

**PRELIMINARY INDIAN OCEAN YELLOWFIN TUNA STOCK
ASSESSMENT 1950-2020 (STOCK SYNTHESIS)**

**PREPARED BY: DAN FU¹, AGURTZANE URTIZBEREA IJURCO², MASSIMILIANO CARDINALE³, RICHARD
METHOT⁴, SIMON HOYLE⁵, GORKA MERINO⁶**

02 OCTOBER 2021

¹ IOTC Secretariat, Dan.Fu@fao.org;

² AZTI, aurtizberea@azti.es;

³ SLU, massimiliano.cardinale@slu.se;

⁴ NOAA, richard.methot@noaa.gov;

⁵ NIWA, simon.hoyle@gmail.com;

⁶ AZTI, gmerino@azti.es;

Contents

1. INTRODUCTION	5
1.1 Biology and stock structure.....	6
1.2 Fishery overview.....	6
2. OBSERVATIONS AND MODEL INPUTS.....	10
2.1 Spatial stratification	10
2.2 Temporal stratification.....	11
2.3 Definition of fisheries	11
2.4 Catch history	13
2.5 CPUE indices	14
2.5.1 Longline CPUE.....	14
2.5.2 Purse seine CPUE	17
2.5.3 Pole and line CPUE	18
2.5.4 Other abundance indices.....	18
2.6 Length frequency data.....	19
2.7 Tagging data.....	24
2.8 Environmental data	28
3. Model structural and assumptions	28
3.1 Population dynamics	28
3.1.1 Recruitment.....	28
3.1.2 Growth and Maturation.....	29
3.1.3 Natural mortality.....	31
3.1.4 Movement.....	32
3.2 Fishery dynamics	32
3.3 Dynamics of tagged fish.....	33
3.4 Modelling methods, parameters, and likelihood	34
4. ASSESSMENT model runs	35
4.1 The basic model	35
4.2 Exploratory model runs.....	36
4.3 Final model assessable (grid).....	36

5. model RESULTS	42
5.1 The basic model	42
5.1.1 Model fits.....	42
5.1.2 Model estimates	52
5.2 Exploratory analysis.....	59
5.3 Final model options.....	62
5.4 Diagnostics.....	66
5.4.1 Jitter analysis	66
5.4.2 ASPM analysis.....	66
5.4.3 Retrospective analysis.....	68
5.4.4 Hindcasting analysis	69
6. Stock status	71
6.1 Current status and yields	71
7. DISCUSSION	77
8. ACKNOWLEDGMENTS	78
9. REFERENCES	79
Appendix A: Revised model Fishery definition	84
Appendix B: RESULTS FROM THE EXPLORATORY MODELLING	85

SUMMARY

This report presents a preliminary stock assessment for Indian Ocean yellowfin tuna (*Thunnus albacares*) using *Stock Synthesis 3* (SS3). The assessment uses an age-structured and spatially-explicit population model and is fitted to catch rate indices, length-composition data, and tagging data. The assessment covers 1950 – 2020 and represents an update of the previous assessment model, taking into account progress and improvements made since the previous assessment. The assessment assumes that the Indian Ocean yellowfin tuna constitute a single spawning stock, modelled as spatially disaggregated four regions, with 21 fisheries. Standardised CPUE series from the main longline fleets 1975 – 2020 were included in the models as the relative abundance index of exploitable biomass in each region. The CPUE indices from EU Purse seine sets on free schools were included in a subset of models with the spatial and fleet structure revised to better accommodate the distribution and size structure of the purse seine fisheries. A new index based on associative dynamics of yellowfin tuna with floating objects and an additional index from the Maldivian pole and line fishery were also available, and the utility of these indices was examined in the assessment. Tag release and recovery data from the RTTP-IO program were included in the model to inform abundance, movement, and mortality rates.

A range of exploratory models is presented to address issues in observational datasets, improve the stability of the assessment model, and explore the effects of alternative model assumptions. The proposed final assessment model options correspond to a combination of model configurations, including alternative assumptions about the spatial structure (2 options), longline CPUE catchability (2 options on the effect of piracy), weighting of the tagging dataset ($\lambda = 0.1$ or 1), steepness values (0.7, 0.8, and 0.9), natural mortality values (2 options), and growth parameters (2 options). The model ensemble (a total of 96 models) encompasses a range of stock trajectories. Estimates of stock status were combined across from the 96 models and incorporated uncertainty estimates from individual models as well as across the model ensemble. Overall stock status estimates do not differ substantially from the previous assessment. Biomass is estimated to have been declining in recent years, and since the previous assessment. Spawning biomass in 2020 was estimated to be 78% of the level that supports the maximum sustainable yield ($SB_{2020}/SB_{MSY} = 0.78$). Current fishing mortality is estimated to be 27% higher than F_{MSY} ($F_{2020}/F_{40\%SB} = 1.27$). The probability of the stock being currently in the red Kobe quadrant is estimated to be 67%. The catches in the last five years have been higher than the estimated MSY. Considering the quantified uncertainty, the stock is considered to be overfished and is subject to overfishing in 2020. The estimated stock status is summarized as below:

• Catch in 2020:	432623
• Average catch 2016–2020:	434568
• MSY (1000 mt) (plausible range):	394 (325 – 463)
• F_{MSY} :	0.18 (0.14–0.21)
• SB_0 (1000 mt) (80% CI):	4192 (3228–5156)
• SB_{2020} (1000 mt) (80% CI):	1162 (773–1550)
• SB_{MSY} (80% CI)	1515 (1146 – 1885)
• SB_{2020}/SB_0 (80% CI):	0.28 (0.21–0.34)
• SB_{2020}/SB_{MSY} (80% CI):	0.78 (0.57–0.98)
• F_{2020}/F_{MSY} (80% CI):	1.27 (0.64–1.91)

1. INTRODUCTION

This paper presents a preliminary stock assessment of yellowfin tuna (*Thunnus albacares*) in the Indian Ocean (IO) including fishery data up to the end of 2020. The assessment implements an age- and spatially-structured population model using the Stock Synthesis software (Methot et al. 2020, Methot & Wetzel 2013).

Prior to 2008, Indian Ocean yellowfin tuna was assessed using methods such as VPA and production models (Nishida & Shono 2005 & 2007). In 2008, a preliminary stock assessment of IO yellowfin tuna was conducted using MULTIFAN-CL (Kleiber et al 2003, Langley et al. 2008) enabling the integration of the tag release/recovery data collected from the large-scale tagging programme conducted in the Indian Ocean in the preceding years (Langley et al. 2008). The MULTIFAN-CL assessment was revised and updated in the following years (Langley *et al.* 2009, 2010 and 2011, Langley et al. 2012a, 2012b).

In 2015, the assessment of IO yellowfin tuna was implemented using the Stock Synthesis software (SS3) (Langley 2015). The SS3 modelling framework is very similar to MFCL conceptually and the two platforms have yielded similar results. On basis of that assessment, the yellowfin tuna stock was determined to be overfished and subject to overfishing. At its 20th meeting, the Indian Ocean Tuna Commission adopted an Interim Plan for Rebuilding the Indian Ocean Yellowfin Tuna Stock (Res. 16/01).

The SS3 assessment was updated in 2016 (Langley 2016) and was revised and updated in 2018 (Fu et al. 2018a). Recent assessments have utilised new composite longline CPUE indices derived from the main distant water longline fleets, replacing the Japanese longline CPUE indices used previously. The 2018 assessment also included a comprehensive analysis of the main assumptions of the stock assessment. A model ensemble covering major components of structural uncertainty was used to characterise the stock status (IOTC 2018a). The assessment estimated that the spawning stock biomass in 2017 was below SB_{MSY} , and that fishing mortality was above F_{MSY} . Therefore, the stock status was determined to remain overfished and experiencing overfishing.

However, the model forecasts to evaluate fishery risk indicated a problem with the model structure or software since a substantial number of model projections yielded non-sensical results, with the stock crashing within a few years into the projection period even under low catch scenarios (Fu et al. 2018b). It was later discovered that the problem was mostly related to an assumption about regional recruitment distribution in the forecast (IOTC 2020). Further, it was considered that model uncertainty had not been adequately captured in the projections. Consequently, the SC considered that the assessment forecast was too uncertain to provide management advice to underpin the yellowfin tuna rebuilding plan (IOTC 2018b).

The Scientific Committee then initiated a work plan to address the problems identified in the 2018 assessment and carried out various work in following years to reduce the uncertainty of the assessment and to coordinate modelling decisions (Merino et al. 2019, 2020). An external review of the assessment provided recommendations to improve model parametrisations (Methot 2019). An attempt was made to update the assessment in 2019, with extensive investigations of alternative spatial structures, data weighting and biological parameters (Ijurco et al. 2019). Further analysis was conducted in 2020 to refine the process of model selection through an objective scoring system based on diagnostic metrics (Ijurco et al. 2020).

The IOTC Commission has thus tasked the Scientific Committee via its Working Party on Tropical Tunas, to conduct a new assessment of the status of the Yellowfin stock in 2021 using all available data. This report documents the next iteration of the stock assessment of the IO yellowfin tuna stock for consideration at 23rd WPTT meeting. This stock assessment is based on 2018 modelling framework of

IO yellowfin tuna but has incorporated some revisions made through additional analysis carried out in 2019 and 2020.

1.1 Biology and stock structure

Yellowfin tuna (*Thunnus albacares*) is a cosmopolitan species distributed mainly in the tropical and subtropical oceanic waters of the three major oceans, where it forms large schools. The sizes exploited in the Indian Ocean range from 30 cm to 180 cm fork length. Smaller fish (juveniles) form mixed schools with skipjack and juvenile bigeye tuna and are mainly limited to surface tropical waters, while larger fish are found in surface and sub-surface waters. Intermediate age yellowfin are seldom taken in the industrial fisheries, but are abundant in some artisanal fisheries, mainly in the Arabian Sea.

Spawning occurs mainly from December to March in the equatorial area (0–10°S), with the main spawning grounds west of 75°E. Secondary spawning grounds exist off Sri Lanka and the Mozambique Channel and in the eastern Indian Ocean off Australia (Froese & Pauly 2009). Yellowfin size at 50% maturity has been estimated at around 75 cm based on cortical alveolar stage (Zudaire et al 2013) and recruitment to the purse seine fishery occurs predominantly in July (as evident in the high catch rates of the Purse seine associated sets in region 1b in the third quarter). Newly recruited fish are primarily caught by the purse seine fishery on floating objects and the pole-and-line fishery in the Maldives. Males are predominant in the catches of larger fish at sizes larger than 150 cm (this is also the case in other oceans). Medium sized yellowfin concentrate for feeding in the Arabian Sea. Feeding behaviour is largely opportunistic, with a variety of prey species being consumed, including large concentrations of crustacean that have occurred recently in the tropical areas and small mesopelagic fishes which are abundant in the Arabian Sea.

Longline catch data indicates that yellowfin are distributed continuously throughout the entire tropical Indian Ocean, but some more detailed analysis of fisheries data suggests that the stock structure may be more complex. The tag recoveries of the RTTP-IO provide evidence of large movements of yellowfin tuna within the western equatorial region, although there are very few observations of large scale transverse movements of tagged yellowfin. This may indicate that the western and eastern regions of the Indian Ocean support relatively discrete sub-populations of yellowfin tuna. Studies of stock structure using DNA techniques have indicated that there may be genetically discrete subpopulations of yellowfin tuna in the north western Indian Ocean (Dammannagoda et al 2008) and within Indian waters (Kunal et al 2013). A recent study of stock structure using the gene sequencing technology along with a basin-scale sampling design indicated genetic differentiation between north and south of the equator within the Indian Ocean for yellowfin, and possibly additional genetic structure within the locations north of the equator (Grewe et al. 2020). These studies generally support the potential presence of population units of yellowfin tuna within the Indian Ocean, despite that there remains considerable uncertainty on sub-regional population structure in the India Ocean. The assessment assumes that the IO yellowfin tuna stock consists of several interconnected regional populations that have the same biological characteristics (see Figure 1). Isotope studies have suggested relatively limited movement, with resident behaviour at the temporal scale of their muscle turnover, which is 3 months (Ménard et al 2007).

1.2 Fishery overview

Yellowfin tuna, an important component of tuna fisheries throughout the Indian Ocean, are harvested with a diverse variety of gear types, from small-scale artisanal fisheries (in the Arabian Sea, Mozambique Channel and waters around Indonesia, Sri Lanka and the Maldives and Lakshadweep Islands) to large gillnetters (from Oman, Iran and Pakistan operating mostly but not exclusively in the

Arabian Sea) and distant-water longliners and purse seiners that operate widely in equatorial and tropical waters (Figure 1). Purse seiners and gillnetters catch a wide size range of yellowfin tuna, whereas the longline fishery takes mostly adult fish.

Prior to 1980, annual catches of yellowfin tuna remained below about 80,000 mt and were dominated by longline catches (Figure 2). Annual catches increased markedly during the 1980s and early 1990s, mainly due to the development of the purse-seine fishery as well as an expansion of the other established fisheries (fresh-tuna longline, gillnet, baitboat, handline and, to a lesser extent, troll). A peak in catches was recorded in 1993, with catches over 400,000 mt, the increase in catch almost fully attributable to longline fleets, particularly longliners flagged in Taiwan, which reported exceptional catches of yellowfin tuna in the Arabian Sea. The Taiwanese longline fishery in the Indian Ocean has been equipped with super-cold storage. Since around 1986, the fleet has fished more frequently with deep sets.

Catches declined in 1994, to about 350,000 mt, remaining at that level for the next decade then increasing sharply to reach a peak of about 520,000 mt in 2004/2005 driven by a large increase in catch by all fisheries, especially the purse-seine (free school) fishery. Total annual catches declined sharply from 2004 to 2007 and remained at about 300,000 mt during 2007–2011. In 2012, total catches increased to about 400,000 mt and were maintained at about that level through 2013 to 2015. Total catches increased to an average of 430,000 mt between 2016 and 2019, and a maximum of close to 450,000 mt in 2019 (Figure 2), despite IOTC Resolution 17/01 which requested major fleets to substantially reduce their yellowfin catches below the 2014 or 2015 catch level. Furthermore, catch levels of about 440,000 t reported for 2018 might be under-estimated (to some extent) because of changes in data processing methodology by EU, Spain for its purse seine fleet for that year (IOTC 2021a).

In recent years (2015–2020), purse seine has been the dominant fishing method harvesting 36% of the total IO yellowfin tuna catch (by weight), with the gillnet and handline fisheries, comprising 20% and 18% of the catch, respectively. There was a substantial increase in the catch by handline in 2020 (Figure 2). A smaller component of the catch was taken by industrial longline (5%), and the regionally important baitboat (4%) and troll (4%) fisheries. The recent increase in the total catch has been mostly attributable to an increase in catch from the gillnet and handline fisheries.

The purse-seine catch is generally distributed equally between free-school and associated (log and FAD sets) schools, although the large catches in 2003–2005 were dominated by fishing on free-schools. Conversely, during 2015–2019 the purse-seine catch was dominated (70%) by the associated fishery.

Historically, most of the yellowfin catch has been taken from the western equatorial region of the IO (44%; region 1b, see Figure 1) and, to a lesser extent, the Arabian Sea (26%), the eastern equatorial region (24%, region 4) and the Mozambique Channel (5%; region 2). The purse-seine and baitboat fisheries operate almost exclusively within the western equatorial region, while catches from the Arabian Sea are principally by handline, gillnet, and longline (see Figure 1). Catches from the eastern equatorial region (region 4) were dominated by longline and gillnet (around Sri Lanka and Indonesia). The southern Indian Ocean (region 3) accounts for a small proportion of the total yellowfin catch (1%) taken exclusively by longline (see Figure 1).

In recent years (2008–2012), due to the threat of piracy, the bulk of the industrial purse seine and longline fleets moved out of the western waters of Region 1b to avoid the coastal and off-shore waters off Somalia, Kenya and Tanzania. The threat of piracy particularly affected the freezer longline fleet and levels of effort and catch decreased markedly from 2007. The total catch by freezing longliners

declined to about 2,000 mt in 2010, a 10-fold decrease in catch from the years before the onset of piracy. Purse seine catches also dropped in 2007–2009 and then started to recover. Piracy off the Somali coast was almost eliminated by 2013 but longline catches have not recovered.

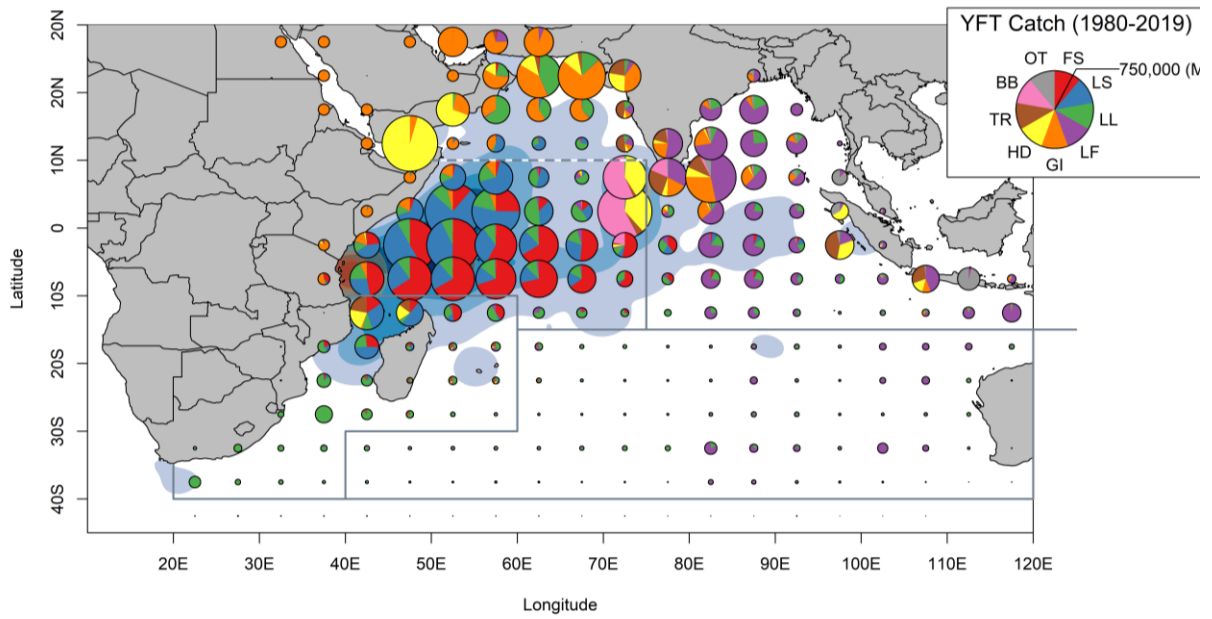


Figure 1: Spatial distribution of Indian Ocean yellowfin catches by main gear types aggregated for 1980-2020, overlaid with tag dispersion (see section 2.7). Gear codes are described in Table 1. Gray lines delineates the spatial structure used in the assessment model (see Section 2.1).

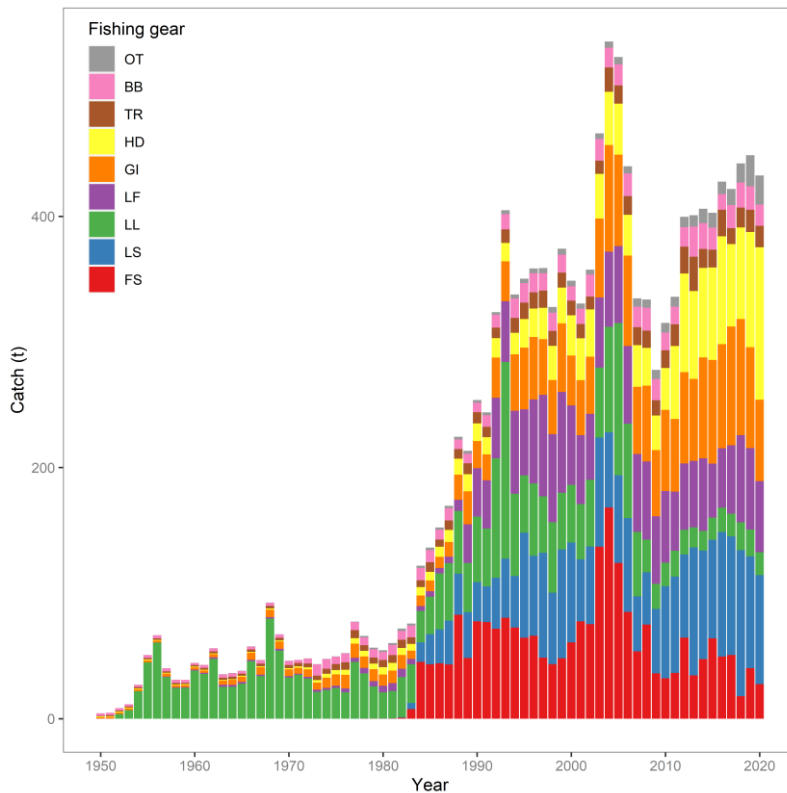


Figure 2: Total annual catch (1000s mt) of yellowfin tuna by main gear types from 1950 to 2020 (Gear codes are described in Table 1).

