comparison of estimates of $SB_{recent}/SB_{F=0}$ and their respective SD estimated for 2020 bigeye tuna models that have PDH and non-PDH solutions.

	Log ($SB_{recent}/SB_{F=0}$)	Standard deviation (SD)	($SB_{recent}/SB_{F=0}$)	($SB_{recent}/SB_{F=0}$) ± 2 *SD
non-PDH	-8.62233000e-01	1.38528909e-01	0.42	0.32 –0.55
PDH	-8.62214000e-01	2.96706066e-02	0.42	0.39–0.44

We note that for this bigeye tuna example, the point estimates of $SB_{recent}/SB_{F=0}$ from the non-PDH and PDH cases are virtually identical. However, the SD in normal space for the non-PDF case is larger than for the PDH solution. From this example we might expect attaining a PDH for the skipjack diagnostic model would result in a smaller CV and therefore tighter confidence intervals. In the case of the current skipjack diagnostic model, we have been able to generate a standard deviation report for the derived parameters, including the $SB_{recent}/SB_{F=0}$ management reference point variable. The estimated CV for $SB_{recent}/SB_{F=0}$ is approximately 0.02. If the bigeye example is indicative for the skipjack model, then we would expect an even smaller CV for a model with PDH.

15.4 Retrospective analyses

Retrospective analysis involves rerunning the 2022 diagnostic model by consecutively removing successive years of data to estimate model bias. A series of 10 retrospective models were fitted starting with the full data-set (through 2021), followed by models with the retrospective removal of all input data. Spawning potential and spawning depletion trajectories are shown in [Figure 15.4](#). Mohns rho is calculated for the retrospective models of spawning depletion.

The models with each year of data removed sequentially produced estimates of spawning potential and spawning depletion with very similar temporal dynamics to the full diagnostic model. Some retrospective bias begins to occur for the years starting 2012 and 2013, but only for the recent years of those trajectories. Mohns rho for the spawning depletion retrospectives is -0.097, and for spawning potential, is -0.154 ([Figure 15.4](#)), indicative of a low and acceptable retrospective bias for a relatively short lived species like skipjack ([Hurtado-Ferro et al., 2015](#)).

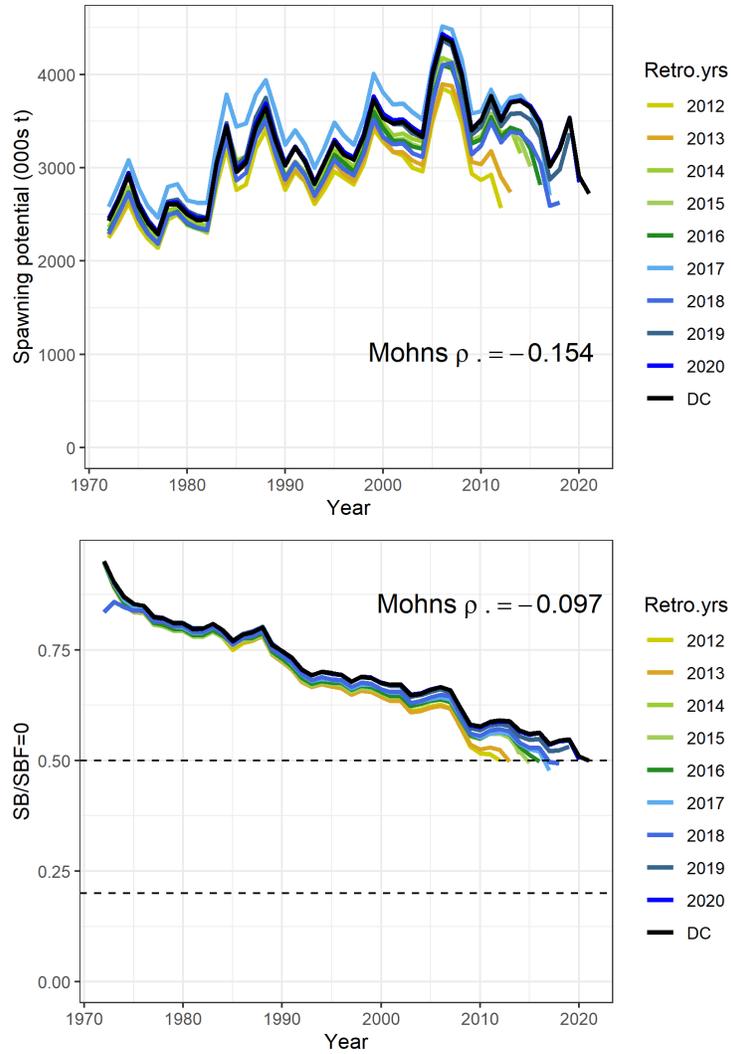


Figure 15.4: Estimated spawning potential (Top), and spawning depletion (Bottom) for each of the retrospective models.

15.5 ‘Status quo’ deterministic stock projections for WCPO skipjack tuna

To evaluate the potential consequences of future fishing levels for the skipjack stock, based on the uncertainty grid of 18 models developed within the 2022 assessment, stock projections were performed. At present, the new ‘catch conditioned’ formulation of the skipjack assessment model developed using MULTIFAN-CL does not allow stochastic projections to be performed. Stock projections were therefore deterministic, with future recruitment assumed to represent the average level defined by the stock recruitment relationship for a given underlying adult biomass.

Three scenarios for future fishing were examined:

- ‘2021’ fishing levels, where the stock was projected under 2021 fishing levels
- ‘2012’ fishing levels, where the stock was projected under 2012 fishing levels
- ‘2012+’ fishing levels, where the stock was projected under 2012 fishing levels, except for the domestic fisheries of Indonesia, Philippines and Vietnam which were projected at 2016-18 average levels (consistent with the analyses of SC18-MI-IP-09).

Purse seine fisheries were projected based upon effort; all other fisheries were projected based on catch. Catchability was assumed to remain constant in the projection period. The lognormal bias correction to the stock recruitment relationship was enabled.

Deterministic projection results were combined across the 18 models. The outputs of the projections are median depletion ($SB_{year}/SB_{F=0}$) for specific years of the projection period and F/F_{MSY} at the end of the projection period. Risk $SB_{2051}/SB_{F=0} < LRP$ and $F > F_{MSY}$ are also calculated (Table 15.6). However, it should be noted that as this is based upon the outputs from deterministic projections, risk is likely to be underestimated.

We stress that the results of these deterministic projections should be used as a guide. Stochastic projections will be performed based on SC18 decisions for the skipjack uncertainty grid, once the functionality is available within MULTIFAN-CL.

Results of the projections are summarized in the table and figures of spawning depletion below.

Table 15.6: Summary of median WCPO skipjack tuna stock outcomes under alternative average future fishing conditions.