2. OBSERVATIONS AND MODEL INPUTS

The data used in the yellowfin tuna assessment consist of catch and length composition data for the fisheries defined in the analysis, CPUE indices and tag release-recapture data. The details of the configuration of the fishery specific data sets are described below.

2.1 Spatial stratification

The geographic area considered in the assessment is the Indian Ocean, defined by the coordinates 40°S–25°N, 20°E–150°E. Earlier yellowfin stock assessments have adopted a five region spatial structure (see Langley 2012). Preliminary analyses conducted during the 2015 assessment highlighted a number of issues related to the five region model structure (see Langley 2015). There have been no CPUE abundance indices available from the Arabian Sea region (region 1a) since 2010 although the area has yielded very high catches from the handline and gillnet fisheries during recent years. Assessments since 2015 thus adopted the four region model structure, combining the Arabian Sea (region 1a) and western equatorial region (region 1b) (Figure 3), although the two sub regions were retained for the definition of spatially distinct fisheries that operate in each area. The spatial structure retains two regions that encompass the main year-round fisheries in the tropical area and two austral, subtropical regions where the longline fisheries occur more seasonally.

The current spatial structure separates the purse-seine fishery in the northern Mozambique Channel (10–15°S) from the equatorial region, as the fishery in the northern Mozambique Channel exhibits strong seasonal variation in effort and operates differently from the equatorial region (Langley 2015). There is also a separation of the purse-seine fishery between the western and eastern tropical region with the current boundary between region 1b and region 4. In the assessment we also considered a revised regional structure where the southern boundary of region 1b is shifted to 20S, and eastern boundary of 1b is shifted slightly to 80E (Figure 3). The purse-seine fishery is completely located within the revised region 1, resulting in a simplification of the fleet structure. The revision also helps improve the modelling of the length composition data from the purse seine fishery, promote the inclusion of the new purse-seine CPUE index (see Section 5.2) and improve the stability of the model by constraining all tag recoveries to its main dispersion area (Figure 3).

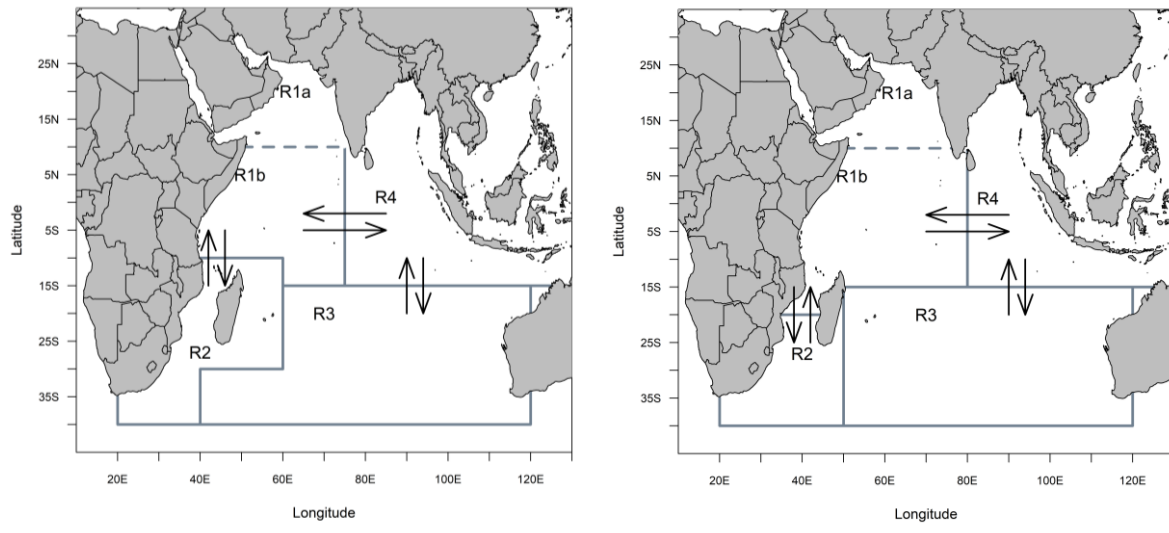


Figure 3: Four region spatial stratification of the Indian Ocean for the basic assessment model (left), and a revised spatial stratification (right). The black arrows represent the configuration of the movement parameterisation of the assessment model

2.2 Temporal stratification

The time period covered by the assessment is 1950–2020 representing the period for which catch data are available from the commercial fishing fleets. Langley (2015) suggested that the assessment results were not sensitive to the early catches from the model (pre-1972) and commencing the model in 1950 or 1972 (assuming unexploited equilibrium conditions) yielded very similar results.

Within this model period, the annual data were compiled into quarters (Jan–Mar, Apr–Jun, Jul–Sep, Oct–Dec) (representing a total of 272 time steps). The time steps were used to define model “years” (of 3 month duration) enabling recruitment to be estimated for each quarter to approximate the continuous recruitment of yellowfin in the equatorial regions.

The quarterly time step (model “year”) precluded the estimation of seasonal model parameters, particularly the movement parameters. Fu et al. (2018a) explored an alternative annual/seasonal model structure which explicitly estimated seasonal movement dynamics. However, the alternative temporal structure did not yield substantially different results.

2.3 Definition of fisheries

The assessment adopted the equivalent fisheries definitions used in the previous SS3 stock assessment. These “fisheries” represent relatively homogeneous fishing units, with similar selectivity and catchability characteristics that do not vary greatly over time. Twenty-one fisheries were defined based on location (region), time period, fishing gear, purse seine set type, and type of vessel in the case of longline fleet (Table 1).

The longline fishery was partitioned into two main components:

Freezing longline fisheries, or all those using drifting longlines for which one or more of the following three conditions apply: (i) the vessel hull is made up of steel; (ii) vessel length overall of 30 m or greater; (iii) the majority of the catches of target species are preserved frozen or deep-frozen. A composite

longline fishery was defined in each region (LL 1–4) aggregating the longline catch from all freezing longline fleets (principally Japan and Taiwan).

Fresh-tuna longline fisheries, or all those using drifting longlines and made of vessels (i) having fibreglass, FRP, or wooden hull; (ii) having length overall less than 30 m; (iii) preserving the catches of target species fresh or in refrigerated seawater. A composite longline fishery was defined aggregating the longline catch from all fresh-tuna longline fleets (principally Indonesia and Taiwan) in region 4 (LF 4), which is where the majority of the fresh-tuna longliners have traditionally operated. The catches of yellowfin tuna recorded in regions 1 to 3 for fresh-tuna longliners, representing only 3% of the total catches over the time series, were assigned to area 4.

The purse-seine catch and effort data were apportioned into two separate method fisheries: catches from sets on associated schools of tuna (log and drifting FAD sets; PS LS) and from sets on unassociated schools (free schools; PS FS). Purse-seine fisheries operate within regions 1a, 1b, 2 and 4 and separate purse-seine fisheries were defined in regions 1b, 2 and 4, with the limited catch, effort and length frequency data from region 1a reassigned to region 1b.

In the previous assessment, the region 1b purse-seine fisheries (log and free-school) were divided into three time periods: pre-2003, 2003–2006 and post-2006. This was mainly to maintain historical consistency (the temporal stratification was initially implemented to account for change in the length composition during the 2000s, but no selectivity changes were identified in the assessment and consequently the same selectivity was shared among the three time periods. For the current assessment, the temporal stratification was removed, reducing the 6 purse-seine fisheries in region 1b to 2 fisheries (log and free-school).

A single baitboat fishery was defined within region 1b (essentially the Maldives fishery). As with the purse-seine fishery, a small proportion of the total baitboat catch and effort occurs on the periphery of region 1b, within regions 1a and 4. The additional catch was assigned to the region 1b fishery.

Gillnet fisheries were defined in the Arabian Sea (region 1a), including catches by Iran, Pakistan, and Oman, and in region 4 (Sri Lanka and Indonesia). A very small proportion of the total gillnet catch and effort occurs in region 1b, with catches and effort reassigned to area 1a.

Three troll fisheries were defined, representing separate fisheries in regions 1b (Maldives), 2 (Comoros and Madagascar) and 4 (Sri Lanka and Indonesia). Moderate troll catches are also taken in regions 1a and 3, the catch and effort from this component of the fishery reassigned to the fisheries within region 1b and 4, respectively.

A handline fishery was defined within region 1a, principally representing catches by the Yemenis fleet. Moderate handline catches are also taken in regions 1b, 2 and 4, the catch and effort from these components of the fishery were reassigned to the fishery within region 1a.

For regions 1a and 4, a miscellaneous (“Other”) fishery was defined comprising catches from artisanal fisheries other than those specified above (e.g. trawlers, small purse seines or seine nets, sport fishing and a range of small gears).

For models that are based on the revised regional structure, the fisheries definition was revised accordingly (Table 1, Appendix A). For most fisheries, this simply means redistributing the catch according to the new region boundaries. The regional purse seine fisheries in the original region 1, 2, 4 are now merged into a single fishery (log and free schools separately) located in the new region 1. For both log and free schools, the purse fisheries are now divided into a small fish (≤ 80 cm) and large fish (≥ 80 cm) components. The division allows the model to better account for the variable proportion

of the small and large fish in the length composition of the purse seine fishery. The separation is also necessary given that purse seine CPUE indices have been standardised for the juvenile and adult parts of the catch independently. Another minor change to the fisheries definition is to merge the troll fishery with the “Other” fishery group given that there is limited size data from the troll fishery, and it is suggested that some CPCs might be reporting their catch under the “trolling” gear codes as an umbrella term encompassing several miscellaneous small-scale gears (GTA 2011).

Table 1: Definition of fisheries for yellowfin tuna assessment model for the basic four-region spatial structure.

Fishery	Gear	Region
1. GI 1a	Gillnet	1a
2. HD 1a	Handline	1a
3. LL 1a	Longline (distant water)	1a
4. OT 1a	Other	1a
5. BB 1b	Baitboat	1b
6. PS FS 1b	Purse seine, school sets	1b
7. LL 1b	Longline (distant water)	1b
8. PS LS 1b	Purse seine, log/FAD sets	1b
9. TR 1b	Troll	1b
10. LL 2	Longline (distant water)	2
11. LL 3	Longline (distant water)	3
12. GI 4	Gillnet	4
13. LL 4	Longline (distant water)	4
14. OT 4	Other	4
15. TR 4	Troll	4
16. PS FS 2	Purse seine, school sets	2
17. PS LS 2	Purse seine, log/FAD sets	2
18. TR 2	Troll	2
19. PS FS 4	Purse seine, school sets	4
20. PS LS 4	Purse seine, log/FAD sets	4
21. LF 4	Longline (fresh tuna)	4

2.4 Catch history

Catch data were compiled based on the fisheries definitions (only catches for the original fisheries definition are shown in Figure 4). An update of quarterly catches by fishery was provided by the IOTC Secretariat, including catches from 2018 to 2020 (as at 30/9/2021). For most fisheries, the time series of catches were very similar to the catch series included in the 2018 assessment (Figure 4). The current estimates for “OT 1a” and “TR 4” are lower than the catches included in the previous assessment,

whereas catch estimates for “GI 1a” and “TR 1b” fisheries have been revised upward. The changes are mostly attributed to revisions of catch estimation by the IOTC Secretariat.

For the revised fishery definitions, the quarterly catch for purse seine fisheries (log and free schools) is divided into the small (≤ 80 cm) and large (>80 cm) fish catch components based on proportion by weight as derived from fishery-level size distribution for each quarter and length-weight relationship.

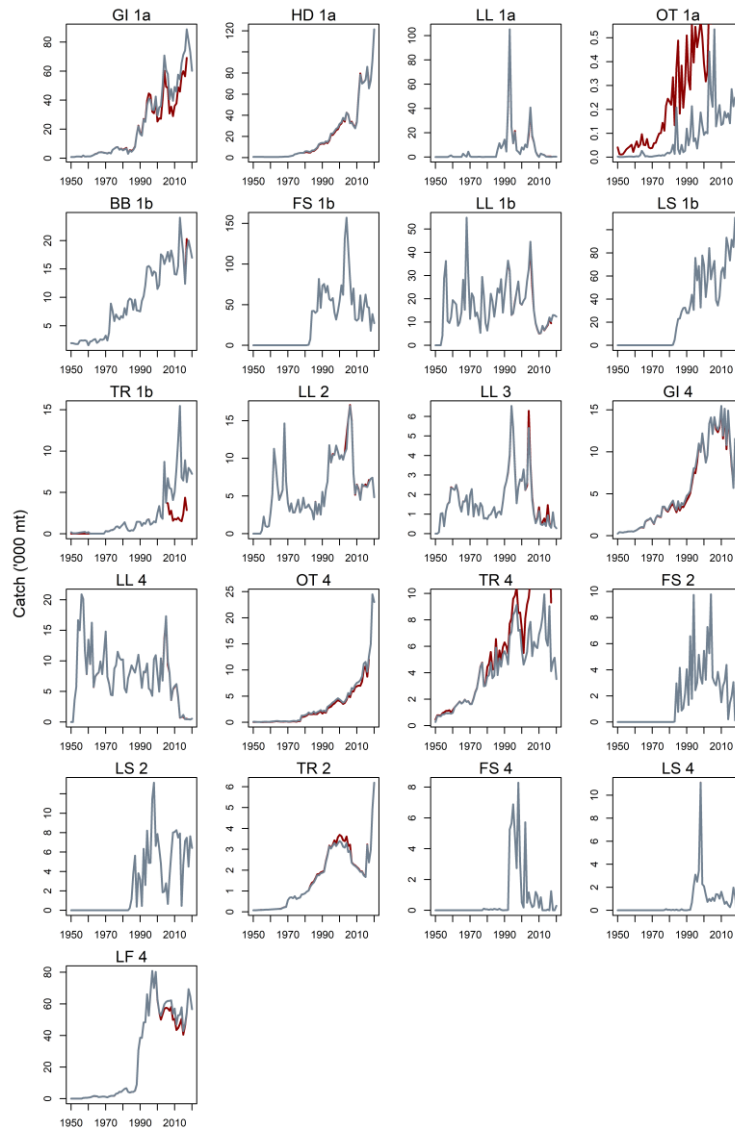


Figure 4: Fishery catches (in 1000 metric tonnes) aggregated by year. Note the y-axis differs among plots. Red lines are catches used in the 2018 assessment.

2.5 CPUE indices

2.5.1 Longline CPUE

Standardised CPUE indices were derived using generalized linear models (GLM) from longline catch and effort data (aggregated by month and 1° grid resolution) provided by Japan, Korea, Taiwan, China (Kitakado et al. 2021). Cluster analyses of species composition data for each fleet were used to separate datasets into fisheries understood to target different species. Selected clusters were then combined and

standardized using generalized linear models. Yellowfin catch (numbers of fish) was the dependent variable of the positive catch model (lognormal error structure), while the presence/absence of yellowfin tuna in the catch was the dependent variable in the binomial model. In addition to the year-quarter, models included covariates for 5° square location, number of hooks, and vessels (that accounted for 50% of the total effort). The CPUE for the temperate regions 2 and 3 incorporated the cluster variable to indicate the targeting effect, whereas the tropical regions 1 and 4 used hooks between floats (HBF). The CPUE indices represented the time series of abundance (1975–2020) for each of the four model regions (1, 2, 3, and 4). The data from region 1a is not included in the standardisations and the index for region 1b is assumed to index the abundance for the whole of region 1.

The standardised quarterly CPUE indices included in the assessment are shown in Figure 5. In general, the overall trends in the updated CPUE indices are similar to those included in the previous assessment, but the updated indices in regions 1, 3, and 4 have larger declines, especially in the early years. For region 3, the updated analysis is based on an aggregated dataset in which there were fewer vessels targeting yellowfin tuna and much less catch and effort, so the information is more scarce (Kitakado et al. 2021).

For the regional longline fisheries, a common catchability coefficient (and selectivity) was estimated in the assessment model, thereby, linking the respective CPUE indices among regions. This significantly increases the power of the model to estimate the relative (and absolute) level of biomass among regions. However, as CPUE indices are essentially density estimates it is necessary to scale the CPUE indices to account for the relative abundance of the stock among regions. For example, a relatively small region with a very high average catch rate may have a lower level of total biomass than a large region with a moderate level of CPUE.

The approach used was to determine regional scaling factors that incorporated both the size of the region and the relative catch rate to estimate the relative level of exploitable longline biomass among regions. This approach is similar to that used in the WCPO regionally disaggregated tuna assessments. Hoyle & Langley (2018) proposed a set of regional weighing factors for IO yellowfin based on aggregated longline catch effort data. The authors recommended the estimates by method ‘8’ for the period 1979–1994 (referred to as ‘7994m8’, see Table 2 of Hoyle & Langley (2018)) to be included in the current assessment. The relative scaling factors calculated for regions 1–4 are 1.674, 0.623, 0.455 and 1.000 respectively.

For each of the principal longline fisheries, the GLM standardised CPUE index was normalised to the mean of the period for which the region scaling factors were derived (i.e., the GLM index from 1979–1994). The normalized GLM index was then scaled by the respective regional scaling factor to account for the regional differences in the relative level of exploitable longline biomass among regions

A number of important trends are evident in the CPUE indices.

- 1980-1989 the western tropical (region 1b) CPUE increased during the 1980s, then declined until 1995, increased again until 2005, and then decreased again. The low CPUE indices followed the period of exceptionally high catches from the purse seine fishery in region 1b during 2003–2005. The drop in CPUE occurred before the peak in the number of piracy incidents in the western Indian Ocean (2008–2011). After that time, it remained close to the lowest level observed but showed very large seasonal variations.
- The eastern tropical region 4 followed a similar pattern until 1990 but then declined steadily, and by 2016 was also close to the lowest level in the time series. The recent decline in CPUE in this region is consistent with a decline in the proportion of yellowfin in the combined tuna catch from the Japanese longline fleet in the eastern Indian Ocean (see Figure 44 from Hoyle

et al 2015). It is unclear whether the change in species proportion is related to a decline in the abundance of yellowfin in the region (relative to the other species) or a regional change in the targeting of the fishing fleet. However, there is an indication that there has been a differential shift towards deeper longline gear (greater HBF) in the eastern Indian Ocean since 2000 and this may indicate a shift in targeting toward bigeye tuna in this region (Hoyle pers. comm. additional JP LL analyses). Such factors may not be adequately accounted for in the standardisation of the yellowfin CPUE data.

- The CPUE indices in western temperate region 2 followed a similar pattern to the western tropical indices, with a decline until the late-1970s followed by an increase until the late 1980s, and subsequently a slow decline with significant variability. However, the two sets of CPUE indices diverge somewhat from about 2007 with the CPUE indices from R2 being maintained at a higher level relative to R1.
- The CPUE indices from region 3 are low compared to the other three regions reflecting the low regional scaling factor. However, the overall trend in the CPUE indices is broadly comparable to the other regions. The eastern temperate region 3 the pattern was similar to the western temperate area before 1979. After 1979 catch rates increased until the mid-2000's, but then declined rapidly and reached their lowest observed levels by 2016.
- There is an exceptionally high peak in CPUE indices 1976–78. Hoyle et al. (2017) showed that this discontinuity exists in Japanese, Taiwanese and Korean data, and in multiple regions in multiple oceans, and for both bigeye and yellowfin tuna. Hoyle et al (2017) suggested this is unlikely to be explained by changes to the population or catchability but may be associated with catch reporting and data management.
- The spike in the CPUE indices around 2012 in the west equatorial region (region 1) was evident for most fishing fleets. Several hypotheses have been proposed on what could have caused CPUE to have increased, including a return to fishing in areas that were most affected by piracy. However, further investigation is required.

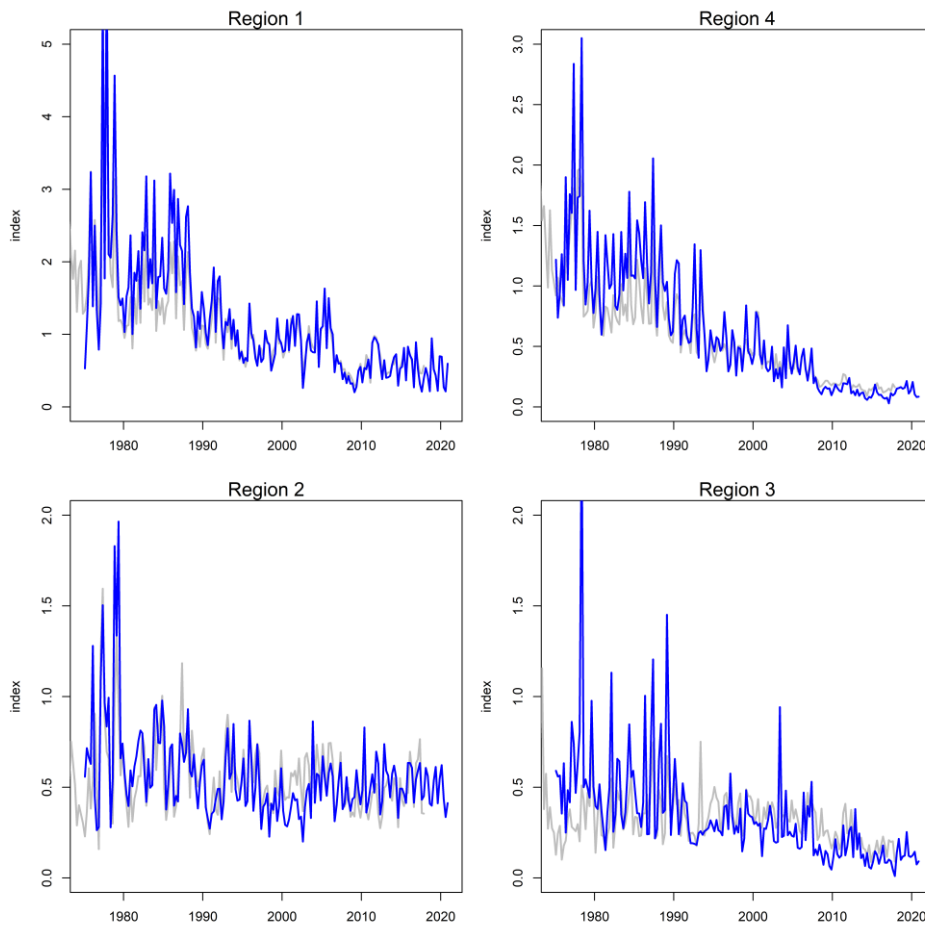


Figure 5: A comparison of the longline CPUE indices included in the 2018 stock assessment (grey line) and the 2021 stock assessment (blue line). The 2018 indices are rescaled to have the same mean as the 2021 indices for each region.

2.5.2 Purse seine CPUE

The European and associated flags purse seine fishing activities in the Indian Ocean during 1981–2020 have been monitored through the collection of logbooks and observer sampling. Standardised indices of the biomass of yellowfin caught by European purse seiners (Spain and France) from sets on free swimming schools (1991 – 2020) as well as sets on associated tuna schools (2010 – 2017) were developed by Guery et al (2021) (Figure 6). The standardisation was based on the application of a generalized linear mixed model which considered a comprehensive list of candidate covariates, including the effect of the technological improvement related to the use of echosounder buoys. A number of standardised indices were developed, based on:

- Randomly encountered free school sets with adult (> 10kg) 1991 – 2020
- Randomly encountered free school sets with juvenile (< 10kg) 1991– 2020
- Randomly encountered log school sets with juvenile (< 10kg) 2010– 2020
- Non-randomly encountered log school sets with juvenile (< 10kg) 2010– 2020

As those indices are based on either the juvenile or adult part of the population vulnerable to the purse seine fishery, they are incorporated into the assessment model that is based on the revised region/fishery structure in which purse seine fisheries are separated into the small and large fish components.

It is well recognised that the relationship between PS CPUE and abundance is unlikely be proportional, as the improvement of catch efficiency due to technology development is difficult to quantify, and the changes in catchability are not fully accounted for in the standardisation process. The WCPO assessments have often estimated substantial changes in PS FAD-associated fisheries (e.g., McKechnie et al 2017). Using a similar approach, Kolody (2018) estimated a catchability increase of approximately 1.25% per year for the standardised purse seine effort for yellowfin from sets on associated schools.

2.5.3 Pole and line CPUE

Medley et al. (2021) provided an abundance index for yellowfin tuna from Maldives pole and line catch and effort data. The index was derived from multiple datasets with differing levels of detail over the period 1970–2020. The standardisation undertaken using a Bayesian approach, accounted for missing information about the mechanization of the fleet and included additional fishing power effects estimated using subjective priors based on a meeting with fishery experts.

The main concern regarding this CPUE index is that the relatively small spatial area in which the Maldives pole and line operates may not be representative of abundance trends in other areas of the Indian Ocean, and thus the index may be more appropriate as a regional abundance index. The substantial decline in the abundance index in the early period is thought to be a result of the overestimation of changes in fishing power. The WPTT23 Data Preparatory meeting suggested that the time series covering 1990-2018 are probably more reliable and could be considered as an abundance estimate for juvenile yellowfin tunas (IOTC–WPTT23(DP) 2021).

2.5.4 Other abundance indices

Baidai et al. (2021) developed a novel approach to construct estimates of tropical tuna population size based on their associative dynamics with floating objects and acoustic data collected from echosounder buoys. The approach was implemented to provide time series of abundance for yellowfin tuna in the western Indian Ocean, over the period 2013 to 2019.

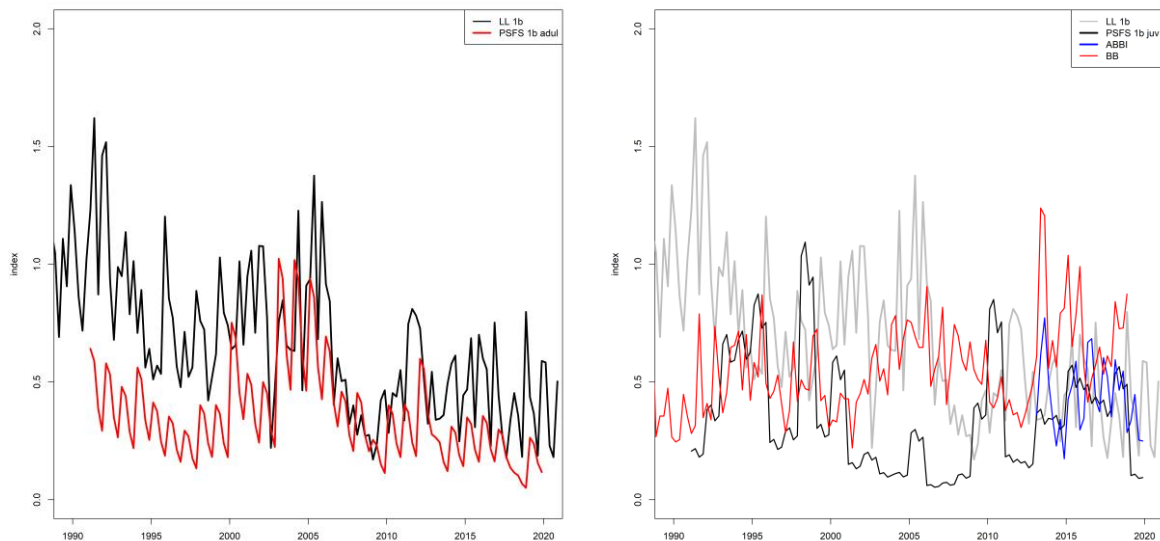


Figure 6: Other CPUE indices for yellowfin tuna: Index from the purse seine free schools on adults (left), Index from the Purse seine free schools on juvenile (PSFS 1b juv), associated behaviour abundance index (ABBI), and pole and line index (BB) (right). Longline CPUE indices are included for comparison (LL 1b).

2.6 Length frequency data

Available length-frequency data for each of the defined fisheries were compiled into 95 2-cm size classes (10–12 cm to 198–200 cm). Each length frequency observation consists of the actual number of yellowfin tuna measured. A graphical representation of the availability of length samples is provided in Figure 7. No the length samples are available for TR 2 and TR 1b. The data were collected from a variety of sampling programmes, which can be summarized as follows:

Purse seine: Length-frequency samples from purse seiners have been collected from a variety of port sampling programmes since the mid-1980s. The samples are comprised of very large numbers of individual fish measurements. The length frequency samples are available by set type with sets catches from associated sets typically composed of smaller fish than free school catches (Figure 8). The size composition of the catch from the free-school fishery is bimodal, being comprised of the smaller size range of yellowfin and a broad mode of larger fish (Figure 8). The bimodal distribution is likely to have reflected different types of schools in the catch composition (e.g., free schools of mostly large adult yellowfin, or mixed species schools consisting of smaller yellowfin, M. Chassot, pers. comm.). Hence the relative composition of large (≥ 80 cm) vs. small (< 80 cm) yellowfin in the purse seine free schools fluctuates considerably over time. Between 2010 and 2020 there was both a dip in the average size of large fish caught in the FAD fishery, and a temporary increase in the average sizes of large fish caught in the free school fishery (Figure 9). There is also considerable catch of smaller fish taken during free school fishing operation in the Mozambique Channel area in region 2 (Chassot 2014). The free-school fishery in region 4 appears to catch larger fish.

Longline freezing: Length and weight data have been collected from sampling at ports and aboard Japanese commercial, research vessels, and observer programmes. Weight frequency data collected from the fleet have been converted to length frequency data via a processed weight-whole weight conversion factor and a weight-length key. Length frequency data from the Taiwanese longline fleet from 1980–2003 are also included in the previous assessment, although data from the more recent years were excluded due to concerns regarding the reliability of these data (Geehan & Hoyle 2013). length data are also available from other fleets (e.g., Seychelles, Korean, China, etc/) in more recent years. However, a recent review of the longline size data shows that the sampling behaviour of Taiwanese and Seychelles fleets (mostly reflagged Taiwanese vessels) have changed over time, with patterns in the logbook length data inconsistent with other fleet (Hoyle et al. 2021), and as such the WPTT23 (DP) recommended omitting all Taiwanese and Seychelles logbook length data from the current assessment (IOTC–2021).

Analyses of size data show that the average lengths of yellowfin caught by the longline fleet are generally larger in the southern regions, particularly in the southwest (Hoyle et al. 2021). There is considerable temporal variation in the length of fish caught, but some of this variation is inconsistent between datasets, such as temporal patterns of variation in the 1970s that differ between length and weight data from the Japanese fleet (Figure 9). For all longline fisheries there was a marked decline in the size of fish caught by Japan during the 1950s and 1960s, while the size of fish caught stabilised during the 1970s and 1980s. Attempts to fit these data in past assessments suggested that the large decline before 1965 is inconsistent with model assumptions about population dynamics. Hoyle et al. (2021) suggest that selectivity may have changed during this early period and recommend avoiding fitting to these data with the same selectivity, which may be achieved by omitting these pre-1965 data from the assessment.

Based on these recommendations, the length frequency data from the Taiwanese and Seychelles longline logbooks were excluded from the final length frequency data sets. The pre-1965 length data from the Japanese fleet were also excluded, encompassing the period of considerable decline in the length composition in all regions. The length data collected by scientific observers (including Taiwanese observer data) in the period 2005–2020 were included in the assessment.

Longline fresh: Length and weight data were collected in port, during unloading of catches, for several landing locations and time periods, especially on fresh-tuna longline vessels flagged in Indonesia and Taiwan/China (IOTC-OFCE sampling). However, the quality of these data is highly variable. Length data from 1998–2008 were included in the previous assessment, but most samples were subsequently found to be biased (F. Fiorellato personal communication). For the current assessment, only 2013–2020 data are included.

Gillnet: Length data are available from both GN 1 and 4 fisheries. The sizes of yellowfin taken by the gillnet fleet range from 40 to 140 cm.

Baitboat: Size data are available from the fishery from 1983 to 2020.

Troll: No size data are available from the TR 2 and TR 1b fisheries. The troll fishery in region 4 was sampled during two periods: 1985–1990 (Indonesian fishery) and 2015–2020 (Sri Lankan fishery). The samples from 2015–2019 appear to be problematic and are excluded from the current assessment.

Handline: Limited sampling of the handline fishery was conducted over the last decade. Samples are available for the Maldivian handline fisheries for this period. Given the considerable catches taken from the fishery, the significant and problematic data gap may pose a problem.

Other: Length samples are available from the “Other” fishery in region 4 (OT 4) fishery and limited data are available from the “Other” fishery in region 1a (OT 1a) (2009–2017).

Length data from each fishery/quarter were simply aggregated assuming that the collection of samples was broadly representative of the operation of the fishery in each quarter.

For the revised fisheries definition (see Table 1, Appendix A), the length compositions the purse seine fishery (both log and free schools) were divided into small (≤ 80 cm) and large (> 80 cm) fish components. The length composition from the Troll fishery were assimilated into the ‘Other’ fishery category to form the length composition for the new composite “Other” fishery group.

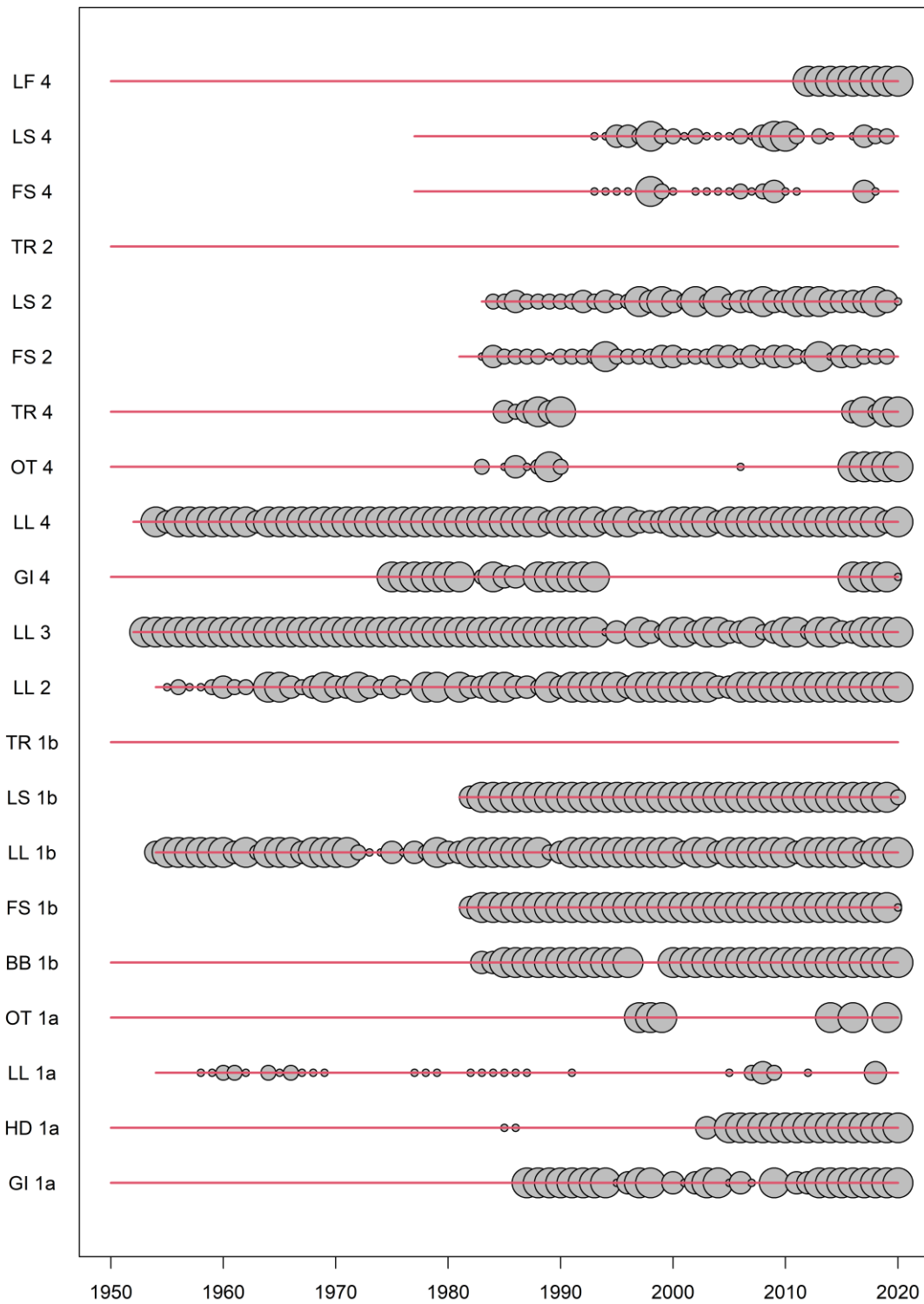


Figure 7: The availability of length sampling data from each fishery by year. The grey circles denote the presence of samples in a specific year. The red horizontal lines indicate the time period over which each fishery operated. The relative size of the bubble indicates number of quarters (in each year) being sampled.

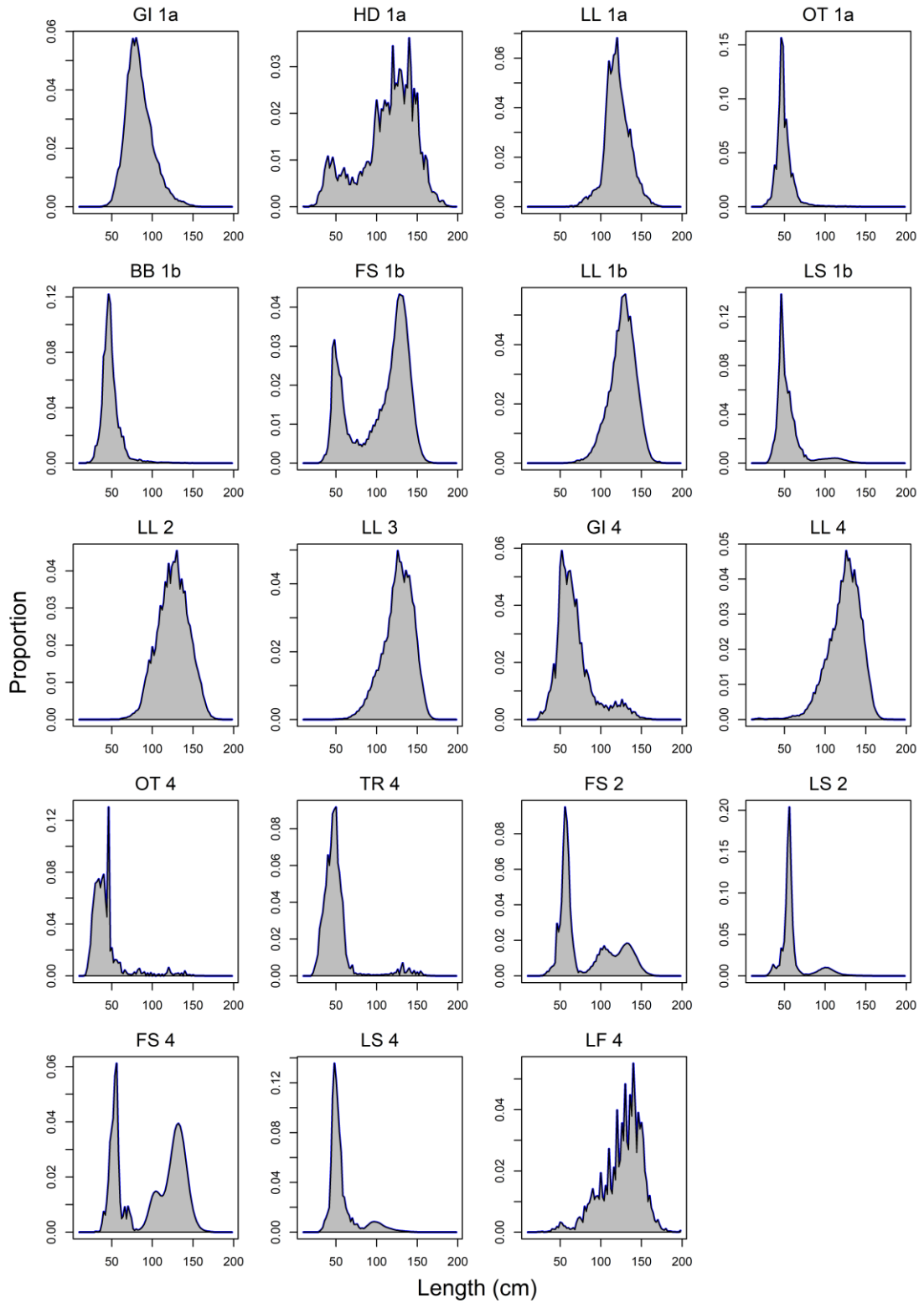


Figure 8: Length compositions of bigeye tuna samples aggregated by fishery.