Fishing level	$SB_{2025}/SB_{F=0}$	$SB_{2035}/SB_{F=0}$	$SB_{2051}/SB_{F=0}$	Risk $SB_{2051}/SB_{F=0}$ < LRP	F/F_{MSY}	Risk $F > F_{MSY}$
2021	0.51	0.56	0.56	0%	0.28	6%
2012	0.48	0.52	0.52	0%	0.31	6%
2012+	0.47	0.51	0.51	0%	0.32	6%

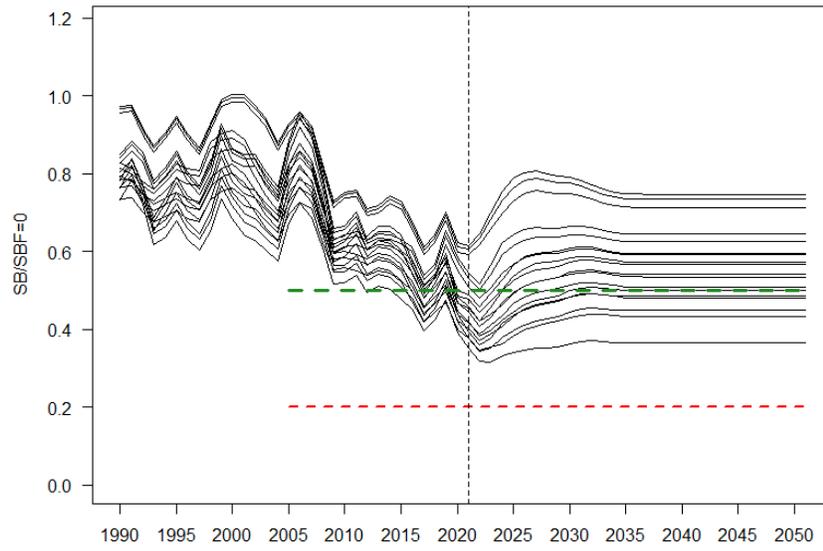


Figure 15.5: Estimated spawning depletion trajectories for each of the 18 grid models where the stock was projected under 2021 fishing levels. Red dashed line is the limit reference point 20% $SB_{F=0}$, green dashed line is the interim target reference point as specified in CMM 2021-01 50% $SB_{F=0}$, vertical dashed line is the terminal year of the 2022 assessment.

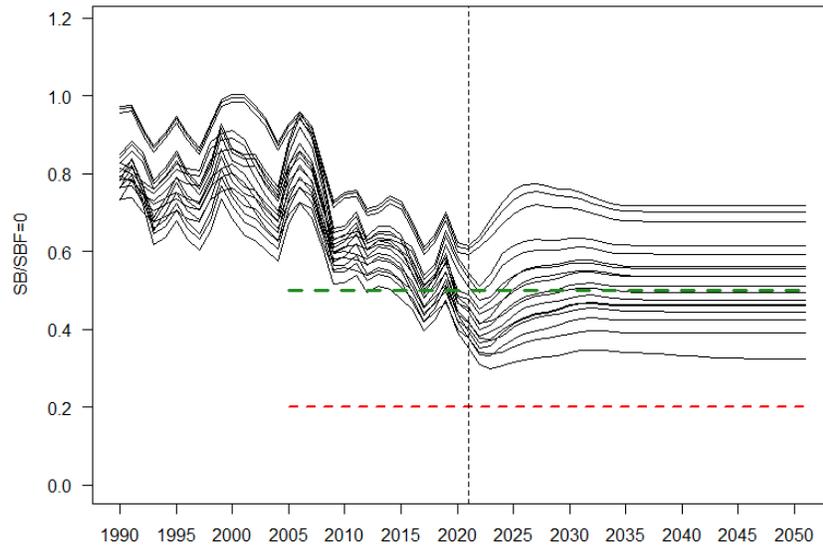


Figure 15.6: Estimated spawning depletion trajectories for each of the 18 grid models where the stock was projected under 2012 fishing levels. Red dashed line is the limit reference point 20% $SB_{F=0}$, green dashed line is the interim target reference point as specified in CMM 2021-01 50% $SB_{F=0}$, vertical dashed line is the terminal year of the 2022 assessment.

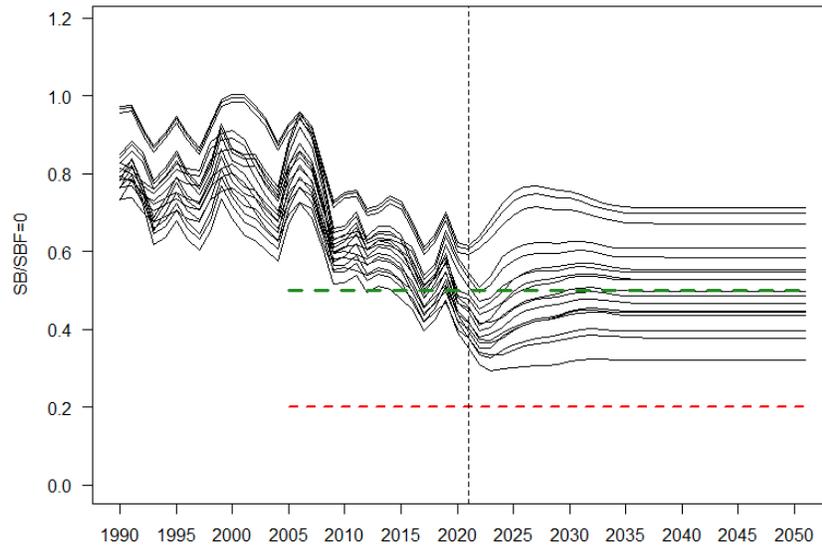


Figure 15.7: Estimated spawning depletion trajectories for each of the 18 grid models where the stock was projected under 2012 fishing levels, except for the domestic fisheries of Indonesia, Philippines and Vietnam which were projected at 2016-18 average levels. Red dashed line is the limit reference point 20% $SB_{F=0}$, green dashed line is the interim target reference point as specified in CMM 2021-01 50% $SB_{F=0}$, vertical dashed line is the terminal year of the 2022 assessment.

15.6 5 region model

A simplified 5-region model was developed in response to recommendations from the PAW (Hamer, 2022). This involved collapsing regions 1-4 of the 8 region model into one large northern region (Figure 15.8). The fishery structures were maintained, but the size composition re-weighting and estimation of the pole-line-survey index were recalculated at the scale of single northern model region. The tag mixing simulations could not be rerun for the larger model region in the time available, therefore the median values across the simulated release groups for the JPTP were applied for each of the KS D values used in the uncertainty grid. Other model settings were consistent with the main 8 region assessment model. The 5 region model was run with the same structural uncertainty grid as the 8 region assessment model. A summary of the main results are provided in the figures and tables below. The summary of reference points (Table 15.1) indicates similar estimates of the management reference points to the 8 region model (Table 4).