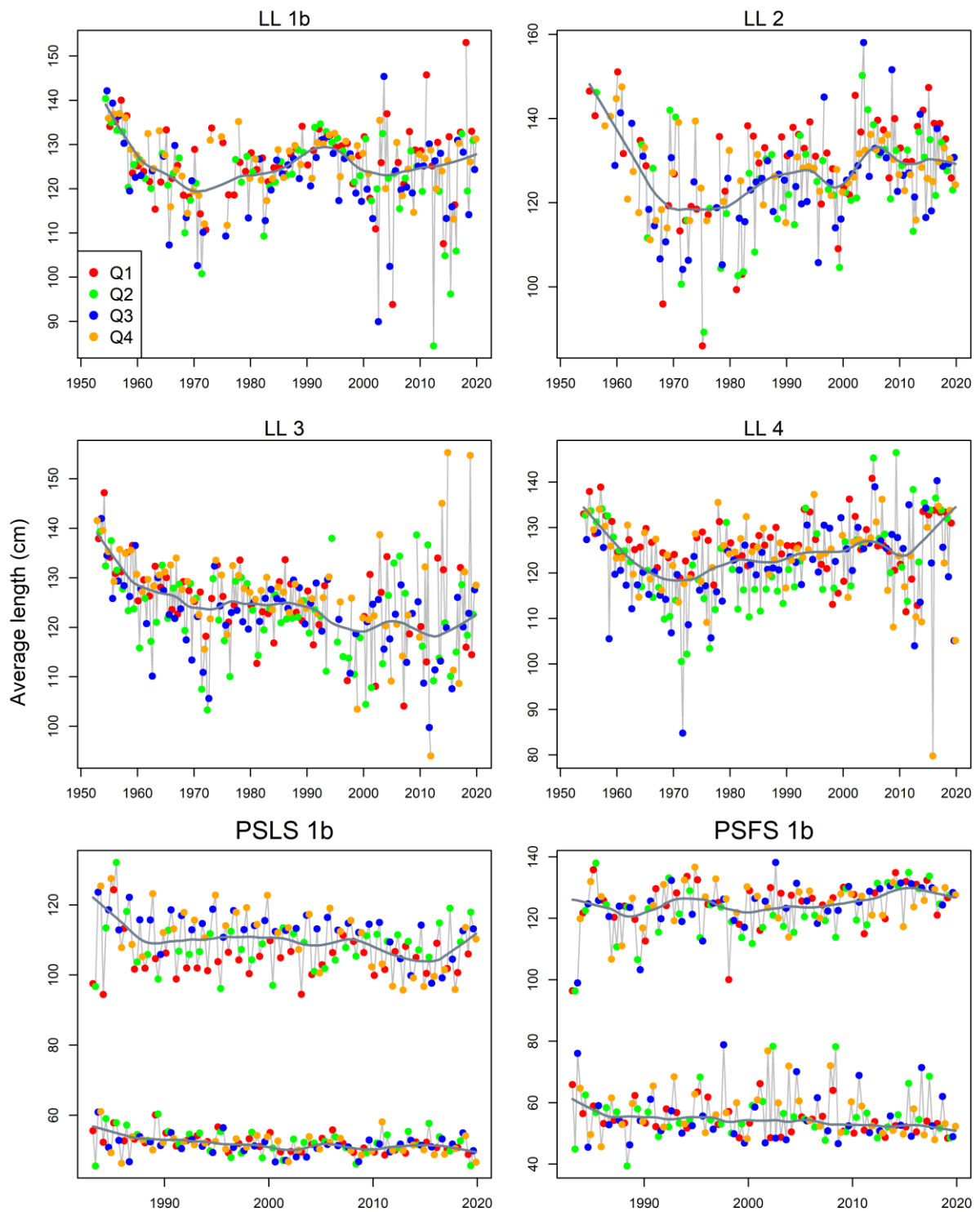


Figure 9: Mean length (fork length, cm) of yellowfin sampled from the principal fisheries (LL 1a-4, PSLS 1b and PSFS 1b) by year quarter. The grey line represents the fit of a lowess smoother to each data set. For the purse seine fisheries, the mean is calculated for fish ≤ 80 cm and > 80 cm separately.

2.7 Tagging data

A considerable amount of tagging data was available for inclusion in the assessment model. The data used consisted of yellowfin tuna tag releases and returns from the Indian Ocean Tuna Tagging Programme (IOTTP), and mainly from its main phase, the Regional Tuna Tagging Project-Indian Ocean (RTTP-IO) conducted during 2005–2009. The IOTC has compiled all the release and recovery data from the RTTP-IO and the complementary small-scale programmes in a single database.

A total of 54,688 yellowfin tuna were released by the RTTP-IO program. Most of the tag releases occurred within the western equatorial region (region 1b) and a high proportion of these releases occurred in the second and third quarters of 2006 (see IOTC 2008a for further details) (Figure 10). Limited tagging also occurred within regions 1a and 2. The model included all tag recoveries up to the end of 2014 and there were no further recoveries since the last assessment. The spatial distributions of tag releases and recoveries are presented in Figure 11 .

In total, 9916 tag recoveries (after removing tags with unknown recovery date or length) could be assigned to the fisheries included in the model. Almost all of the tags released in region 1 were recovered in the home region, although some recoveries occurred in adjacent regions, particularly region 2. A small number of tags were recovered in region 4 (from tags released in region 1b) and there were no tags recovered from region 3 (Table 2). Most of the tag recoveries occurred between mid-2006 and mid 2008 (Figure 11). The number of tag recoveries started to attenuate in 2009 although small numbers of tags were recovered up to the end of 2014.

Most of the tags were recovered by the purse seine fishery within region 1b (Figure 11). A significant proportion (35%) of the tag returns from purse seiners were not accompanied by information concerning the set type. These tag recoveries were assigned to either the free-school or log fishery based on the expected size of fish at the time of recapture; i.e. fish larger than 80 cm at release were assumed to be recaptured by the free-school fishery; fish smaller than 80 cm at release and recaptured within 18 months at liberty were assumed to be recovered by the floating object fishery; fish smaller than 80 cm at release and recaptured after 18 months at liberty were assumed to be recovered by the free-school fishery.

For incorporation into the assessment model, tag releases were stratified by release region, time period of release (quarter) and age class. The recaptures by fishery for each release group inform the assessment model on fishing mortality and abundance and fish movement. Therefore, factors that might have affected the interpretation of tag returns need to be accounted for to minimise potential bias. Fu (2020) provides a summary of how the tag data were incorporated into the assessments of IOTC tropical tuna species, and below is a description of the procedure applied to yellowfin tuna.

Age assignment of tag release

The age at release was assumed based on the fish length at release and the average length-at-age from the yellowfin growth function (see Section 3.1.2). Fish aged 15 quarters and older were aggregated in a single age group. Tag releases in regions 1a and 1b were stratified in separate release groups due to the spatial separation of the individual release events. A total of 54,392 releases were classified into 131 tag release groups. Most of the tag releases were in the 5–8 quarter age classes (Figure 11).

Initial tagging mortality

Hoyle et al. (2015) examined the effects of various covariates (e.g., individual tagger effect) on tag failures for the RTTP program and estimated a combined effect of 20% for all tropical tuna species relative to a base failure rate. No formal estimate was made for the base failure rate but a 7.5% was

suggested by the WPTT in 2018 based on the assessment of the Western and Central Pacific tuna species. This equates to a total tag failure rate of 27.5%. For the current assessment, the number of tags in each release group was reduced by 27.5% to account for initial tag mortality

Chronic tag loss

Tag recoveries were also corrected for long-term tag loss (tag shedding) based on an update of the analysis of Gaertner and Hallier (2015). Tag loss for yellowfin was estimated to be approximately 20% at 2000 days at liberty. This was accounted for through the SS3 chronic tag loss parameter (an annual rate of 0.03).

Reporting rate

The returns from tag release group were classified by recapture fishery and recapture time period (quarter). The results of associated tag seeding experiments conducted during 2005–2008, have revealed considerable temporal variability in tag reporting rates from the IO purse-seine fishery (Hillary *et al.* 2008a). Reporting rates were lower in 2005 (57%) compared to 2006 and 2007 (89% and 94%). Quarter estimates were also available but was similar in magnitude (Hillary *et al.* 2008b). This large increase over time was the result of the development of publicity campaign and tag recovery scheme raising the awareness of the stakeholders, *i.e.* stevedores and crew. SS3 assumes a constant fishery-specific reporting rate. To account for the temporal change in reporting rate, the number of tag returns from the purse-seine fishery in each stratum (tag group, year/quarter, and length class) were corrected using the respective estimate of the reporting rate. Following Kolody (2011) and Fu (2017, 2020), an 100% reporting rate was assumed for at-sea recoveries whereas tags recovered from Seychelles landings were corrected for reporting rates based on the quarterly estimates from Hillary *et al.* (2008b), and were also corrected for the portion of the total purse-seine catches examined for tags, based the proportions of EU PS catch landed in the Seychelles relative to the total EU PS catches (Kolody 2011). For example, the adjusted number of observed recaptures for a PSLS fishery as input to the model, R'_L was calculated using the following equation:

$$R'_L = R_L^{sea} + \frac{R_L^{sez}}{p^{sez}r^{sez}}$$

where

R_L^{sea} = the number of observed recaptures recovered at sea for the PSLS fishery.

R_L^{sez} = the number of observed recaptures recovered in Seychelles for the PSLS fishery.

r^{sez} = the reporting rates for PS tags removed from the Seychelles

p^{sez} = the scaling factor to account for the EU PS recaptures not landed in the Seychelles.

The adjusted number of recaptures for a PSFS fishery was calculated similarly. The SS3 reporting parameters for the purse seine fisheries were subsequently fixed at 100% in the model. Some of the other (non purse-seine) fisheries also returned a substantial number of tags. There are no direct estimates of fishery specific reporting rates for these fisheries. The reporting rates for these fisheries are estimated within the assessment model.

Small-scale tagging programmes

Additional tag release/recovery data are available from a number of small-scale tagging programmes. The data set included a total of 7,828 tags released during 2002-08, primarily within regions 1b (70%) and 4 (28%). A total of 366 tag recoveries were reported, predominantly from the Bait boat fishery in region 1a. There has been no comprehensive analysis of these data and there is no information available concerning the fishery specific reporting rate of these tags. The tag release/recovery data from the SS tagging programmes were not incorporated in the current range of assessment models. Earlier analysis indicated that the stock assessment results were relatively insensitive to the inclusion of these data (Langley et al 2012a).

Fu et al. (2018a) investigated a range of alternative options for processing and incorporating the tagging data into the assessment model (see Table 5 of Fu et al. 2018a). These exploratory analyses are not repeated in the current assessment.

Table 2: Tag recoveries by year of recovery (box), region of release (vertical), and region of recovery. Region of recovery is defined by the definitions of the fisheries included in the model.

Recovery year	Release region	Recovery region		
		1	2	4
2005	1	21	-	-
	2	4	47	-
2006	1	2,495	29	23
	2	22	5	-
2007	1	4,127	411	-
	2	13	2	-
2008	1	1,510	277	3
	2	5	-	-
2009	1	464	61	1
	2	3	-	-
2010	1	171	5	-
	2	-	-	-
2011–2014	1	107	12	-
	2	-	-	-

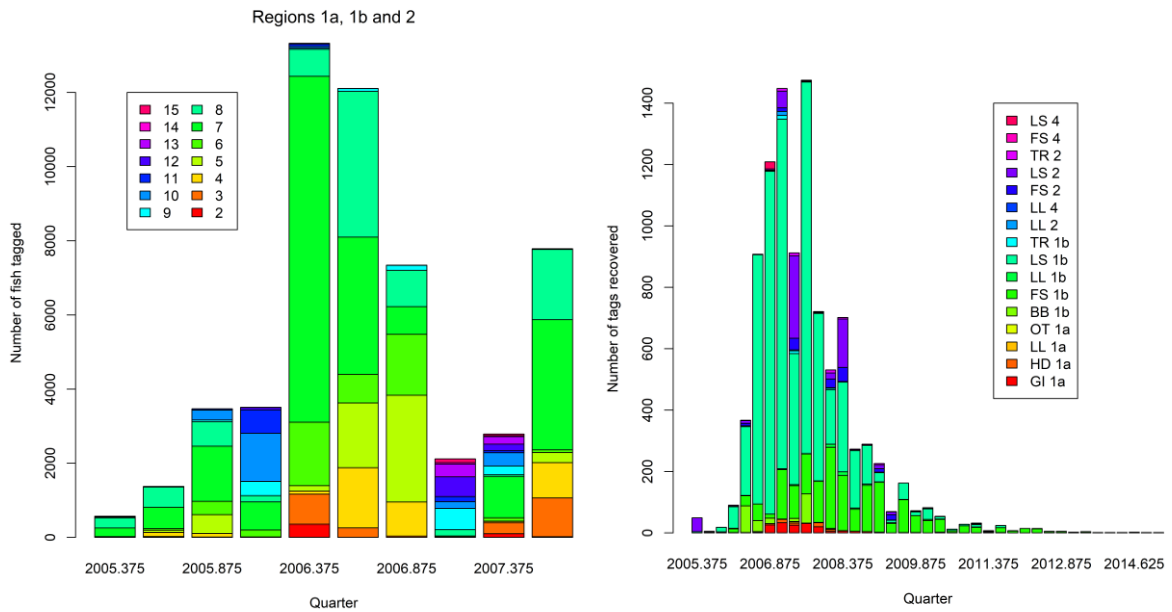


Figure 10: Number of tag releases by quarter and age class (left), and tag recoveries by year/quarter and fishery (right). Ages were assigned based on the length. Purse seine tag recoveries are shown here are not corrected for reporting rate.

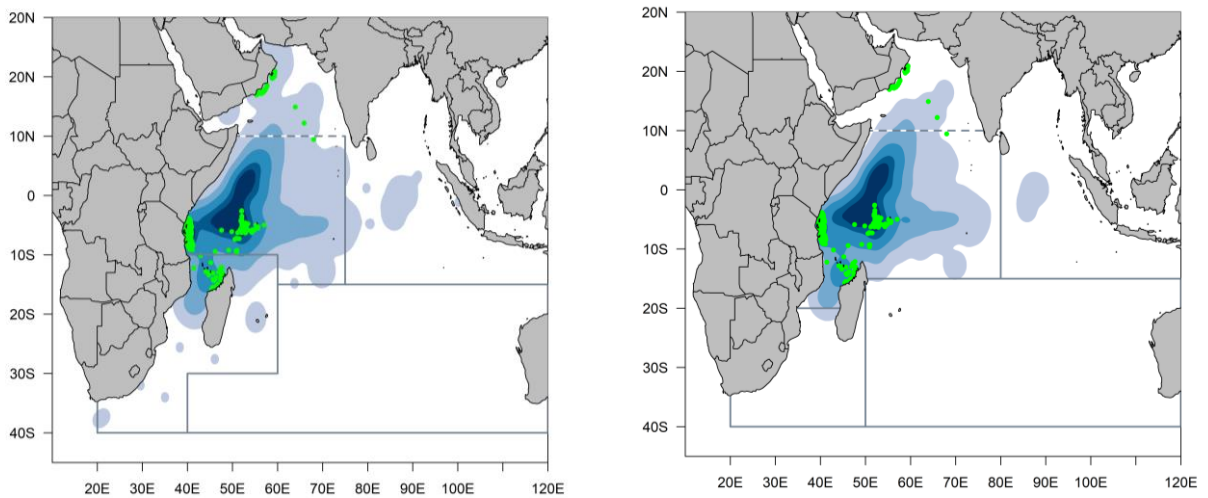


Figure 11 : Locations of releases (green) and density of recoveries for the yellowfin tuna RTTO-IO tag Program: for all recoveries (left), and for purse seine recoveries only (right, overlaid with the revised spatial region boundaries).

2.8 Environmental data

The previous assessment included a range of environmental data to investigate the potential for the incorporation of environmental covariates to inform the movement of fish. However, although there is evidence that there may be an association between the movement of yellowfin tuna and seasonal and temporal changes in ocean conditions in the Indian Ocean, the potential relationship between environmental indices (i.e., monthly current and sea temperature data) and fish movement is unclear. Analyses conducted so far, suggested that these environmental indices had no influence on the estimation of yellowfin tuna movement rates of different life stages between adjacent model regions (Langley 2016a, Fu et al. 2018a), and seasonal variation in movement may be better accounted for by models that can explicitly incorporate seasonal effects (Fu et al. 2018a). The environmental data haven't not been included in the Management Strategy Evaluation for the IO yellowfin tuna (Kolody & Jumppanen 2021). Therefore, the environmental data are not included in the current assessment.

3. MODEL STRUCTURAL AND ASSUMPTIONS

3.1 Population dynamics

The spatially disaggregated model partitions the population into four regions. The population in each region is comprised of 28 quarterly age-classes both sexes combined. The first age-class has a mean fork length of around 22 cm and is assumed to be approximately three months of age based on ageing studies of yellowfin tuna (Fonteneau 2008). The last age-class comprises a “plus group” in which mortality and other characteristics are assumed to be constant. The model commences in 1950 at the start of the available catch history. The initial population age structure in each region was assumed to be in an unexploited, equilibrium state.

3.1.1 Recruitment

Recruitment occurs in each quarterly time step of the model. Recruitment was derived from a BH stock recruitment relationship (SRR) and variation in recruitment was estimated as deviates from the SRR. Recruitment deviates were estimated for 1972 to 2019 (192 deviates), representing the period for which longline CPUE indices are available. Recruitment deviates were assumed to have a standard deviation (σ_R) of 0.6. For 1950-1971, recruitment was derived directly from the SRR. The base model assumed a level of steepness (h) of 0.8 for the SRR, an intermediate value within the plausible range of steepness values generally adopted in the tuna assessments by other tuna RFMOs (0.7, 0.8 and 0.9) (Harley 2011).

Recruitment was assumed to occur in the two equatorial regions only (region 1 and 4). This assumption was based on the temperature preference for the spawning of yellowfin tuna and a minimum temperature for larval survival of about 24°C (Suzuki 1993). The overall proportion of the quarterly recruitment allocated to region 1 and region 4 was estimated (*RecrDist_Area* parameters). Variation in the regional distribution of recruitment was included by estimating temporal deviates of the *RecrDist_Area* parameters for 1977 to 2018 (168 deviates). A time-block was imposed on the temporal deviates which were divided into two periods: 1977 – 2008 and 2009 – 2019 (both assuming a standard deviation of 1.0 for the deviates). The time block makes it possible to use the average recruitment distribution of the most recent period instead of the long-term average in the model prediction. The selection of time period is based on the estimated relative trend of regional recruitment distribution.

A sensitivity model was performed in which recruitment was assumed to occur in all model regions, which allowed the model to have more flexibility to distribute fish in model regions. This is preferable because although spawning and larvae require water of at least 24°C, the growing juveniles can move to other regions before they reach the size of recruitment to the fishery, when the model first needs to predict their distribution. Their pre-recruitment movement behaviour is likely to differ from older fish. However, there is a significant increase in the computational overhead for estimating the Hessian matrix (for variance estimation) associated with the 168*2 additional temporal deviates parameters for the regional recruitment distribution (unless a stationary regional distribution is assumed). As such, this parametrization was not used in the final model ensemble.

3.1.2 Growth and Maturation

For the current assessment, growth parameters were fixed at values that replicated the growth curve derived by Fonteneau (2008) (Figure 12). The non-von Bertalanffy growth of juvenile yellowfin tuna is evident, with slow growth for young age classes and near-linear growth in the 60–110 cm size range. Growth in length is estimated to continue throughout the lifespan of the species, attenuating as the maximum is approached. The estimated variance in length-at-age was assumed to increase with increasing age (Figure 12). Kolody (2011) cautioned that the artefact effect of the size selectivity may lead growth estimates to deviate from the von Bertalanffy growth.

Dortel et al. (2015) estimated growth by integrating otolith readings, growth increments from mark-recapture data, and modal progressions from purse seine length frequency data. These estimates were comparable to the values currently incorporated in the assessment model. The WPTT in 2019 requested that the estimates from model 2 (based on tagging and length frequency data) and model 3 (based on tagging, length frequency, and otolith readings) be investigated. The lognormal variance structure estimated by the Dortel model 3 cannot be easily represented in SS3 (this model estimates a large increase in variance near the asymptotic length), due to the limited options over the variance in length-at-age with SS3. Kolody & Jumppanen 2021 explored the possibility of approximating the lognormal variance structure using SS3 platoons (The approach partitions each age-class into multiple semi-independent units, see Methot et al. 2020). Yet this approach is computationally costly. Hence, the estimates from Model 2 were used for the current assessment. The Model 2 estimates were comparable to the values from Fonteneau (2008) from 1 to 8 quarters and then diverges, with larger length at age due to higher asymptotic length (L_{inf}) (Figure 12)

Estimates by Eveson et al. (2015) using otolith and growth increment from tag data forecast a mean asymptotic length of about 130 cm FL, which was comparatively low compared to the maximum lengths historically reported for yellowfin in the Indian Ocean (Figure 12). Sex-specific estimates from a small subset of samples supported the hypothesis that, on average, males grow to a larger size than females, with the mean asymptotic length estimate being 151 cm for males versus 140 cm for females. Similar differences between sex have been observed in Atlantic yellowfin (Pacicco et al 2021) and other *Thunnus* species. The sex-specific estimates were explored in a two-sex model in the previous assessment.

Length based maturity ogive for Indian Ocean yellowfin are available from Zudaire et al (2013). The paper presents two alternative maturity ogive based on either the cortical alveolar or vitellogenic stages of ovarian development (Figure 13). The length-based ogives were converted to age-based ogive assuming an equilibrium population age-length structure (Figure 13, derived from age-specific natural mortality, growth function and the assumed variation of length-at-age).

The maturity ogive based on cortical alveolar stage development indicates that the onset of maturity occurs at about age 5 quarters (about 75 cm) and full maturity is attained at about 12 quarters). The maturity ogive based on vitellogenic stage development is offset by about 3 quarters. The latter (older)

estimate was used in a sensitivity in the previous assessment, but it did not lead to appreciably different model results. The current assessment included only the ogive based on cortical alveolar stage development. The age-based ogive was provided to the base model as inputs of proportions mature at age. Fu et al. (2018) showed that the assessment estimates are not sensitive to whether the age-based or the length-based ogive was used.

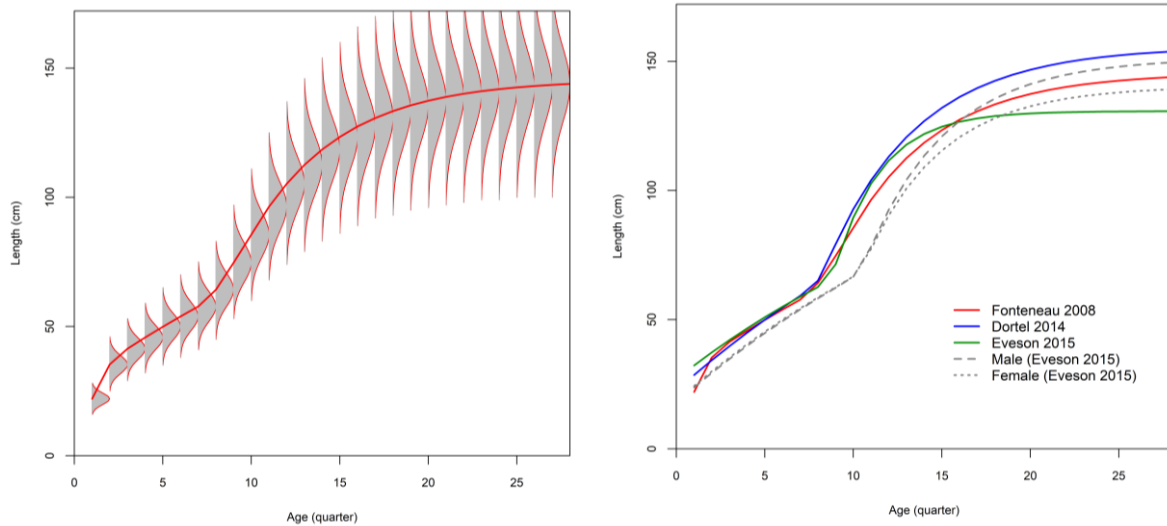


Figure 12: Fixed growth function for yellowfin tuna following Fonteneau 2008 (left – the red line represents the estimated mean length (FL, cm) at age and the grey area represents the assumed distribution of length at age), and alternative growth estimates (right).

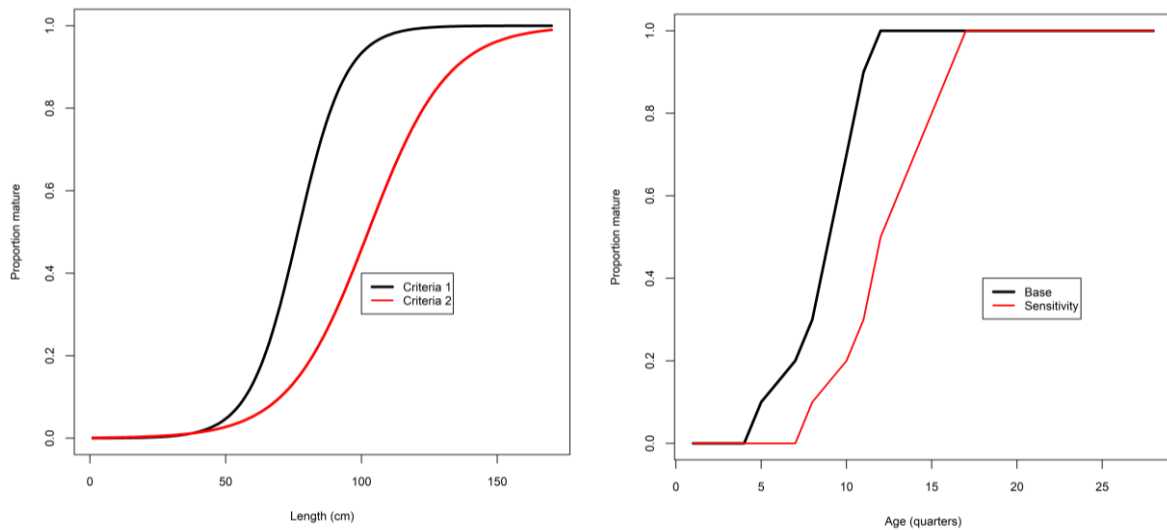


Figure 13: Length-based maturity ogive (left – from Zudaire et al 2013) and age-based maturity ogive for Indian Ocean yellowfin tuna (right – derived from Zudaire et al 2013). The ‘Sensitivity’ was examined in the previous assessment.

3.1.3 Natural mortality

Natural mortality is variable with age with the relative trend in age-specific natural mortality based on the values applied in the Pacific Ocean (western and central; eastern) yellowfin tuna stock assessments.

In the 2012 stock assessment (Langley 2012), the overall average level of natural mortality was initially fixed at a level comparable to a preliminary estimate of age-specific natural mortality from the tagging data (see IOTC 2008b). However, the overall level of natural mortality is low compared to the level of natural mortality used in the stock assessments of other regional yellowfin stocks (WCPO, EPO) (Maunder & Aires-da-Silva 2012). The WPTT considered that the IO tag data set was likely to be reasonably informative regarding the overall level of natural mortality and for the final model options the overall (average) level of natural mortality was estimated, while maintaining the relative age-specific variation in natural mortality (Langley 2012). The estimated level of natural mortality was intermediate between the initial level and the level of natural mortality adopted for the WCPFC and IATTC yellowfin stock assessments (Maunder & Aires-da-Silva 2012). The resulting age-specific natural mortality has been used as the base level of natural mortality for the subsequent stock assessments (Figure 14).

The most recent yellowfin tuna assessment in the Atlantic Ocean adopted natural mortality estimates based on the results of a study of the relationship between maximum observed age and natural mortality (Then et al 2015), and maximum age estimates derived from an aging study using annual otolith increments (Andrews et al. 2020, Pacicco et al. 2021, see also Urtizberea et al. 2020). A Lorenzen (1996) natural mortality form was developed with an M of 0.35 Yr⁻¹ for adult yellowfin, based on a validated maximum observed age of 18 years (Andrews *et al.* 2020). The level of natural mortality used for the ICCAT yellowfin assessment is slightly lower to the low M option considered in the previous assessment of Indian Ocean yellowfin tuna (Figure 14). For the current assessment, the ICCAT natural mortality is included in a model sensitivity (*M_{low}*).

Hoyle (2021) reviewed approaches for estimating natural mortality and provided estimates of age dependent natural mortality based on the method proposed by Maunder et al (submitted). The new approach assumes the Lorenzen type (Lorenzen 1996) of relationship between natural mortality and length (weight) and starts M for younger fish high and declining with age. Mortality increases after individuals become reproductively mature, and this increase is linked to the proportion mature which follows a logistic curve. Senescence is assumed to be either small or to occur at an age when too few fish alive to affect dynamics. The new approach has also been applied in the 2021 South Pacific albacore stock assessment (Castillo-Jordan *et al.* 2021). Two alternative estimates were provided: one based on a maximum age of 10.5 years observed for Indian Ocean yellowfin tuna (Shih *et al.* 2014), and the other based on the validated maximum age of 18 years obtained in the Atlantic Ocean (as the *M_{low}* option above). The range estimates by Hoyle (2021) were somewhat lower than the natural mortality options in the basic and *M_{low}* Models for the sub-adults/adults particularly for juveniles (Figure 14), and do not include a ‘hump’ to allow for higher female natural mortality 1.5 years after maturation. The hump is removed because change in sex ratio at length observed for yellowfin tuna can be partly or wholly explained by males growing larger than females (Pacicco et al. 2021, Farley this meeting). The estimates by Hoyle (2021) were examined in sensitivity models

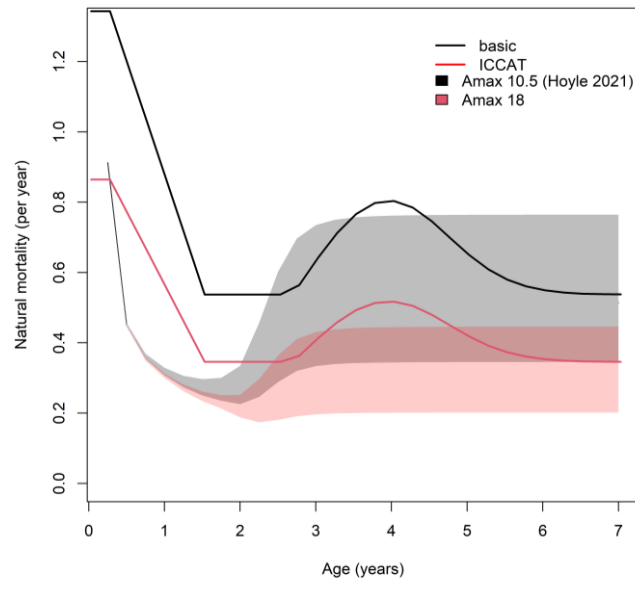


Figure 14: Age specific natural mortality patterns assumed for the basic and sensitivity (Mlow) assessment model options, as well as the range estimates by Hoyle (2021) based on maximum age of either 10.5 or 18 (the shaded area represent the 80% confidence interval of each natural mortality curve).

3.1.4 Movement

For the four region model, reciprocal movement was assumed to occur between adjacent model regions, specifically R1-R2, R1-R4, R3-R4 (3x2) (Figure 3). Movement is parameterised as the proportional redistribution of fish amongst regions, including the proportion remaining in the home region. The redistribution of fish occurs instantaneously at the end of each model time step.

Movement was parameterised to estimate differential movement for young (2–8 quarters) and old (≥ 9 quarters) fish to approximate potential changes in movement dynamics associated with maturation. Thus, for each movement transition two separate movement parameters were estimated. Fish did not commence moving until the end of age 2 quarters.

it is not possible to directly estimate seasonal movements due to the model configurations (quarterly time steps treated as calendar years). The seasonal variation in the longline CPUE indices and the purse-seine catches, particularly in region 2, indicate that there are likely to be significant seasonal changes in the regional abundance of yellowfin. Seasonal movement dynamics were investigated in the previous assessment by correlating movement parameters with the environmental indices or using an alternative model structure that can explicitly estimate seasonal movements. However, neither option was able to explain the magnitude of variability exhibited in the LL CPUE nor have any significant effect on the model results.

3.2 Fishery dynamics

Fishing selectivity for long line fisheries is assumed to be age-specific and time-invariant. A separate selectivity is estimated for each of the five LL fisheries (LL 1a, 1b, 2, 3, 4). The selectivity in each fishery is shared by corresponding set of LL CPUE indices. The longline selectivity was parameterised

with a logistic function that constrains the older age classes to be fully selected (“flat top”). The selectivity of the fresh tuna longline fishery (LF4) was estimated using a separate logistic function.

The length-based purse seine selectivities (FS and LS) in each region were formulated using a cubic spline interpolation with five nodes. The nodes were specified to approximate the main inflection points of the selectivity function. This formulation was sufficiently flexible to provide a reasonable representation of the modal structure of the length composition of the catch from the two purse seine methods, but was not adequate to account for the temporal variations in the compositions of the two length modes, particularly for the FS fisheries. For models that divide the purse seine fisheries (FS and LS) into the small (≤ 80 cm) and large (> 80 cm) fish components, a length-based, double normal selectivity was assumed for each component.

For the other fisheries, selectivity was parameterised using an age-based, double-normal function (Methot 2013). Size data were available from the troll fishery in region 4, but not from the fisheries in regions 1b and 2. The selectivity of the “Other” fisheries was assumed to be equivalent among the two regions (1a and 4), while a common selectivity was assumed for the troll fisheries in regions 1b and 4.

Fishing mortality was modelled using the hybrid method that the harvest rate using the Pope’s approximation then converts it to an approximation of the corresponding F (Methot & Wetzel 2013).

3.3 Dynamics of tagged fish

In general, the population dynamics of the tagged and untagged populations are governed by the same model structures and parameters. An obvious exception to this is recruitment, which for the tagged population is simply the release of tagged fish. The probability of recapturing a given tagged fish is the same as the probability of catching any given untagged fish in the same region. For this assumption to be valid, either the distribution of fishing effort must be random with respect to tagged and untagged fish and/or the tagged fish must be randomly mixed with the untagged fish. The former condition is unlikely to be met because fishing effort is almost never randomly distributed in space. The second condition is also unlikely to be met soon after release because of insufficient time for mixing to take place. Depending on the distribution of fishing effort in relation to tag release sites, the probability of capture of tagged fish soon after release may be different to that for the untagged fish. It is therefore desirable to designate one or more time periods after release as “pre-mixed” and compute fishing mortality for the tagged fish based on the actual recaptures, corrected for tag reporting (see below), rather than use fishing mortalities based on the general population parameters. This in effect desensitizes the likelihood function to tag recaptures in the pre-mixed periods while correctly discounting the tagged population for the recaptures that occurred.

Several analyses of the tag recovery data have been undertaken to determine an appropriate mixing period for the tagging programme (Langley & Million 2012, Kolody & Hoyle 2013). The analysis revealed that the tag recoveries from the FAD purse-seine fishery were not adequately mixed, at least during the first 6 months following release. Conversely, the free-school tag recoveries indicate a higher degree of mixing within the fished population. Most of the tagged yellowfin were in the length classes that are not immediately selected by the free-school fishery (< 90 cm). A mixing period of about 6–12 months is of sufficient duration for most tagged fish to recruit to the free-school fishery (> 90 cm) and no longer be vulnerable to the FAD fishery. However, the maximum displacements of tags reached a plateau within a few weeks of release suggesting rapid movement of yellowfin within the tag release/recovery areas. On basis of the above (see (Figure B1 of Fu et al. 2018a), it was considered that

a mixing period of 3 or 4 quarters was probably sufficient to allow a reasonable degree of dispersal of tagged fish amongst the yellowfin tuna population within the primary region of release.

The release phase of the tagging programme was essentially restricted to the western equatorial region. Fu et al. 2018a showed that the recovery rate of tags after 3 quarters at liberty was similar both in trend and magnitude between the main latitude bands within the western equatorial region (see Figure B3 of Fu et al. 2018a), which suggested a reasonable degree of mixing of tagged fish at the regional scale. The distribution of tags throughout the wider IO appears to have been relatively limited as is evident from the low number of tag recoveries from the fisheries beyond region 1b. Tag recoveries from beyond region 1b are unlikely to significantly inform the model regarding movement rates.

Estimates of tag reporting rates from the purse seine fishery were available from tag seeding trials. These estimates were applied to correct the number of tags included in the recovery dataset for the purse seine fisheries within region 1b and region 2 (see section 2.7 for details). For the other fisheries, there was very limited information available to indicate the tag reporting rates. Fishery-specific reporting rates were estimated based on uninformative priors. All fishery reporting rates were assumed to be temporally invariant and were estimated within the model.

3.4 Modelling methods, parameters, and likelihood

The total likelihood is composed of a number of components, including the fit to the abundance indices (CPUE), tag recovery data, fishery length frequency data and catch data. There are also contributions to the total likelihood from the recruitment deviates and priors on the individual model parameters. The model is configured to fit the catch almost exactly so the catch component of the likelihood is very small. There are two components of the tag likelihood: the multinomial likelihood for the distribution of tag recoveries by fleets over time and the negative binomial distribution of expected total recaptures across all regions. Details of the formulation of the individual components of the likelihood are provided in Methot & Wetzel (2013).

Following the previous assessment, the weighting of the CPUE indices followed the approach of Francis (2011). A series of smoother lines were fitted to the CPUE index and the RMSE of the resulting fit to each set of CPUE indices was determined as a measure of the magnitude of the variation of each set of CPUE indices. The resulting RMSEs were relatively high (0.40–0.50). However, a significant proportion of this variation is related to the relatively high seasonal variation in CPUE in most regions. Performing the same analysis on the annualised CPUE index resulted in considerable reduction in the RMSEs (0.15–0.2). On that basis, a CV of 0.2 was assigned to each set of CPUE indices in the base model, to ensure the stock biomass trajectories were broadly consistent with the CPUE indices while allowed for a moderate degree of variability in fitting to the indices.

The relative weighting of the tagging data was controlled by the magnitude of the over-dispersion parameters assigned to the individual tag release groups. In the previous assessment, the over-dispersion parameters for all tag release groups were set at 7.0 - determined iteratively from the residuals of the fit to the tag recovery data (observed – expected number of tags recovered). The same value was used in the current assessment.

The reliability of the length composition data is variable across fisheries and over time periods. For that reason, it was considered that the length composition data should not be allowed to dominate the model likelihood and directly influence the trends in stock abundance. Following the previous assessment, an overall effective sample size (ESS) of 5 was assigned to all length composition observations (all

fisheries, all time periods) following the Francis (2011) method. This essentially gave the entire length composition data set a relatively low weighting in the overall likelihood. Nonetheless, due to the quality of the length composition data, these data were sufficiently informative to provide reasonable estimates of fishery selectivity and provide some information regarding recruitment trends.

The weightings were applied by the values assigned to components of the likelihood of each observational dataset included in the total model likelihood. A default lambda of 1.0, represented the native weighting of the data. A lower value of Lambda would effectively downweigh a dataset relative to other observations, effectively reducing its influence on the overall model estimates. A lower lambda was applied to the tagging data in a sensitivity model.

The Hessian matrix computed at the mode of the posterior distribution was used to obtain estimates of the covariance matrix, which was used in combination with the Delta method to compute approximate confidence intervals for parameters of interest.

4. ASSESSMENT MODEL RUNS

The approach we have taken here is to explore a range of model assumptions and parameter configurations, and to examine areas of uncertainty that would impact assessment results. The model runs included a basic model that represents a continuity run from the 2018 assessment, and a range of exploratory analyses. Final model options included a grid of models running over permutations of plausible parameters and/or model settings, from which the uncertainty was quantified. The grid approach aims to provide an approximate understanding of variability in model estimates due to assumptions in model structure, which is usually much larger than the statistical uncertainty conditional on any individual model (Polacheck et al 1997, McKechnie et al. 2016, Kolody et al. 2011). The assessment was conducted using the 3.30.17 version of the Stock Synthesis software. The stock status was reported for the terminal year of the model (i.e., 2020).

4.1 The basic model

In the 2018 assessment, final model options selected for quantify stock status by the WPTT included 24 models with alternative assumptions on levels of steepness, tag weighting, tag release mortality, and LL CPUE catchability assumption (IOTC 2018a). The model *io_h80_q1_tm30_dw1* (steepness of 0.8, constant catchability, tag release mortality of 27.5%, and tag likelihood lambda of 1) was considered as a reference model. The 2018 reference model was updated to ensure continuity, but many revisions were incorporated based on the analysis conducted by the yellowfin task force team since the last assessment. These revisions are to improve the configuration of the assessment model, but do not represent any major structural changes to the model. The revisions included:

- Optimising the parametrizations of regional recruitment and movement by removing redundant parameters
- Changing the age-based selectivity to size-based for the purse seine fisheries (to improve the estimates of the 5-node cubic spline function)
- Simplifying the fleet structure by removing the temporal split of the purse seine fisheries in region 1, reducing the number purse seine fisheries in region 1 from 6 to 2.
- The length data from the Taiwanese and Seychelles longline logbooks were excluded. The Pre-1965 length data from the Japanese fleet were also excluded.

- The increase of length bin size from 2 to 4 cm
- Estimating separate/independent longline selectivity in each region
- Environment data are not included.
- Update of Stock Synthesis from version 3.24 to 3.30.17

Ideally, a sequential, step-wise approach should be taken to better understand the effects of these changes on model estimates (e.g., spawning biomass, recruitment trends, movement etc.). However, as the model development process has taken iteratively over a long span of time and the input datasets were finalised very close to the WPTT meeting date, as such, only the main source of impact was identified (see Figure 29).

4.2 Exploratory model runs

The 2018 assessment performed reasonably well in explaining the observational data but also revealed several problems. The assessment model showed a certain degree of instability and the likelihood function seemed to have local minimums which was thought to be related to the estimation of movement. The model had a retrospective pattern in which the model estimated a lower productivity as the most recent data were sequentially excluded. Kolody & Jumppanen (2021) found that the assessment model had a structure problem that meant it was unable to explain the newly reported catches when projecting forward. These issues are likely to be related to various aspects of the model configuration (including the SS3 assumption on future recruitment regional distribution that is now fixed). It is known that movement parameters are not well estimated as the tagging data are not very informative, and the estimation are influenced by other sources. Further the model cannot account for seasonal changes in movement, catchability, and selectivity, which are probably confounded. Therefore, a number of exploratory models were performed to investigate a range of options related to the configuration of key data sets, model structure and parameters. The analysis extended the exploratory models of the previous assessment. A description of the range of alternative model options considered is presented in Table 4.

4.3 Final model ensemble (grid)

On basis of the exploratory analysis, final options were configured to capture the uncertainty related to assumptions on the spatial structure, longline catchability, stock-recruitment steepness, tag data weighting, and biological parameters including growth and natural mortality, which are thought to contribute to the main sources of uncertainty. Thus, the final models involved running a combination of options on spatial structure (2 options), LL catchability (2 scenario), steepness (3 values), growth (2 values), natural mortality (2 levels), and tag lambda (2 values) (Table 5). The final model grid is different from the assumptions of the 2018 assessment: it omits the alternative, lower tag release mortality value, but adds alternative options for spatial structure and biological parameters. The final models with the alternative spatial structure included the purse seine free school CPUE on adults.