Comparison of 5 and 8 region diagnostic models: Similar to the 8 region diagnostic model, the 5 region diagnostic model displays an overall increasing trend in estimated recruitment from 1980 onwards, and this is largely driven by trends in regions 5 and 7 (Figure 15.15, Figure 15.16). The 5 region diagnostic model estimates lower spawning potential than the 8 region diagnostic model, particularly from 1990 onwards when it shows less of an increasing trend, but the temporal dynamics are similar. For spawning depletion, the 5 and 8 region diagnostic models estimate very similar trends and levels, but prior to 1990 the 5 region model estimates the stock is less depleted than the 8 region model, and after 1990, more depleted (Figure 15.17).

Table 15.1: Summary of reference points over the 18 individual models in the structural uncertainty grid for the 5 region model

	Mean	Median	Min	10%	90%	Max	Diag case
C_{latest}	1530209	1530208	1530207	1530207	1530212	1530212	1530207
F_{MSY}	0.23	0.23	0.18	0.19	0.27	0.29	0.26
f_{mult}	2.93	2.52	1.78	2.01	4.37	5.97	2.41
F_{recent}/F_{MSY}	0.38	0.40	0.17	0.23	0.50	0.56	0.41
MSY	2442777	2197800	1830000	1952360	3850040	3874000	2058800
SB_0	6708055	6114000	4191000	4447200	10602000	11990000	4583000
$SB_{F=0}$	7103224	6862191	4909138	5071170	10625443	11105981	5126164
SB_{latest}/SB_0	0.45	0.46	0.36	0.39	0.51	0.55	0.46
$SB_{latest}/SB_{F=0}$	0.42	0.42	0.34	0.36	0.52	0.54	0.41
SB_{latest}/SB_{MSY}	2.52	2.50	1.60	1.82	3.41	3.74	2.29
SB_{MSY}	1251144	1141000	727900	767510	1703400	2547000	913600
SB_{MSY}/SB_0	0.19	0.19	0.13	0.15	0.22	0.23	0.20
$SB_{MSY}/SB_{F=0}$	0.17	0.18	0.12	0.14	0.22	0.23	0.18
$SB_{recent}/SB_{F=0}$	0.49	0.49	0.40	0.42	0.59	0.61	0.48
SB_{recent}/SB_{MSY}	2.92	2.86	1.96	2.09	3.88	4.20	2.72
Y_{Recent}	1761022	1757800	1510200	1579640	1950200	2161200	1637200

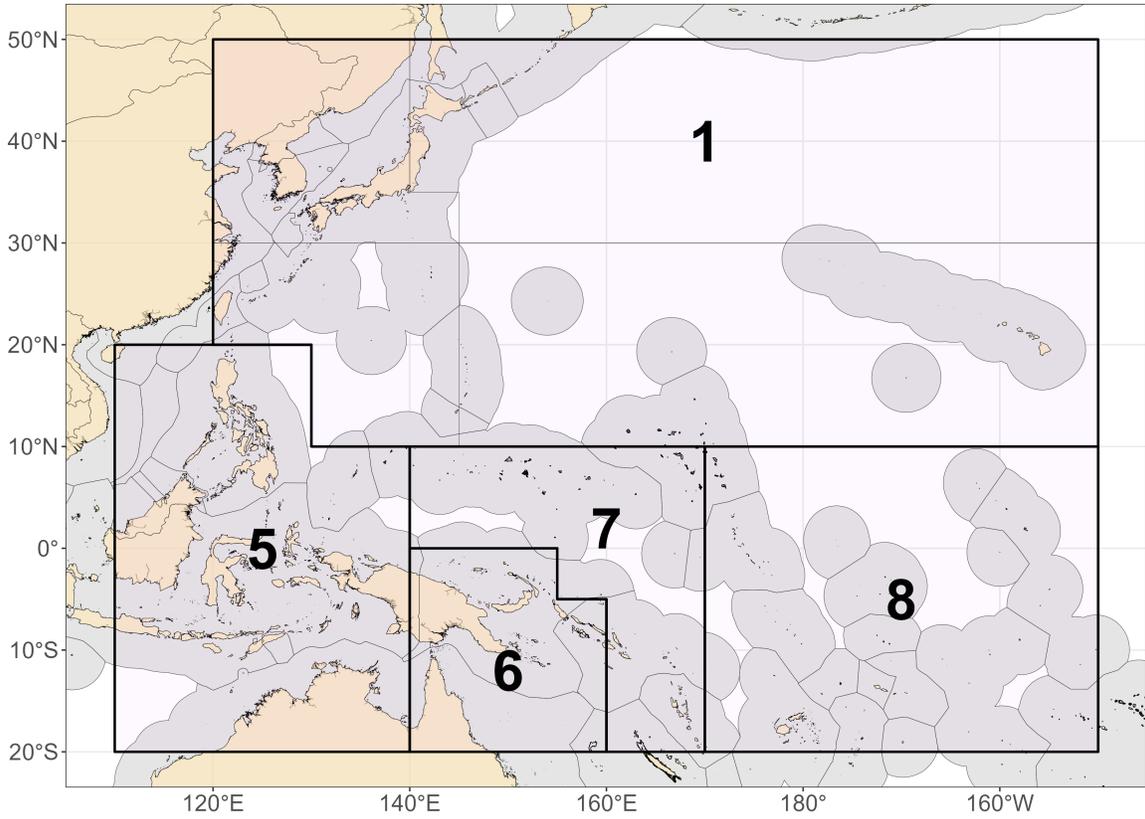


Figure 15.8: Map showing region structure of the 5 region model.