Table 3: Main structural assumptions of the yellowfin tuna basic model and details of estimated parameters. Changes to the 2018 base model are highlighted in red.

Category	Assumptions	Parameters
Recruitment	<p>Occurs at the start of each quarter as 0 age fish.</p> <p>Recruitment is a function of Beverton-Holt stock-recruitment relationship (SRR).</p> <p>Regional apportionment of recruitment to R1 and R4 only.</p> <p>Temporal recruitment deviates from SRR, 1972–2018.</p> <p>Temporal deviates on proportion of recruitment allocated to R1 and R4, 1977–2018, with separate mean for time-block 1977–2008, and 2009–2018</p>	<p>R_0 Norm(10,2); $h = 0.80$</p> <p>$PropR1$ Norm(1.5,0.25); $PropR4$ Norm(0.5,0.25)</p> <p>$SigmaR = 0.6$. 188 deviates.</p> <p>Deviates Norm(0,1), 336 deviates.</p>
Initial population	<p>A function of the equilibrium recruitment in each region assuming population in an unexploited state prior to 1950. Initial fishing mortality fixed at zero for all fisheries.</p>	
Age and growth	<p>28 quarterly age-classes, with the last representing a plus group.</p> <p>Growth based on VonBert growth model with age-specific k to approximate the mean length at age determined by Fonteneau (2008).</p> <p>SD of length-at-age based on a constant coefficient of variation of average length-at-age.</p> <p>Mean weights (W_j) from the weight-length relationship $W = aL^b$ (source IOTC-2016-WPDCS12-INF05).</p>	<p>$L_{infinity} = 145\text{cm}$, k (base) = 0.455, k deviates for ages 2–13.</p> <p>CV =0.10</p> <p>$a = 2.459 \text{ e-}05$, $b = 2.9667$</p>

Natural mortality	Age-specific. Relative variation amongst ages based on WCPO yellowfin assessment and overall scale of natural mortality estimated in 2012 IO yellowfin assessment (see Figure 13). Constant over time and among regions.	
Maturity	Age-dependent, specified. Derived from length based maturity ogive in Zudaire et al (2013). Mature population includes both male and female fish (single sex model).	age-class 0-4: 0; 5: 0.1; 6: 0.15; 7: 0.2; 8: 0.5; 9: 0.5; 10: 0.7; 11: 0.9; 12-28: 1.0
Movement	Age-dependent with two blocks; age classes 2-8 and 9-28. Constant among quarters.	10 movements * 2 age blocks. Norm(0,4).
Selectivity	constant over time. Separate age-based, logistic selectivity parameters for Principal longline fisheries. Common length-based, selectivity for all PSLS fisheries. Common length-based, selectivity for all PSLS fisheries. LF4 fishery logistic selectivity. HD 1a fishery logistic selectivity All other fisheries: double normal selectivity. OT 1a & 4 and TR 1b & 4 share selectivity parameters. GI 1a estimated selectivity for time blocks before and after 2000;	Logistic $p1$ Norm(14,1), $p2$ Norm(4,1) Five node cubic spline Five node cubic spline
Catchability	Constant over years and among regions for LL CPUE indices. CPUE indices are scaled to reflect different region sizes. Scaling factors were revised using the '7994m8' estimates from Hoyle & Langley (2018). No seasonal variation in catchability for LL CPUE.	Unconstrained (nuisance) parameter LLq

	LL CPUE indices have CV of 0.2 for all regions.	
Fishing mortality	Hybrid approach (method 3, see Methot & Wetzel 2013).	
Tag mixing	Tags assumed to be randomly mixed at the model region level four quarters following the quarter of release. Accumulation after 16 quarters	
Tag loss	Initial tag loss represents 27.5% tag mortality (Hoyle et al. 2015) Chronic tag loss represents tag shedding of 20% over 2000 days (Gaertner & Hallier). Applied to all tag release groups.	Parameter -3.5
Tag reporting	All (corrected) reporting rates constant over time. Common tag reporting rate fixed for all PS fisheries. Non PS tag reporting rates	Uninformative priors.
Tag variation	Over dispersion parameter of 7.0. Applied to all tag release groups.	Tag OD 7.0
Length composition	Multinomial error structure, all length samples assigned ESS of 5.0.	

Table 4: Description of the exploratory runs for the 2021 assessment.

<i>Model</i>	Description
Spatial structure	
<i>Revised</i>	A four-region model adopting the revised regional and fleet structure (see Figure 2, table 1 in Appendix A); The model included LL CPUE indices and tagging data from the purse seine fisheries.
<i>bRecruit4</i>	Basic model, allowing recruitment into each of the four regions (R1, R2, R3, R4)
Length data	
<i>bLFdw</i>	Basic model, but down-weighting the LL length frequency data 2003–2020 (sample size fixed at 0.1)
<i>bLFsplit</i>	Basic model, but splitting the four LL fisheries (LL1b, LL2, LL3, and LL4) before and after 2003, estimating a double-normal selectivity for 1950 – 2003, and a logistic selectivity for 2004 – 2020 (independently for each fishery).
<i>bSelrw</i>	Basic model, but estimating deviations on the selectivity parameters that follow a random walk for each of the four LL fisheries after 2003
CPUE	
<i>bLL1bSplit</i>	Basic model, removing the LL1b CPUE indices in the piracy period 2007 – 2011, and estimating separate catchability for the LL1b CPUE before 2007 and after 2011.
<i>b1990</i>	Basic model, starting all four LL CPUE from 1990 (removing indices 1975 – 1989)
<i>rCPUEFSq</i>	Revised model, including the PSFS CPUE on adults for region 1 which had been adjusted for a 3.15% annual catchability increase
<i>rCPUEFSpe</i>	Revised model, including the PSFS CPUE on adults for region 1, estimating an additional variance component for both the PSFS and LL1b CPUE
<i>rCPUEFSLSBBpe</i>	Revised model, including both PSFS CPUE (adults and juveniles), the Associate Behaviour Buoyancy Indices (ABBI), and Pole and Line indices (PL), estimating an additional variance component for each index
Seasonal structure	
<i>bSeason</i>	Basic Model, splitting each of the four LL fisheries (LL1b, LL2, LL3, and LL4) into 4 seasonal fisheries (16 in total) and estimate independent selectivity in

<i>season1, season2, season3, season4</i>	each quarter; splitting each of the four LL CPUE into 4 seasonal indices (16 in total) and estimating independent catchability for each quarter Four sub-models which includes LL size composition data and CPUE from one of quarter 1,2,3, and 4, respectively
<i>bYearSeason</i>	Model configured using the SS3 internal year-season structure in which an annual cycle consists of 4 seasons (3 month each);
Biological parameters	
<i>bDortel</i>	Growth estimates by Dortel et al. (2015) approximated by age-varying <i>k</i>
<i>bMlow</i>	Age-specific natural mortality derived from aging study in the Atlantic Ocean. Overall level of <i>M</i> approximately 60% of the base level.
<i>bMAmax10</i>	Age-specific natural mortality by Hoyle (2021) using the proposed method by Maunder et al. (submitted), based on a maximum age of 10.5 observed in IO
<i>bMAmax18</i>	Age-specific natural mortality by Hoyle (2021) using the proposed method by Maunder et al. (submitted), based on a maximum age of 18 observed in IO

Table 5: Description of the proposed final model options for the 2021 assessment. The model grid consists of a full combination of options below, with a total of 96 models . The options adopted for the basic model is highlighted.

Model options	Description
Spatial structure	<ul style="list-style-type: none"> io – regional structure adopted in the basic model sp – the revised regional & fleet structure, including the PSFS CPUE on adults with additional variance estimated.
Steepness	<ul style="list-style-type: none"> h70 – Stock-recruitment steepness parameter 0.7 h80 – Stock-recruitment steepness parameter 0.8 h90 – Stock-recruitment steepness parameter 0.9
Tag weighting	<ul style="list-style-type: none"> <i>TagLamda01</i> – Tag lambda = 0.1 for both components of tag likelihood. <i>TagLamda1</i> – Tag lambda = 1 for both components of tag likelihood.

<i>LL1b catchability</i>	<ul style="list-style-type: none"> • <i>q1</i> – a single catchability for LL1b CPUE.
	<ul style="list-style-type: none"> • <i>q2</i> – removing the LL1b CPUE indices 2007 – 2011 and estimating separate catchability before 2007 and after 2011.
<i>growth</i>	<ul style="list-style-type: none"> • <i>Gbase</i> – VonBert growth model with age-specific k to approximate the mean length at age determined by Fonteneau (2008).
	<ul style="list-style-type: none"> • <i>GDortel</i> – VonBert growth model with age-specific k to approximate the mean length at age determined by Dortel et al. (2014).
<i>Natural mortality</i>	<ul style="list-style-type: none"> • <i>Mbase</i> – base level natural mortality.
	<ul style="list-style-type: none"> • <i>Mlow</i> – A natural mortality (by age) that is 70% of the base level natural mortality (see Section 5.2).

5. MODEL RESULTS

5.1 The basic model

The basic model proposed here is for the purpose of facilitating the discussions of model diagnostics and performance and is not intended as the final model for providing management advice (which shall be determined by the Working Party on Tropical Tunas after deliberations of all model options explored during the assessment).

Preliminary models showed that the high CPUE peak 1976–78 resulted in estimation of unrealistic biomass pulses in the late 1970s. These high peaks are more likely to be related to data reporting (see 2.5.1). Therefore, the few exceptionally high CPUE indices (if they indicate more than 5 folds of quarterly abundance changes) from 1976–78 for regions 1– 4 were removed (alternatively they can be assigned a very large standard deviation).

5.1.1 Model fits

The model provides a reasonable fit to the overall trend in the CPUE indices for each region (Figure 15). The CPUE indices exhibit a high degree of seasonal variability that is not estimated by the model. There is no discernible temporal trend in the residuals from the fit to the CPUE indices. However, there is a slight downward trend in the residuals in region 4 (Figure 16), as the LL4 CPUE appears to decline more than other regions in the long term. For LL1b, there are more negative residuals in the CPUE fits over the last 8 years. The model estimated a low degree of mixing between regions (see Section 5.1.2), implying the population trends are relatively independent amongst regions. The larger decline in the CPUE index for the tropical regions over the data period appeared to be consistent with the exploitation history in the regions. The model was unable to adequately account for the very high CPUE indices in region 1b in 2005.

For most fisheries, there is a reasonable overall fit to the length composition data (Figure 17). For the main longline fisheries (LL1a, 1b, 2–4), although the model fit the long-term trends in the average length 1960s – 2000s reasonably, it does not capture well the increase in fish size during more recent years (Figure 18). Before 2003, the model appears to have under-predicted fish in the size range 100 –

140cm, but overestimated fish above 140 cm in the length composition, but the opposite is true after 2003 for LL 1b, 2, and 4 (Figure 19). This may indicate that selectivity for the longline fisheries in the early years are dome shaped. For the main purse seine fisheries (particularly the PSFS), the relative proportion of fish in the small (≤ 80 cm) and large (> 80 cm) length mode is variable over time, probably due to size related schooling behaviour of adult yellowfin tuna, resulting in strong residual patterns in the fits (Figure 20). The recent trends in the predicted average fish size for the PSFS1b and PSFS2 fisheries are broadly consistent with the sampling data with larger fish caught during the mid-2000s and smaller fish from 2010 onwards. There is a marked decline in the average size of fish sampled from the purse seine FAD fisheries in both region 1b and region 2 (Figure 18), particularly during the mid-1990s. This trend is not evident in the predicted average fish size derived from the model for region 2. There is an improvement in the fits to the length data from the handline and gillnet fisheries in region 1a.

A comparison of the observed and predicted numbers of tags recovered from each fishery (excluding recoveries during the four-quarter mixing period) by quarterly time period for each release group are presented in Figures Figure 21–Figure 24. Overall, the model provides a poor fit to the tag recoveries during the main recovery period (2007–2009). Most of the tag returns were from the purse-seine fishery in region 1b, to a lesser extent, region 2. In region 1b, there are a number of quarters when the model substantially underestimates the number of tag recoveries from both regional purse seine fisheries. These quarters correspond to the first quarter following the four-quarter mixing period for the large releases of tags in 2006 (quarters 2, 3 and 4) and 2007 quarter 3 (see Figure 21). The lack of fit to the recoveries in those quarters suggests that even the four-quarter mixing period may not be sufficient to allow for adequate dispersal of tagged fish in the population. The lack of fit is also spread though time which may indicate that the fishing mortality estimate may be too low and biomass too high, and/or the natural mortality may be too high. The tag returns in region 2 are highly variable, reflecting the small volume and the high seasonality of the catch in the region. In some cases, the lack of fit to the observed recoveries occurred near the boundary between region 1 and 2 and is possibly influenced by where the boundary is drawn. The maximum tag returns occurred around three to four quarters following release for the purse seine sets on associated schools and around six to eight quarters for the free schools, reflecting the differences in the age of recruitment into these fisheries. In region 1b, the tags remained vulnerable to purse seine associated sets for an extended period. This seems to contradict the length composition data, which indicates that the susceptibility of adult yellowfin tuna in the associated sets is low (see Figure B8). The Tag recoveries from the non-purse seine fisheries are not considered to be very informative and the model has the flexibility to freely estimate reporting rates for these fisheries. Of these fisheries, only the LL fisheries in region 1b and region 2 recovered moderate numbers of tags during the period following the four-quarter mixing phase. The numbers of tags recovered from these fisheries was low relative to the purse-seine fishery and the fishery specific tag reporting rates were estimated to be very low.

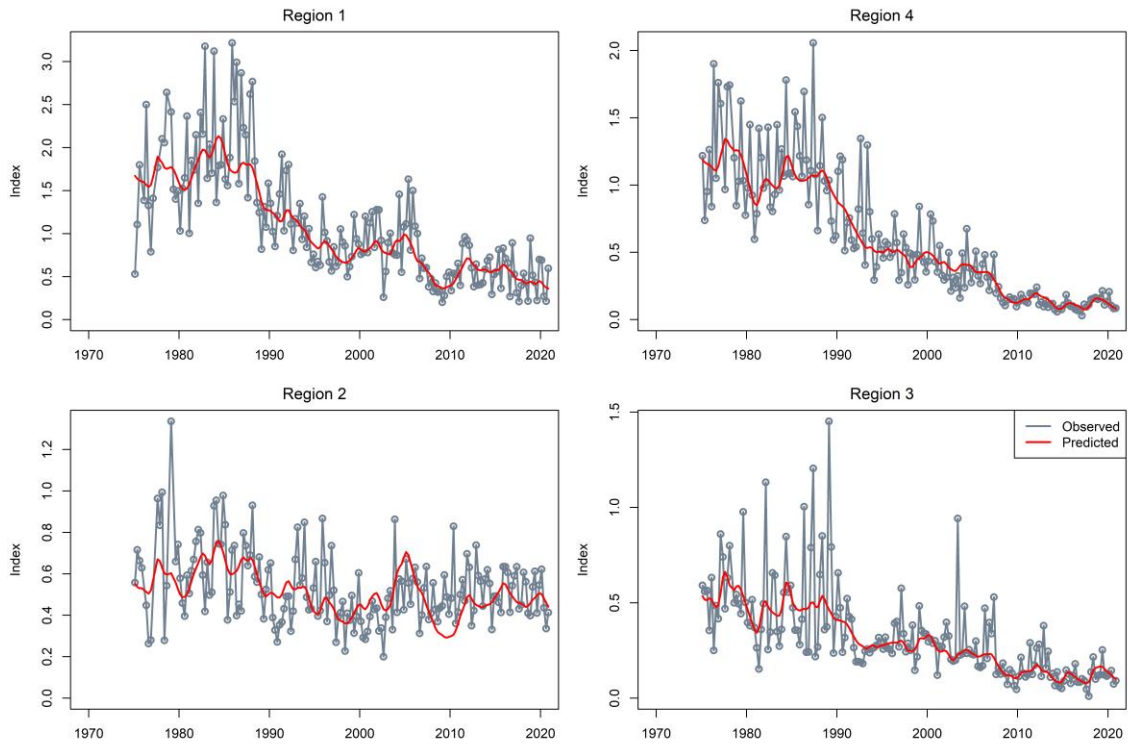


Figure 15: Fit to the regional longline CPUE indices, 1975–2020 from the basic model.

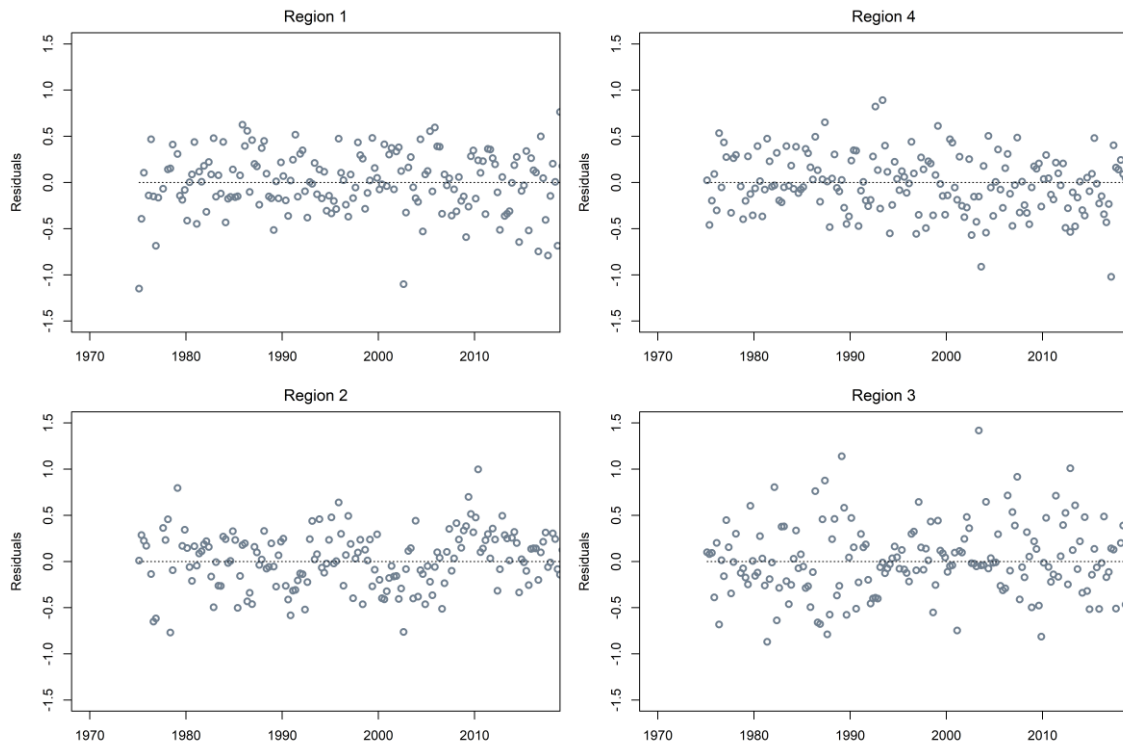


Figure 16: Standardised residuals from the fits to the CPUE indices from the basic model.

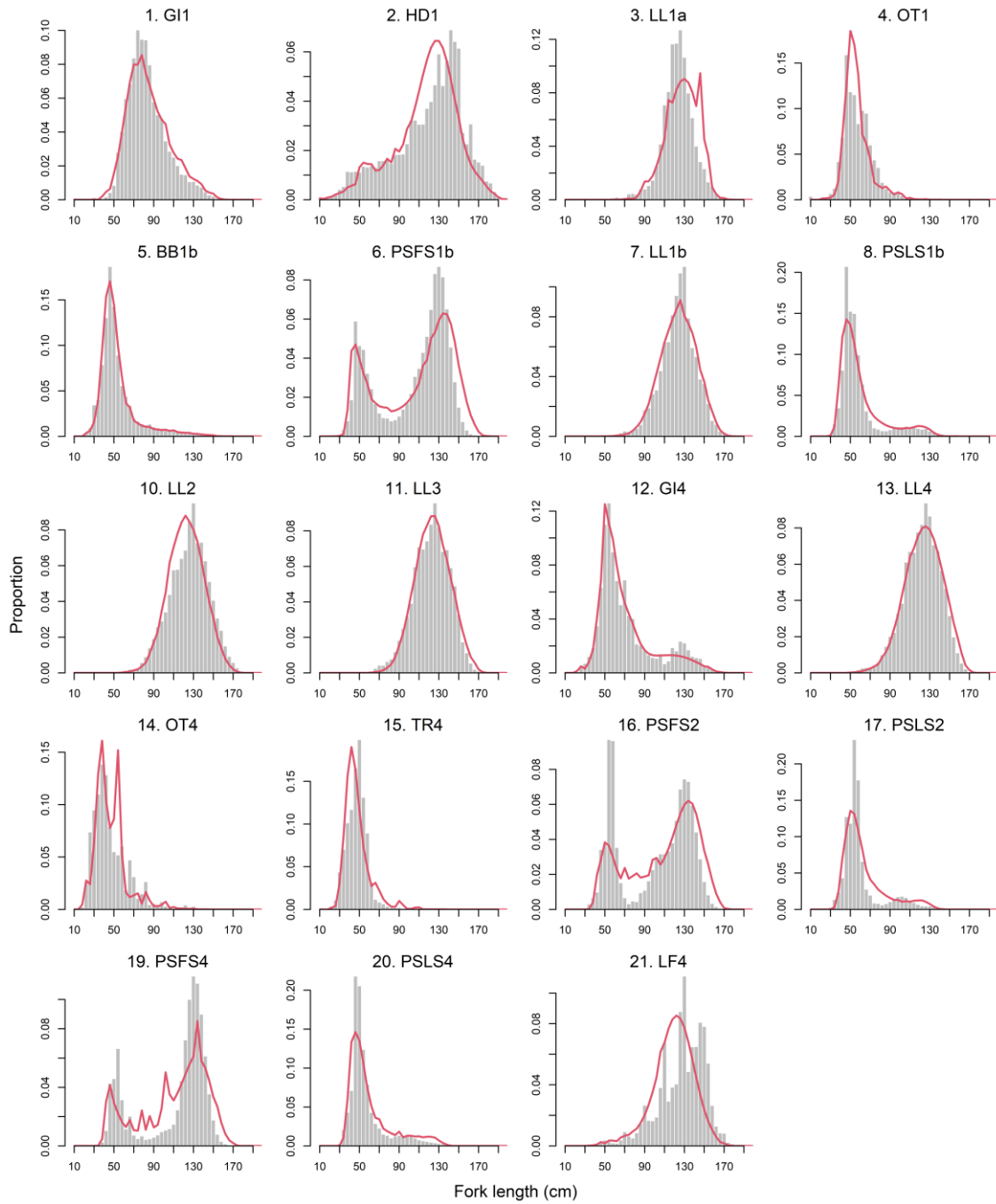


Figure 17: Observed (grey bars) and predicted (red line) length compositions (in 4 cm intervals) for each fishery of yellowfin tuna aggregated over time for the basic model.

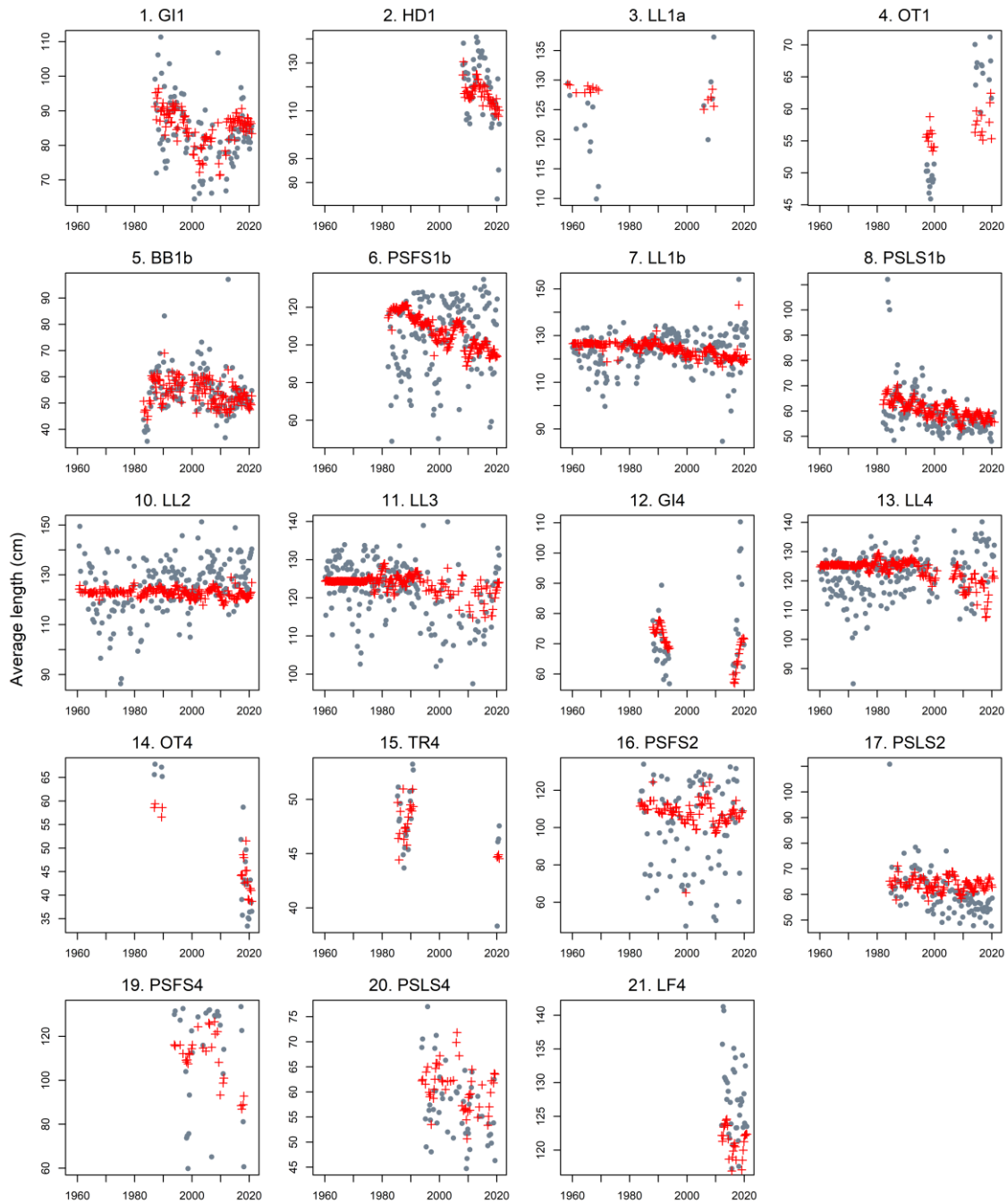


Figure 18: A comparison of the observed (grey points) and predicted (red points and line) average fish length (FL, cm) of yellowfin tuna by fishery for the basic model.

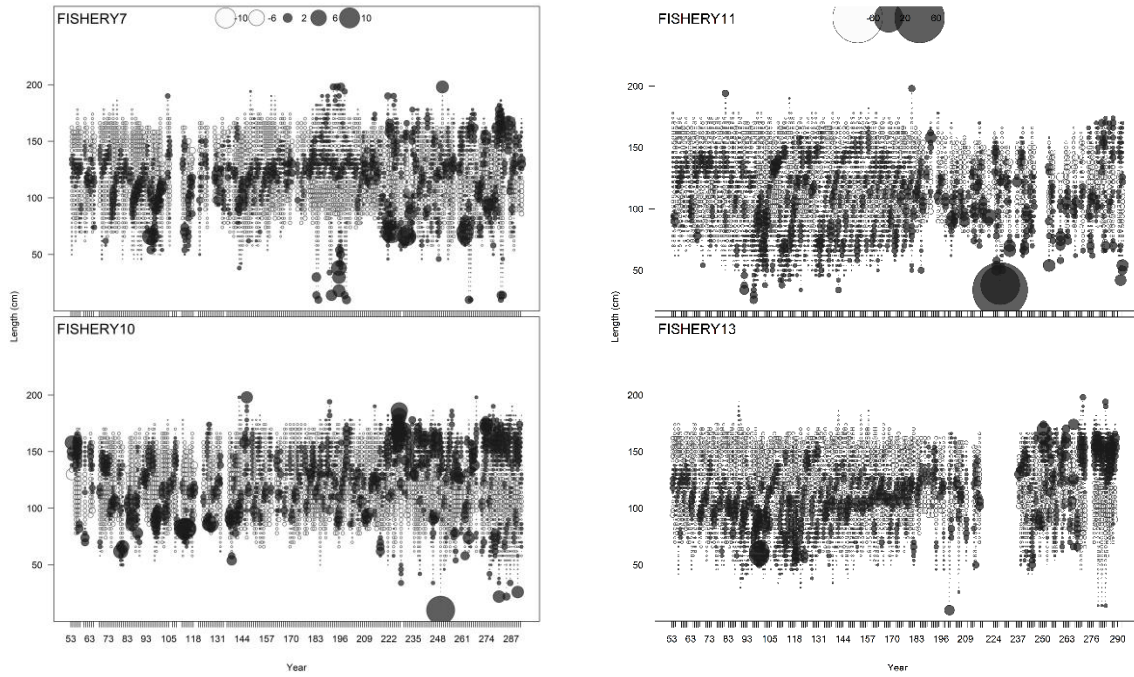


Figure 19: Relative residuals from the fits to the length compositions for LL fisheries 7 (R1b), 10(R2), 11(R3), and 13 (R4) for the basic model.

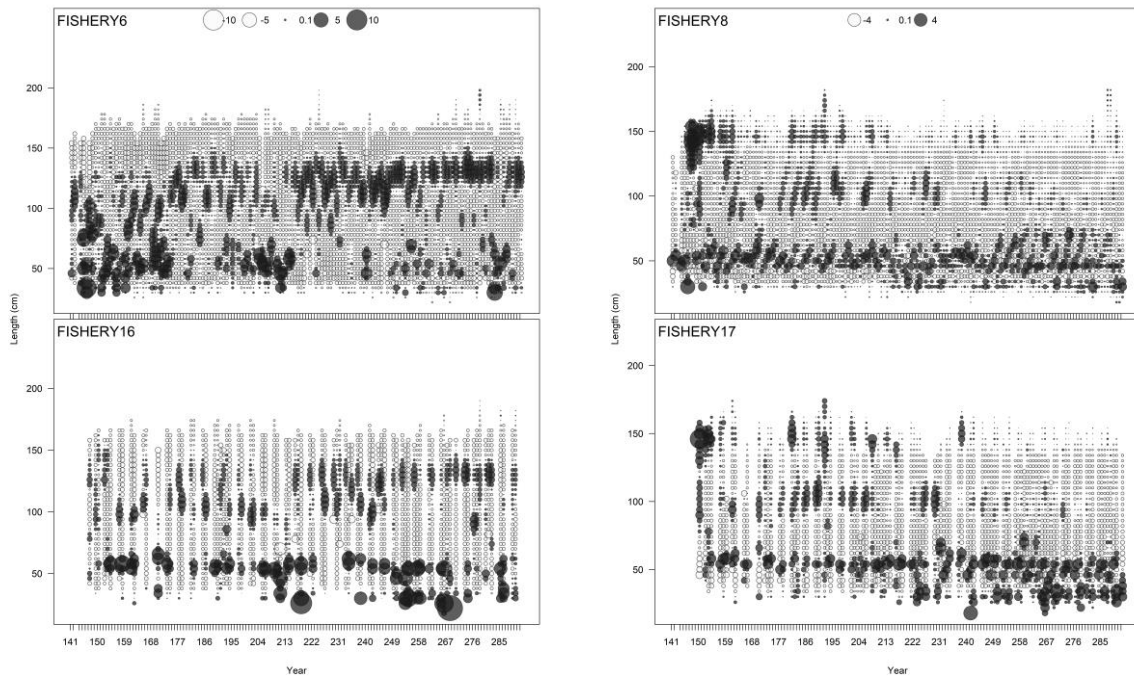


Figure 20: Relative residuals from the fits to the length compositions for PSFS fisheries 6 (R1b) and 16 (R2), and PLS fisheries 8 (R1b) and 17 (R2) for the basic model.

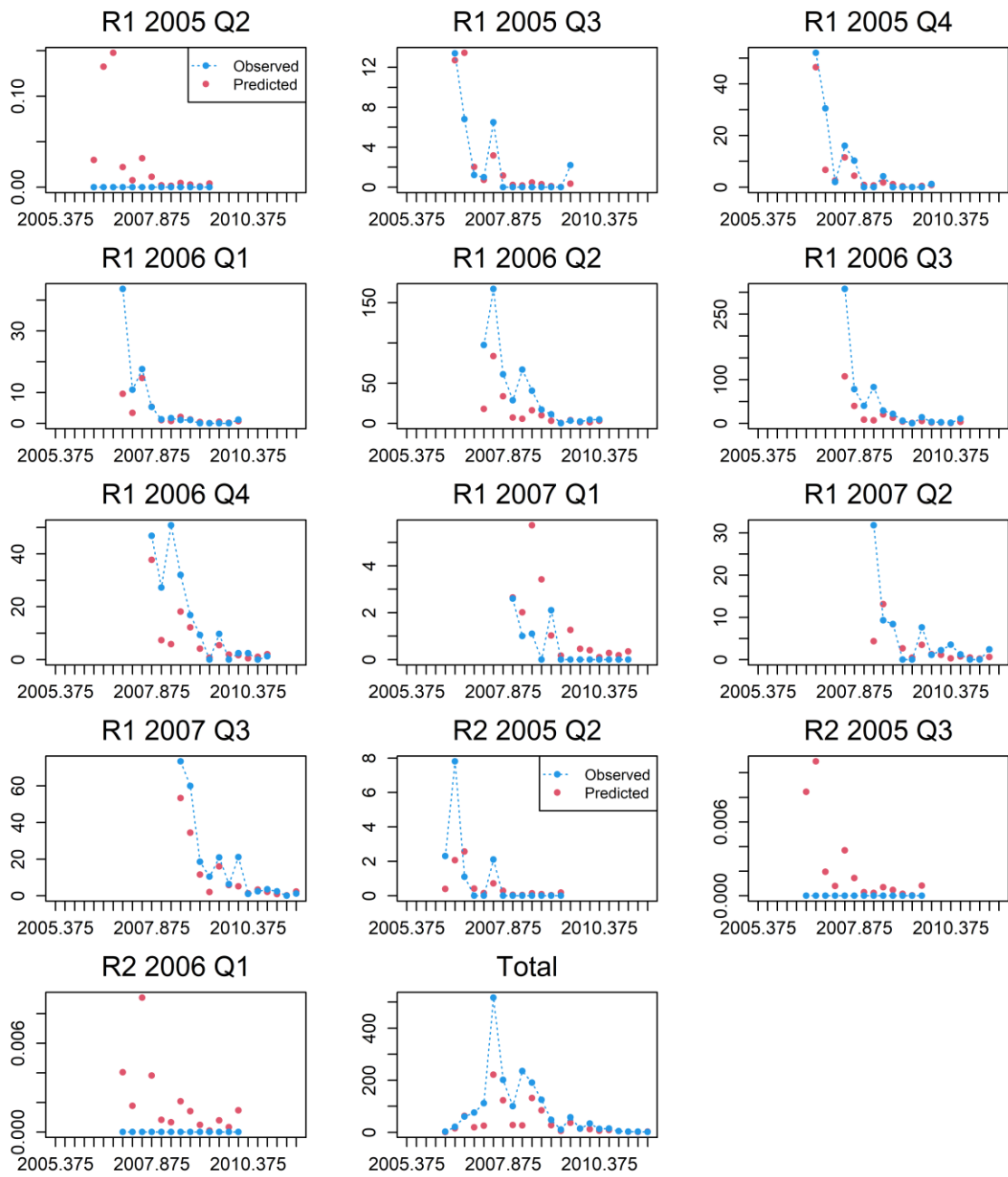


Figure 21: Observed and predicted number of tags recovered by quarter for the PSLS fishery in region 1b (PSLS 1b). Only tags at liberty after the four quarter mixing period are included. Tag recoveries are aggregated for each release group.

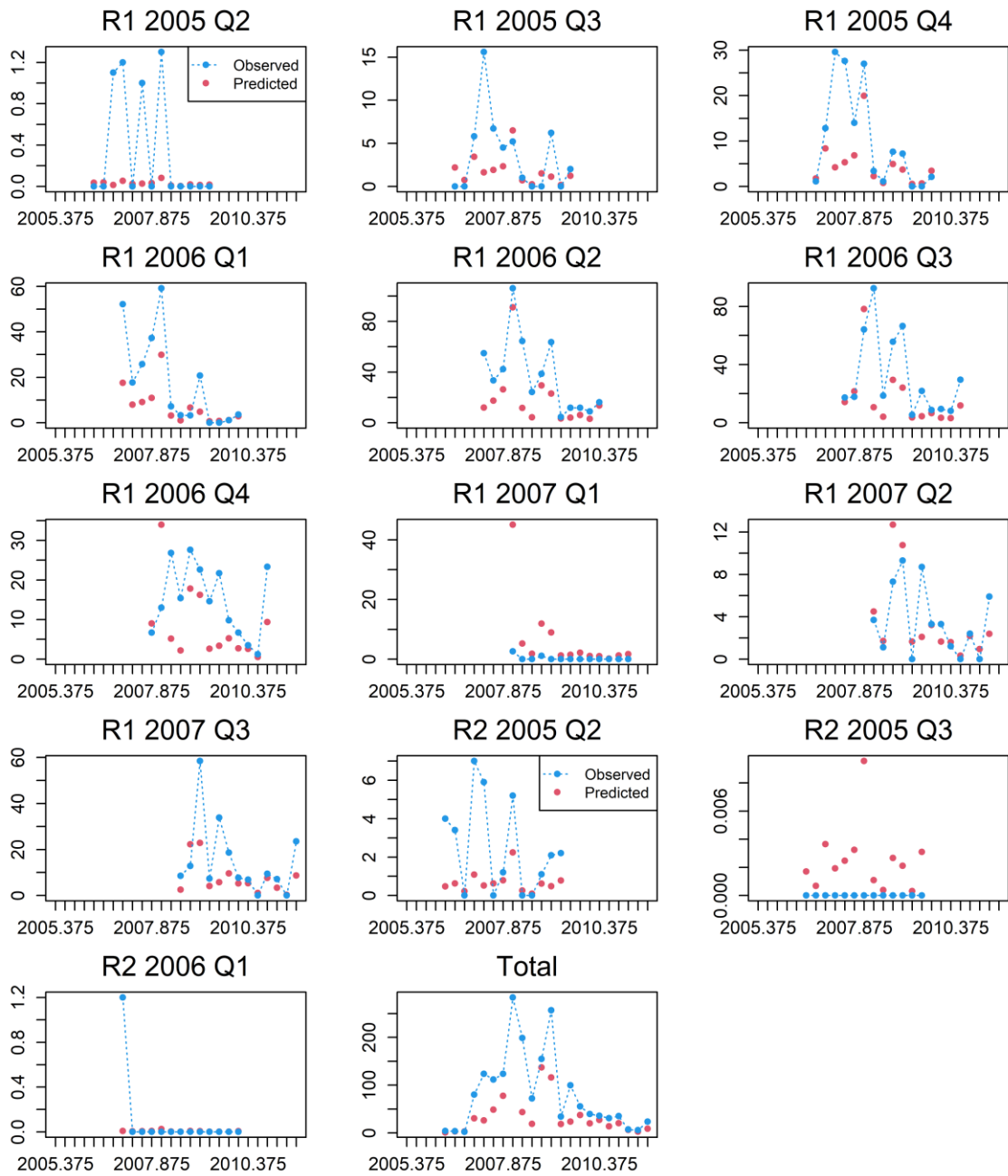


Figure 22: Observed and predicted number of tags recovered by quarter for the PSFS fishery in region 1b (PSFS 1b). Only tags at liberty after the four-quarter mixing period are included. Tag recoveries are aggregated from each release group (region, year, and quarter).

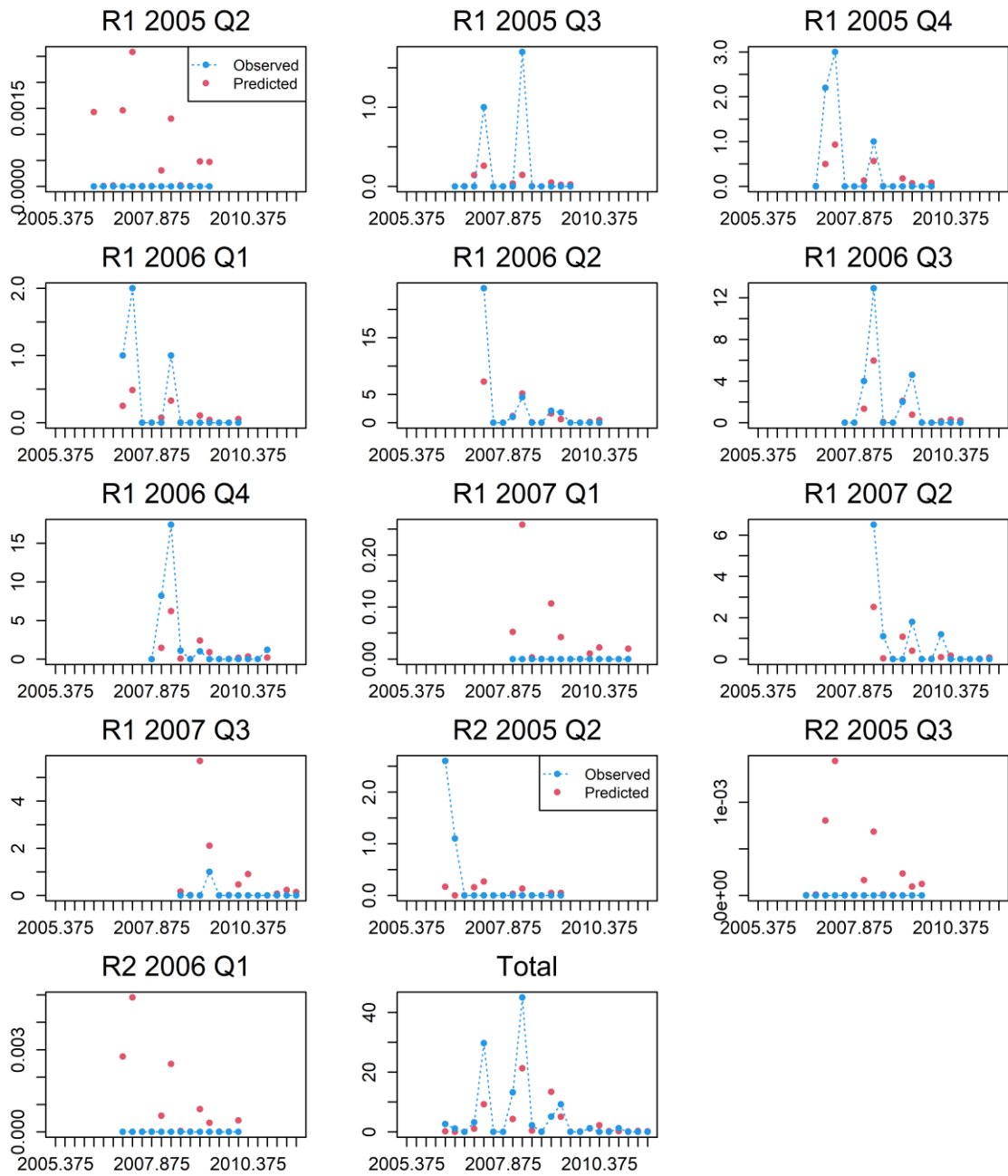


Figure 23: Observed and predicted number of tags recovered by quarter for the PSLS fishery in region 2 (PSLS 2). Only tags at liberty after the four -quarter mixing period are included. Tag recoveries are aggregated from each release group (region, year, and quarter).