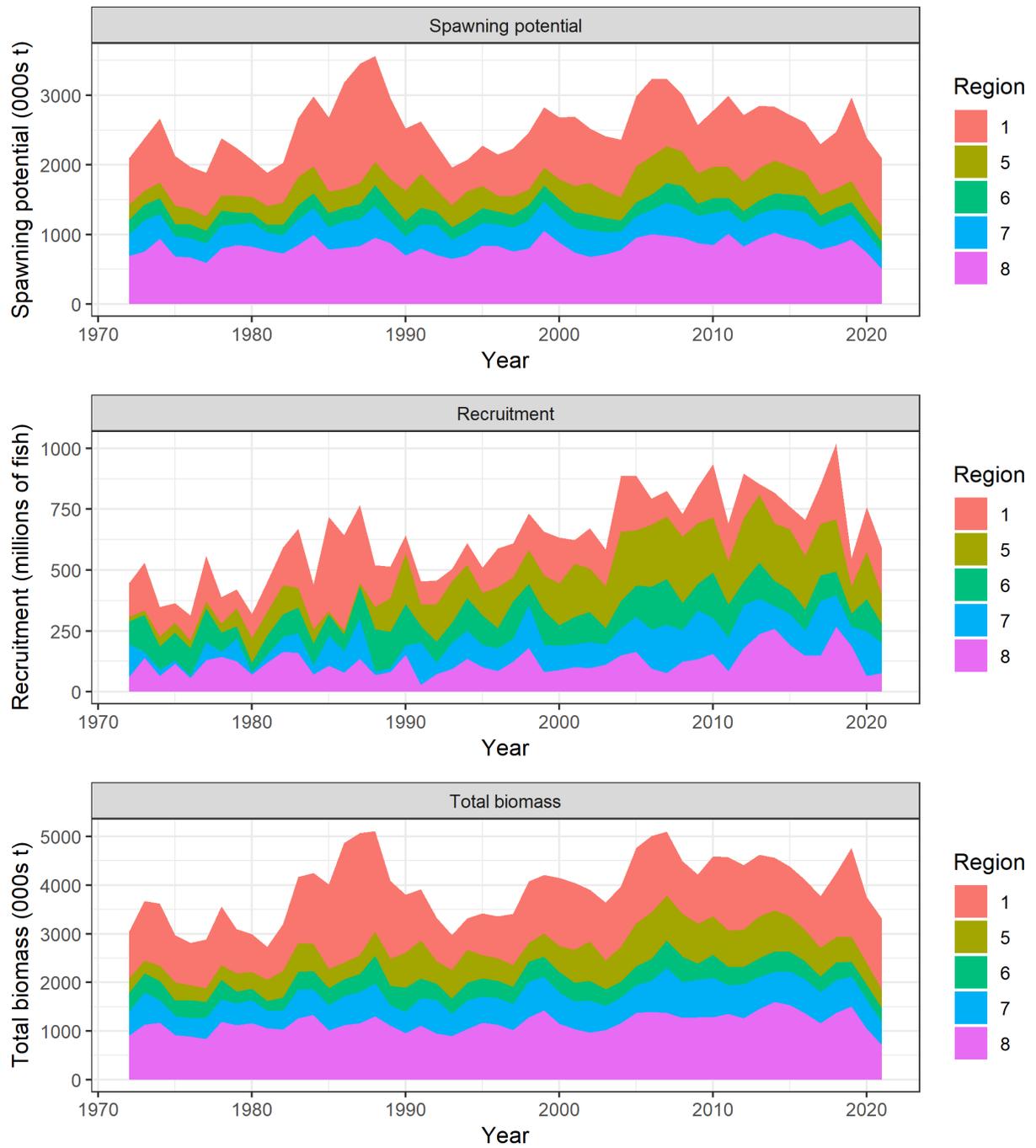


Figure 15.9: Estimated annual average recruitment, spawning potential and total biomass by model region for the 5 region diagnostic model, showing the relative sizes among regions.

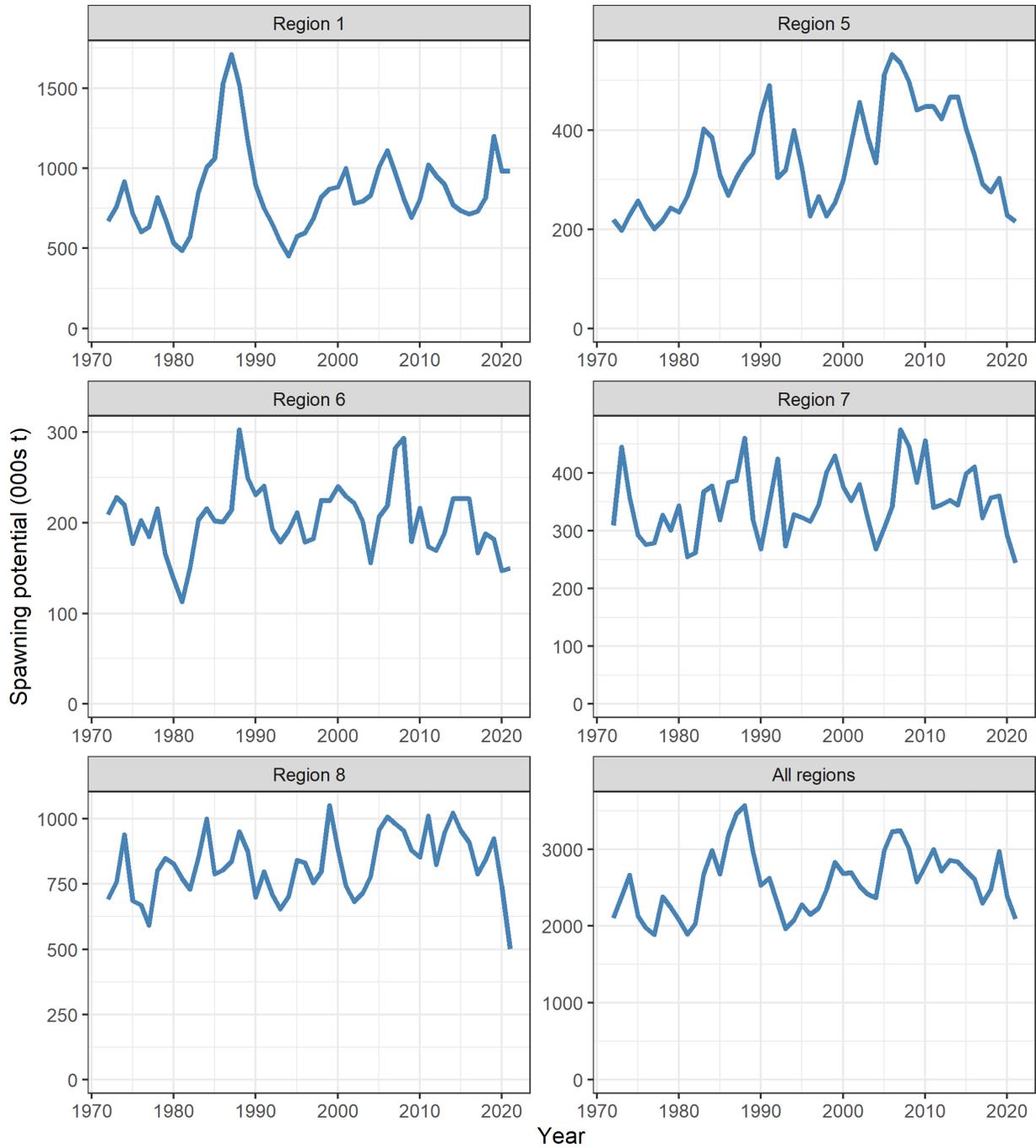


Figure 15.10: Estimated temporal spawning potential by model region, and for all model regions summed for the diagnostic 5 region model. Note that the scale of the y-axis is not constant across regions.

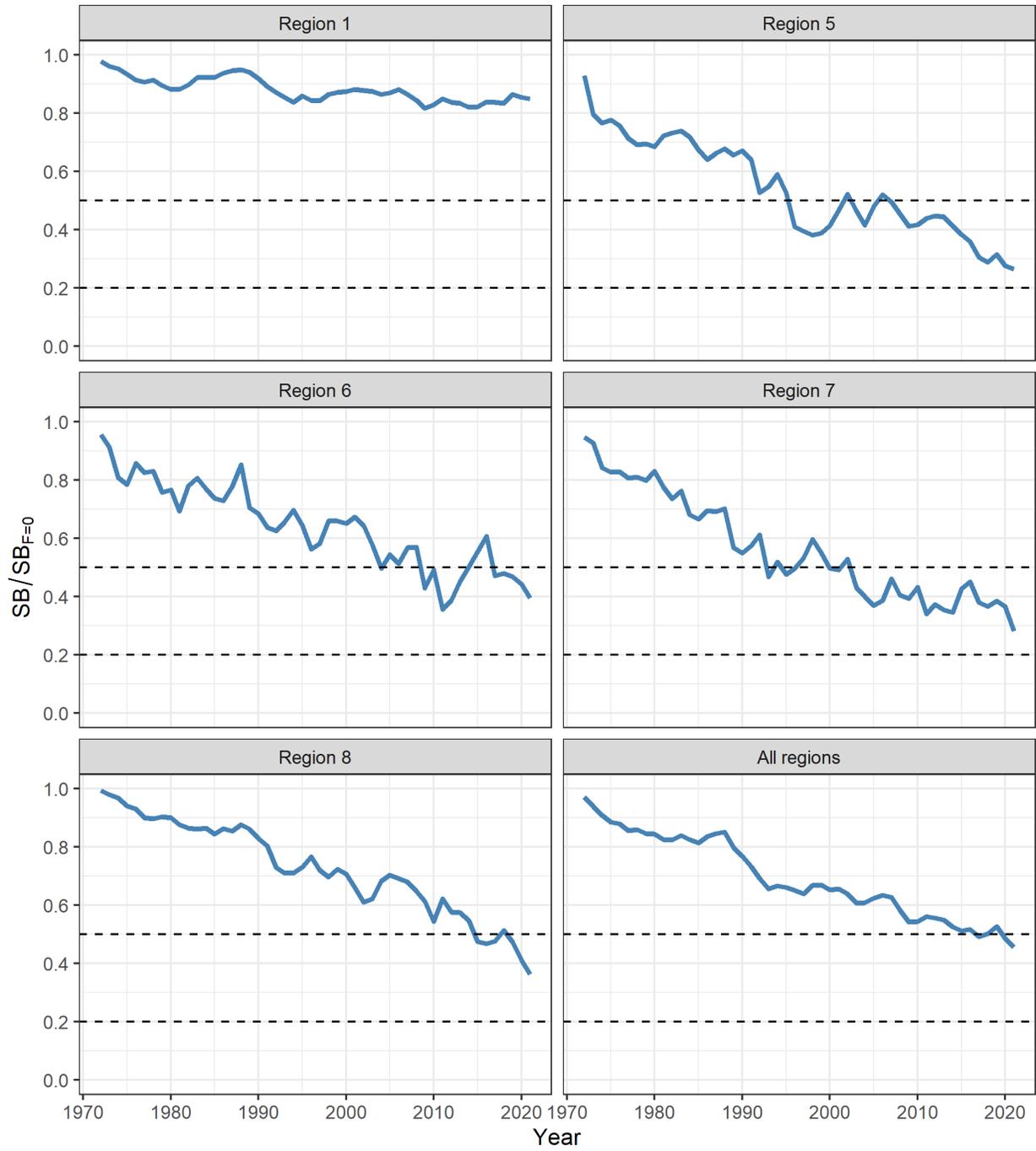


Figure 15.11: Estimated annual spawning potential depletion by model region, and for all model regions summed for the diagnostic model.

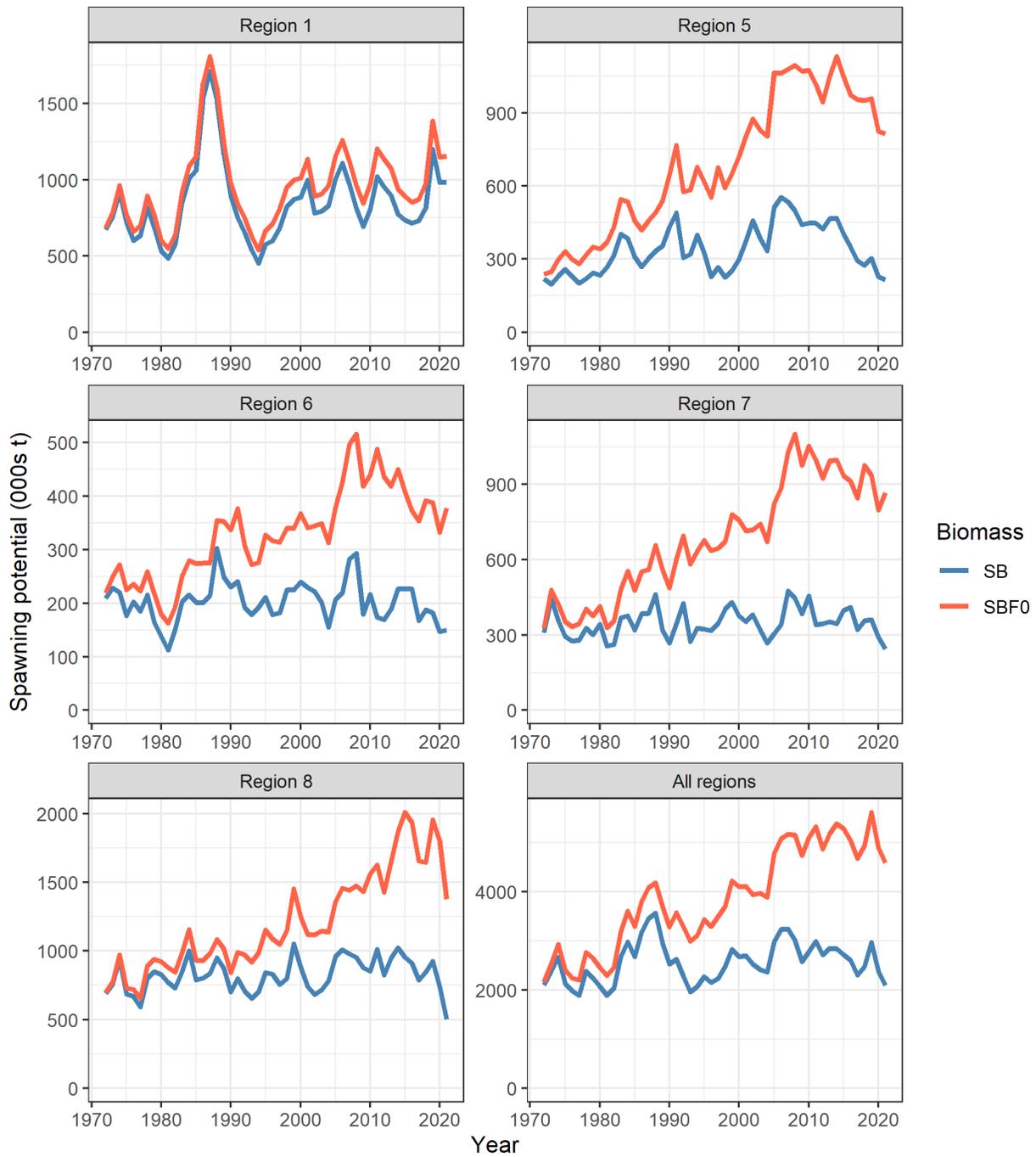


Figure 15.12: Comparison of the estimated annual spawning potential trajectories (lower blue lines) with the spawning potential trajectories predicted to have occurred in the absence of fishing (upper red lines) for each region and overall, for the diagnostic 5 region model. Note the scales of the Y-axes vary.

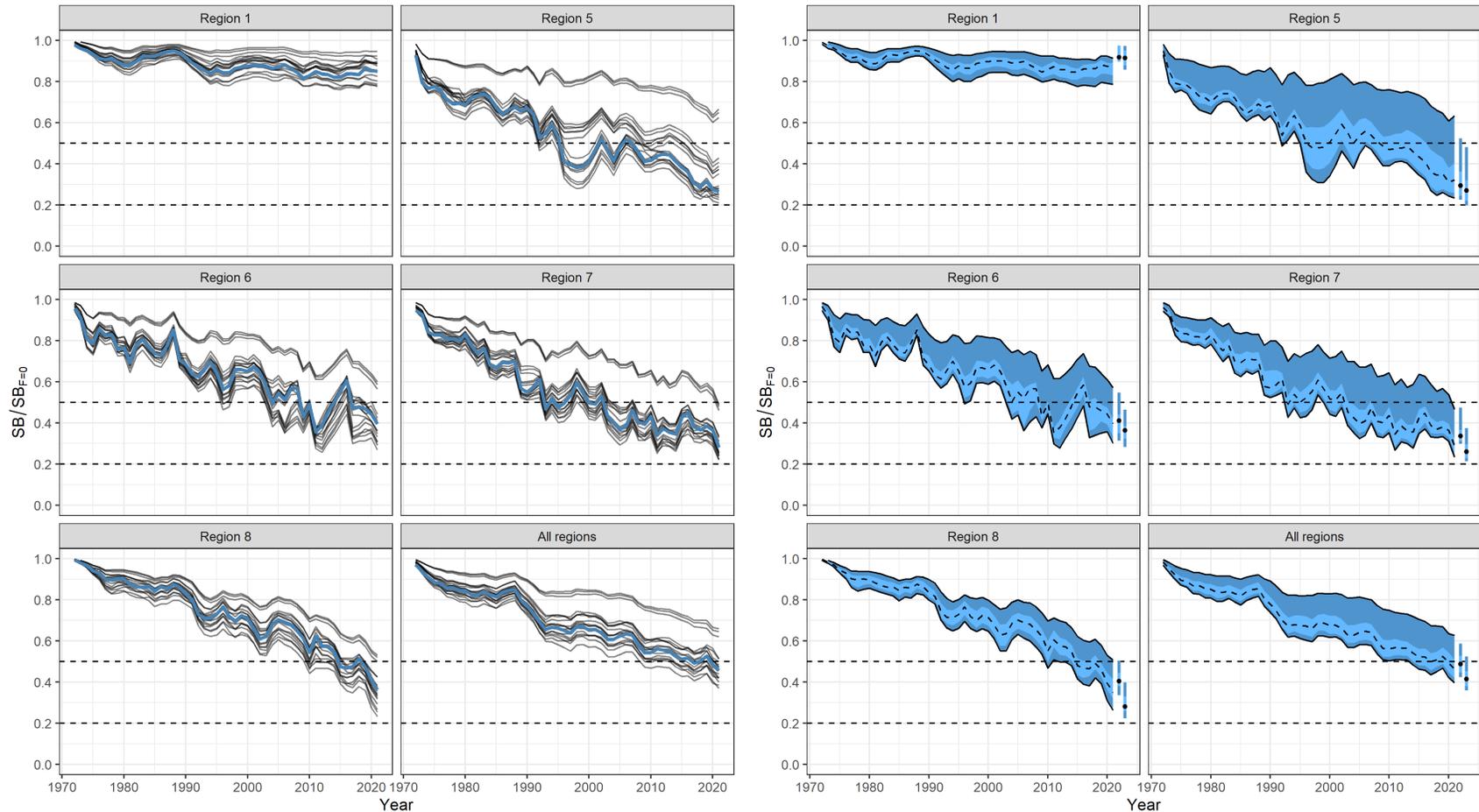


Figure 15.13: 5 region model: (Left) Trajectories of spawning depletion for the individual model runs included in the structural uncertainty grid over the period 1972-2021. (Right) Estimated spawning depletion across all models in the structural uncertainty grid over the period 1972-2021. The dashed line represents the median. The lighter band shows the 50th percentile, and the dark band shows the 80th percentile of the model estimates. The bars at the right of each ribbon indicate the median (black dots) and 80th percentile range for (left bar) $SB_{recent}/SB_{F=0}$ and (right bar) $SB_{latest}/SB_{F=0}$.

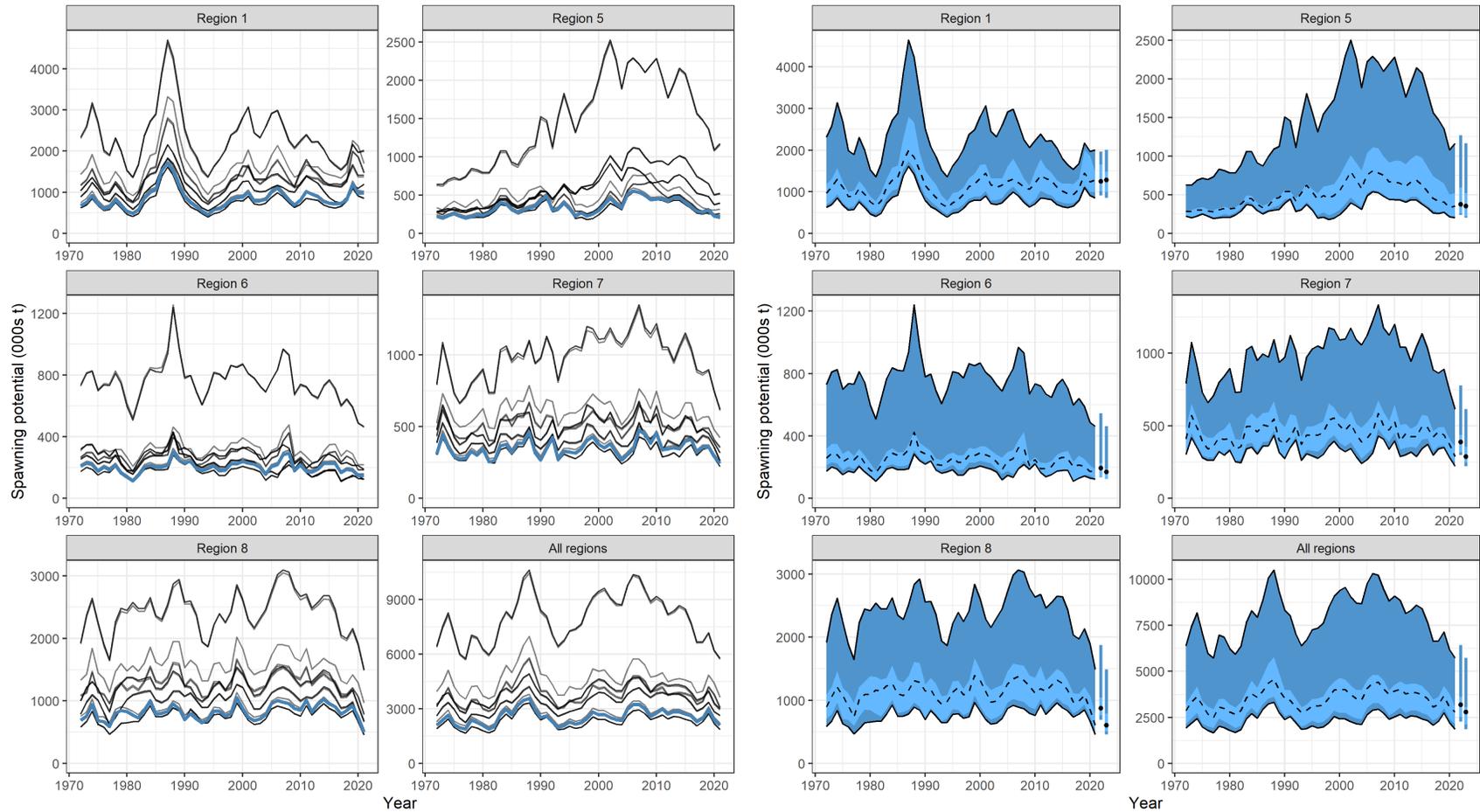


Figure 15.14: 5 region model: (Left) Trajectories of spawning potential for the individual model runs included in the structural uncertainty grid over the period 1972-2021.(Right) Estimated spawning potential across all models in the structural uncertainty grid over the period 1972-2021. The dashed line represents the median. The lighter band shows the 50th percentile, and the dark band shows the 80th percentile of the model estimates. The bars at the right of each ribbon indicate the median (black dots) and 80th percentile range for (left bar) SB_{recent} and (right bar) SB_{latest} .

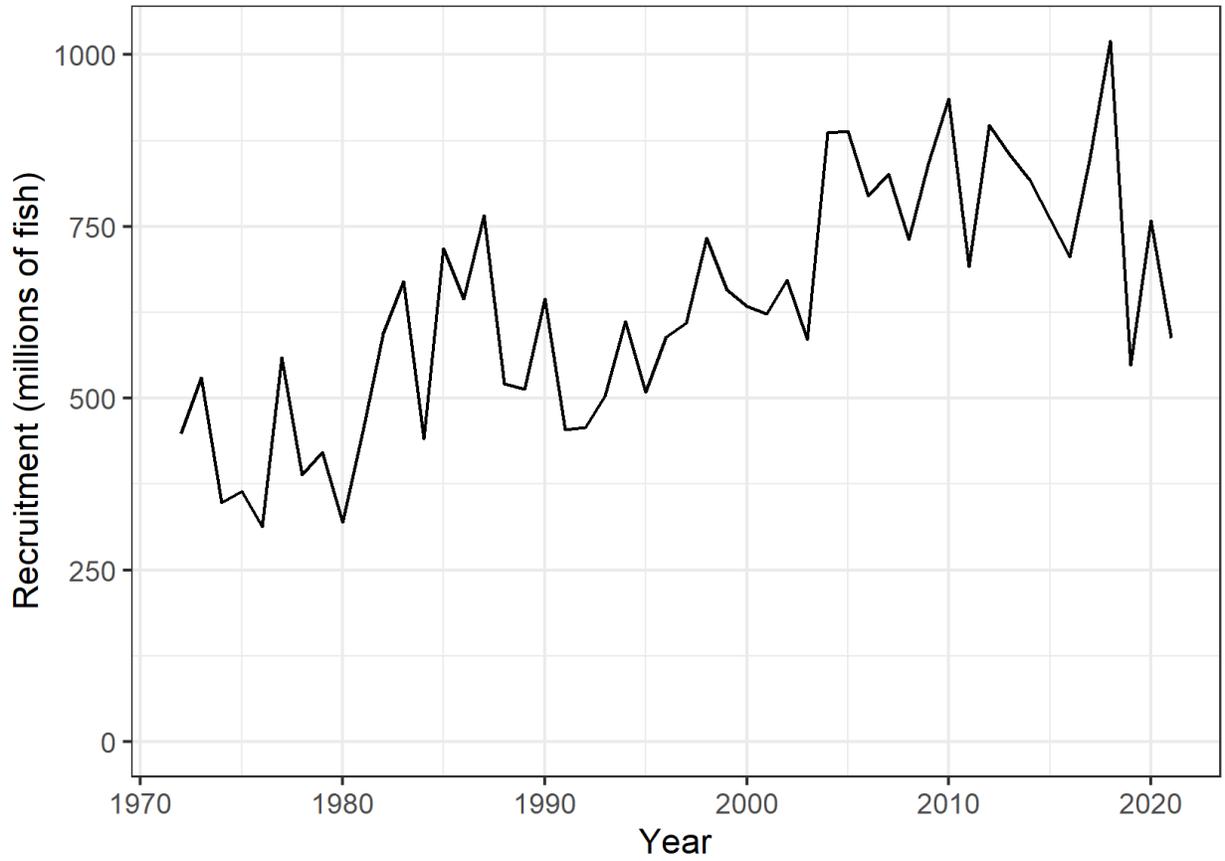


Figure 15.15: Estimated annual (summed across quarters) recruitment summed across regions for the 5 region diagnostic model.

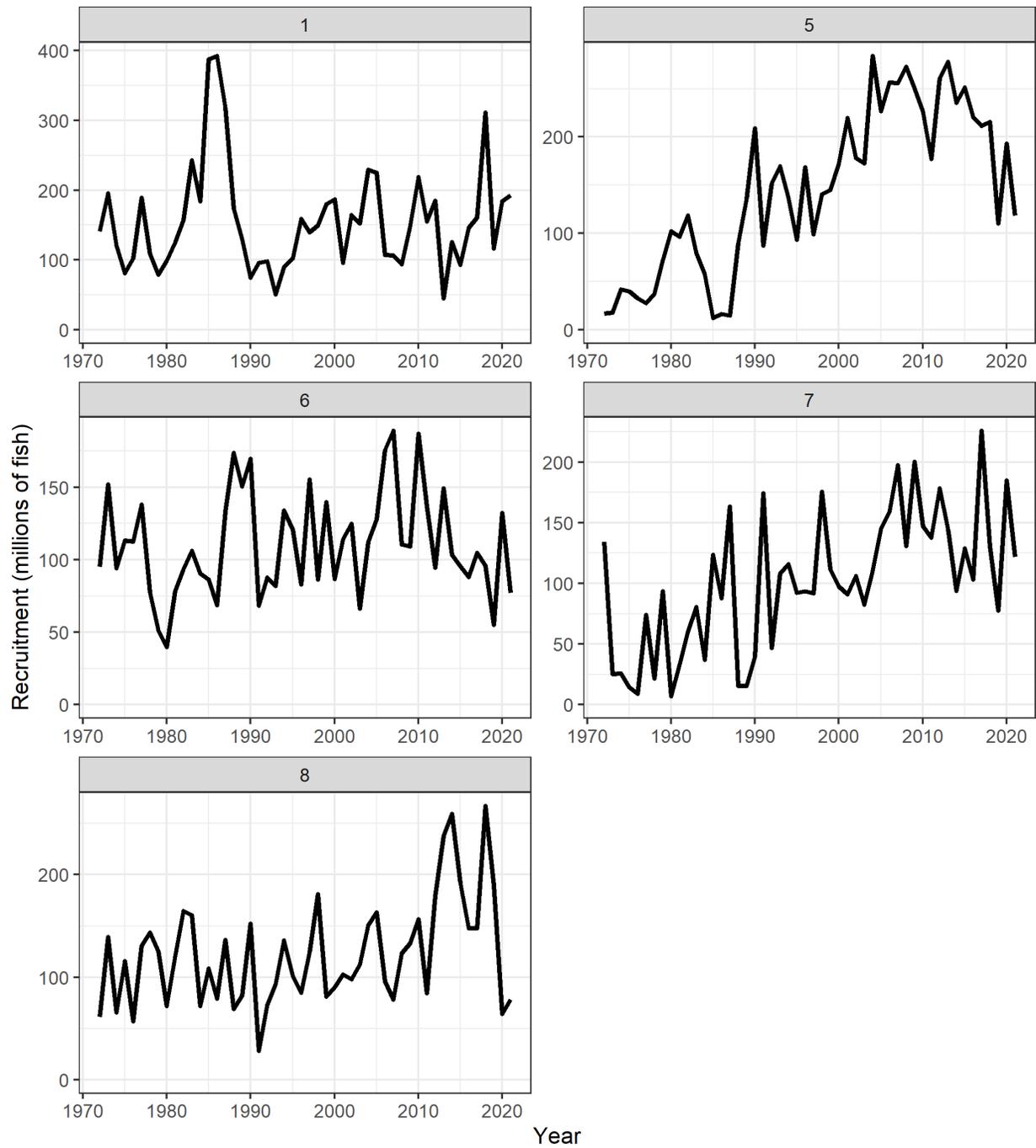


Figure 15.16: Estimated annual (summed across quarters) recruitment by model region for the 5 region diagnostic model. Note that the scale of the y-axis is not constant across regions.

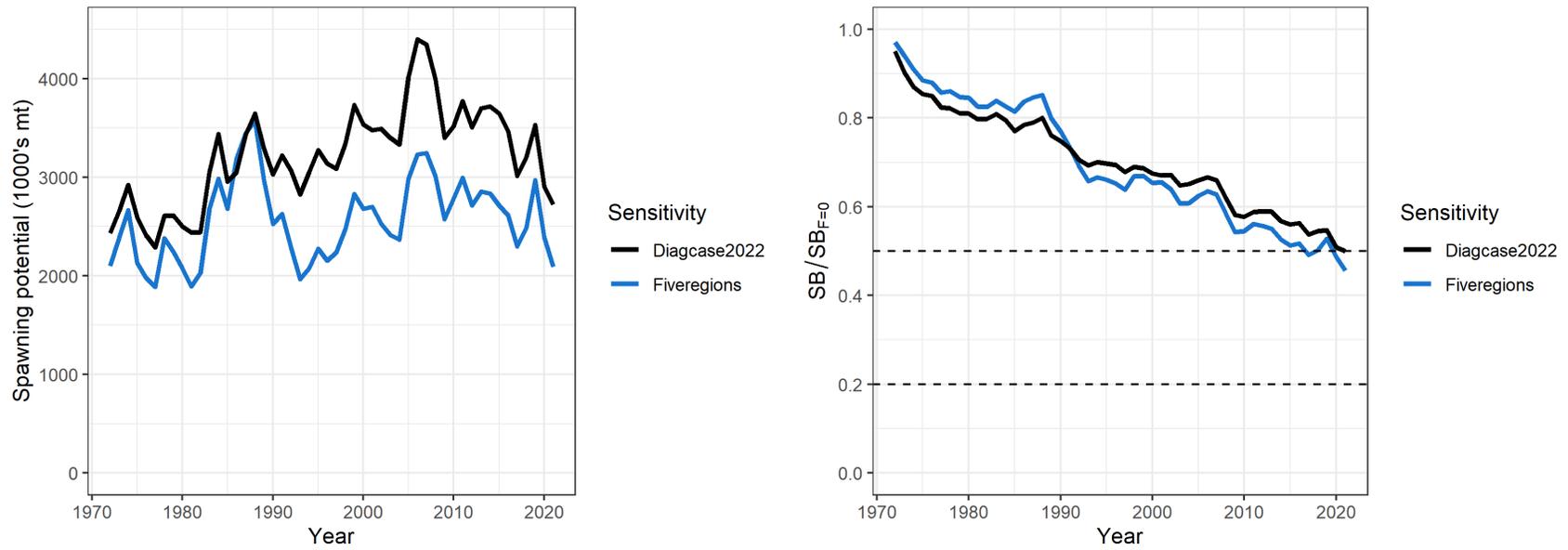


Figure 15.17: Comparisons of the trajectories of spawning potential (Left) and spawning depletion (Right) across all model regions combined for the 8-region diagnostic model (Diagcase2022) and 5-region (Fiveregions) diagnostic model.