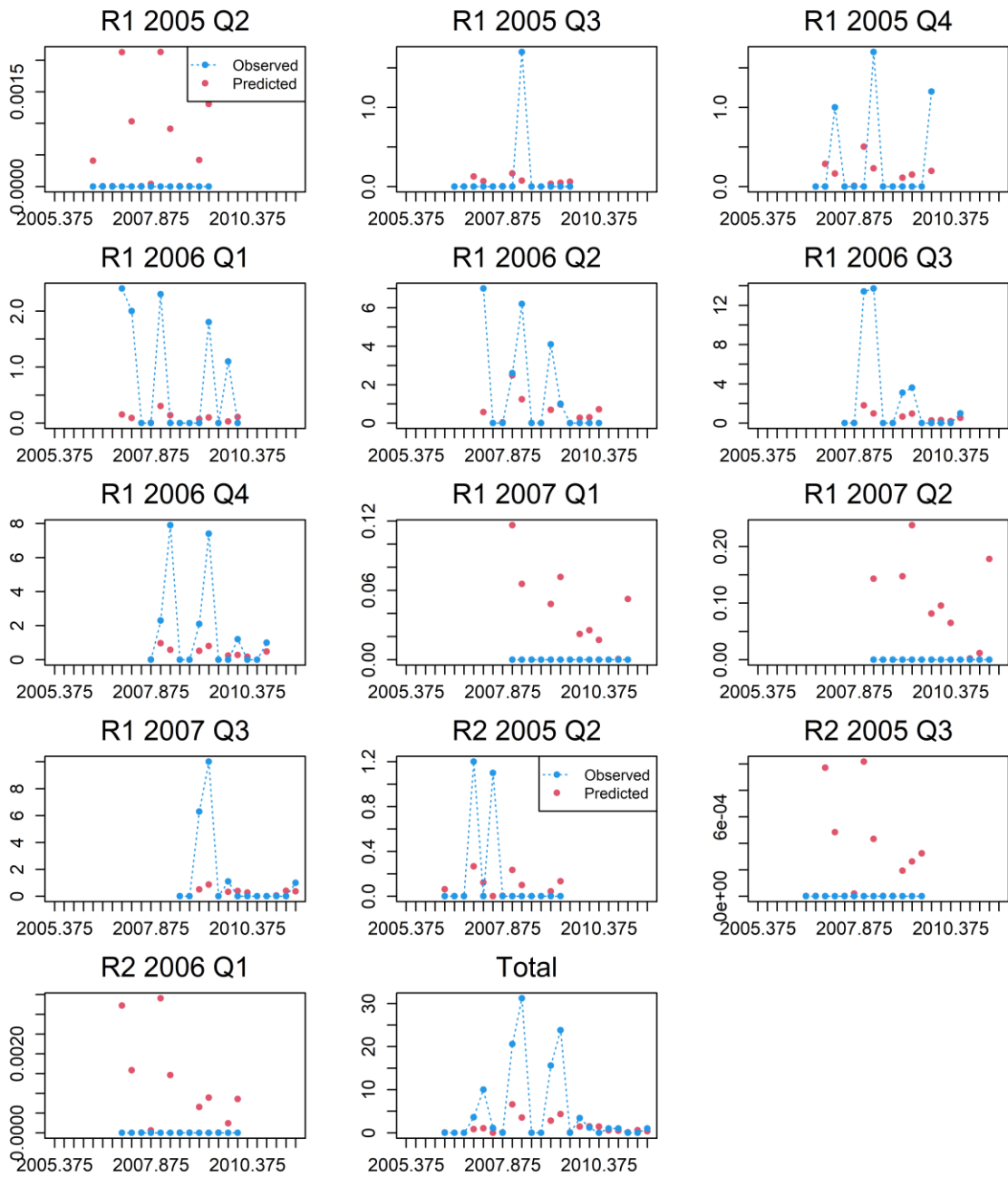


Figure 24: Observed and predicted number of tags recovered by quarter for the PSFS fishery in region 2 (PSFS 2). Only tags at liberty after the four-quarter mixing period are included. Tag recoveries are aggregated from each release group (region, year, and quarter).

5.1.2 Model estimates

The estimated parameters in the basic model include: the overall population scale parameter R_0 , the time series of recruitment deviates, the distribution of recruitment among regions, age specific movement parameters, the fishery selectivity parameters, fishery tag reporting rates and the catchability parameters for the CPUE indices.

The age-based selectivity functions (except for purse seine fisheries which are length-based) are presented in

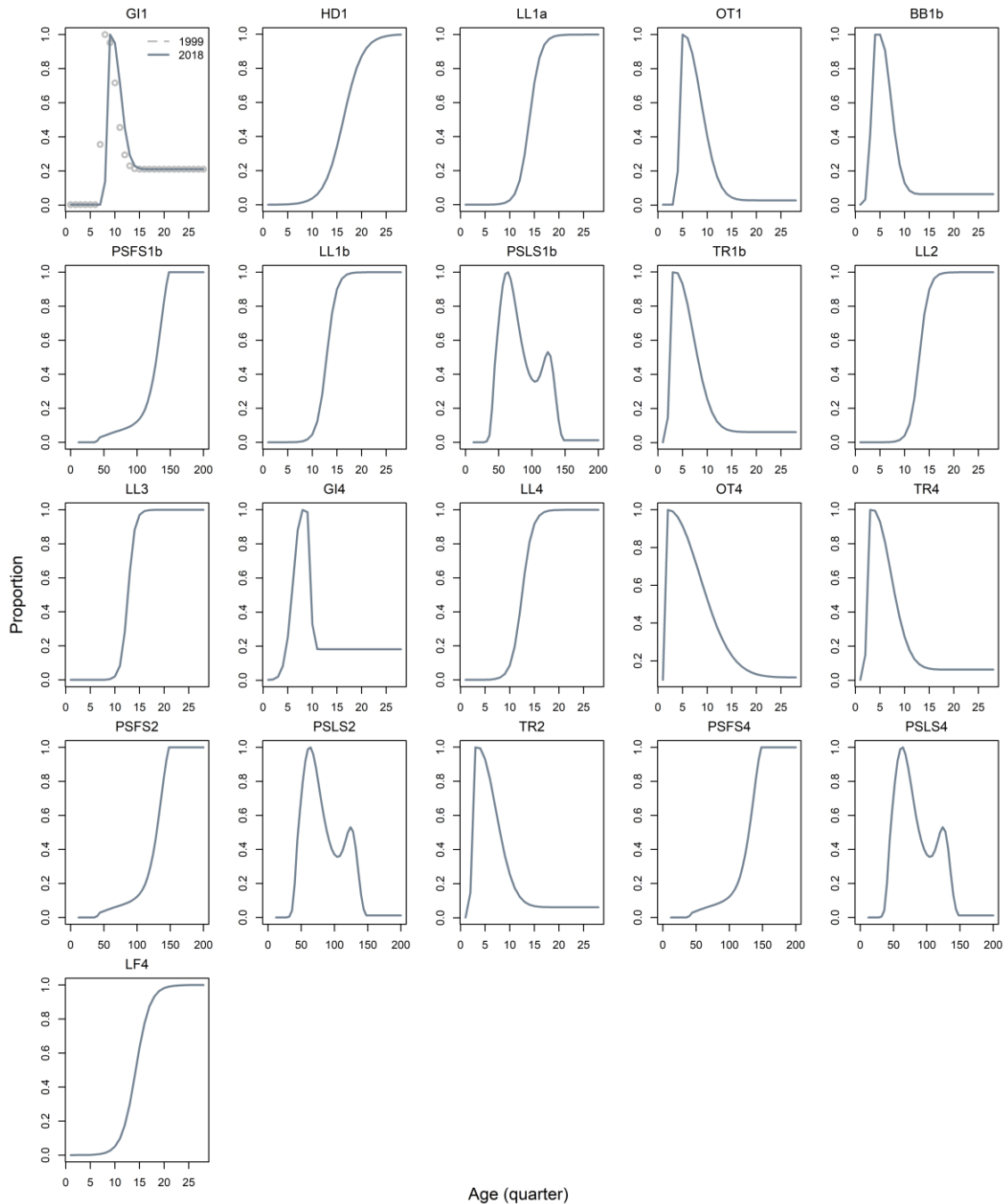


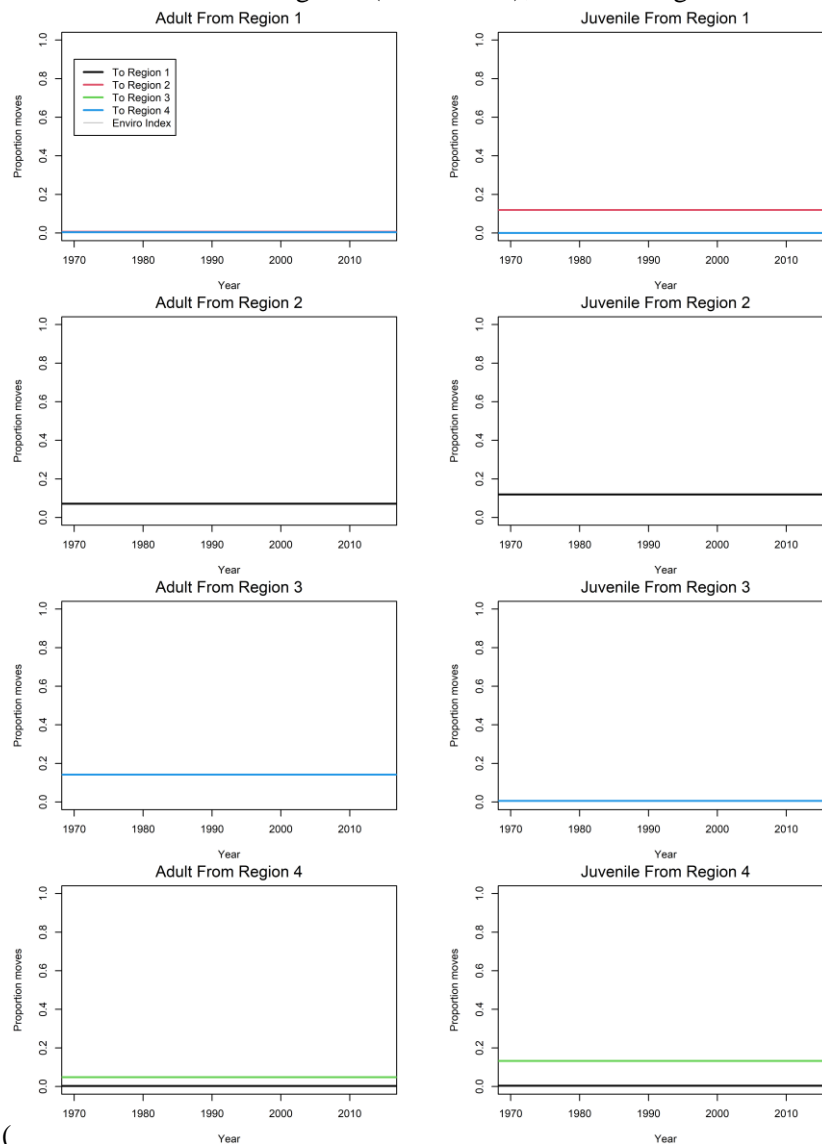
Figure 25. Independent logistic selectivity functions are estimated for the principal longline fisheries (LL 1a, 1b, 2–4) which attain full selectivity at about age 15–17 quarters. The fresh tuna fishery (LF 4) is estimated to have a relatively similar selectivity to the principal longline fisheries, albeit slightly

skewed towards older fish. The logistic selectivity of the handline (HD1a) fishery is estimated to have a relatively gentle slope (the fishery caught a broader range of age classes). The associated purse-seine fisheries have a high selectivity for small fish, while the free-school purse-seine fishery selects substantially larger fish. For all regions, the selectivities of the free school and associated purse-seine fisheries was held constant through time. The selectivity of associated purse-seine sets is relatively broad compared to the modal structure of the length frequency data. The pole and line fishery is also highly selective for juvenile fish. Limited or no size data were available for a number of fisheries, specifically the artisanal fisheries (OT 1a & 4) and the troll fishery in regions 1b and 2 (TR 1b & 2). Consequently, selectivity for these fisheries is poorly estimated or, in the absence of size data, assumed equivalent to a fishery with the same gear code in another region. The model did not estimate a significant change of selectivity for gillnet fishery in region 1a, despite the fishery appeared to have caught smaller fish after the 2000s than the early period.

The quarterly recruitment deviates indicate that recruitment varies seasonally (Figure 26). Recruitment deviates were low during 2004–2006, especially during 2005. This low recruitment occurred shortly before the tagging program and may be related to the intention of the model to achieve the estimation of a lower biomass (and a higher fishing mortality) to better predict the tag returns. The low recruitment estimate may also be due to the subsequent decline in CPUE rates in the 2007–2011 piracy period. However, the pattern persists in models where either the tagging data or LL 1b CPUE indices (after 2007) were removed, but disappeared when both datasets were removed (see Figure C22 of Fu et al. 2018), suggesting that the estimation of low recruitment in 2004–2006 was likely related to both factors.

Recruitment is parameterised to occur in region 1 and 4 only. The model estimates about 60% and 40% of the total annual recruitment is assigned to regions 1 and 4 in the initial period 1950–1977, respectively. The proportion of total recruitment assigned to either region varies temporally during the estimation period (1977–2018) and the proportion allocated to region 1 has increased to be above 80% since mid-2000s, (and vice versa for region 4) (Figure 27). The large increase of the recruitment to region 1 coincided with the exceptionally high catches that occurred in the western tropical region between 2003–2006. In a hypothetical model which assumed that the sharp increase in catches in region 1 occurred in region 4 instead, the regional recruitment trend is reversed.

The model estimates that there is a relatively low degree of connectivity between the two western regions (R1 and R2) and between the eastern regions (R3 and R4), and no longitudinal movement



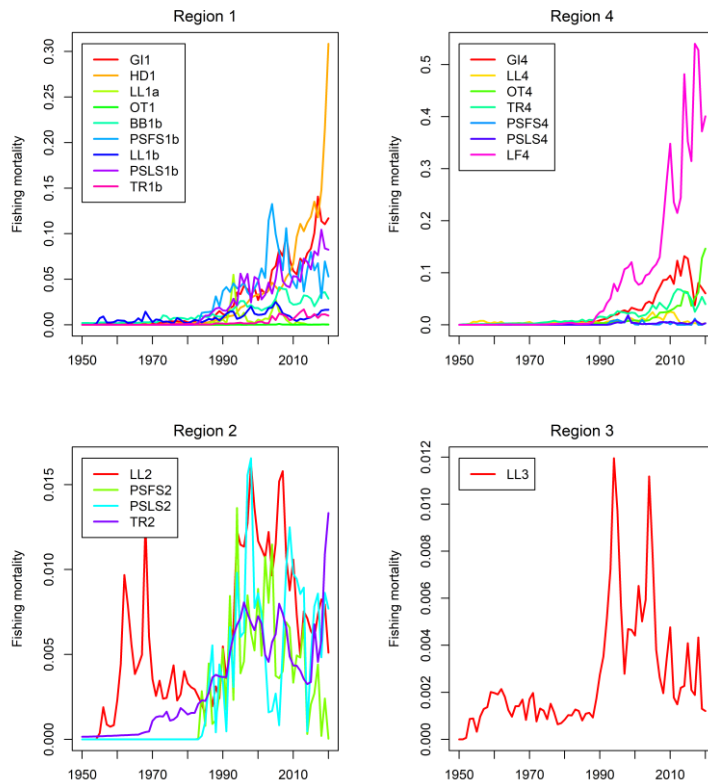
between regions 1 and 4 (

Figure 28). This contrasts with the estimates from the early assessment which indicated that the movement between R1 and R2, and between R3 and R4 is relatively high, especially for the juveniles (Fu et al. 2018). A likelihood profile on movement parameters (e.g., R1 to R2) suggests some conflicts between datasets with CPUE indices supporting low movement rates while tagging data supporting higher rates.

The estimated initial spawning biomass (SB0) is higher than the reference model in the previous assessment, but the spawning biomass levels in recent years are similar (Figure 29). The difference mainly stems from the update to the longline CPUE indices (as illustrated by the same basic model which included the 2018 CPUE indices instead, see Figure 29). It is estimated that in the 1950s, 1960s, and early 1970s, the spawning biomass of the IO population was still relatively high, reflecting the relatively low catch and the assumption of equilibrium recruitment during this period. Total spawning biomass declined rapidly from the late 1980s to the mid-1990s, recovered slightly in the late 1990s and early 2000s, and then fell to low levels in 2008–2009. The spawning biomass rebounded slightly from 2009 to 2011 and then declined to the current year with fluctuations, with the SB in 2020 estimated to

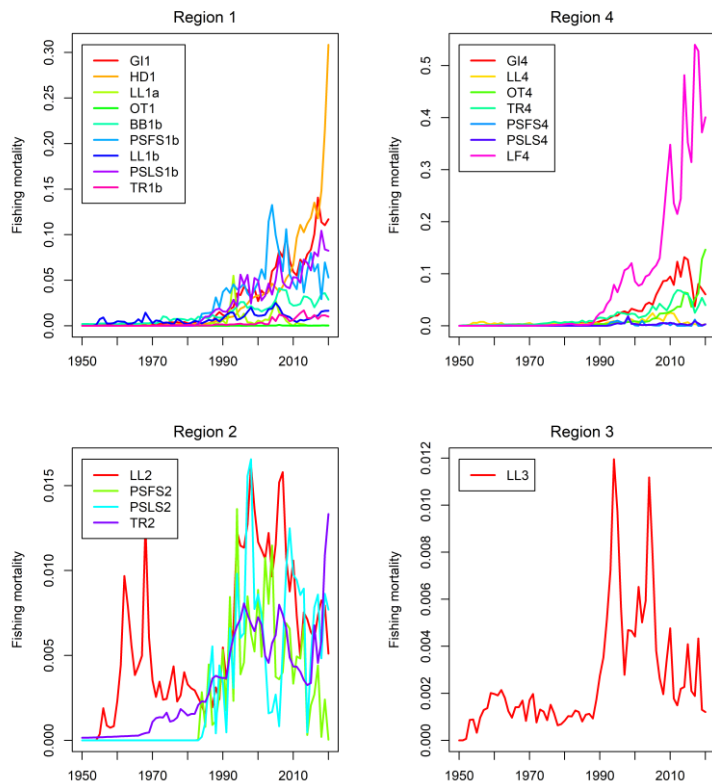
be close to the historically low level. The relative trends of the four model regions are largely comparable (Figure 30), although the overall magnitude of decline is substantially higher in Region 1 and 4. The biomass in region 4 declined steadily throughout the 1990s and 2000s following the trend in the LL CPUE index. For the most recent years, region 4 biomass is estimated to be at a very low level (Figure 30).

Fishing mortality rates for each fishery are defined as apical fishing mortality rates, i.e., the fishing mortality for the fully selected age class (or age classes). The fishing mortality rates approximate the Baranov continuous F (Methot & Wetzel 2013). Relatively high recent fishing mortality rates have been estimated for a number of fisheries in Region 1, specifically PSL1b, PSFS1b, GI1a, HD1a and BB1b



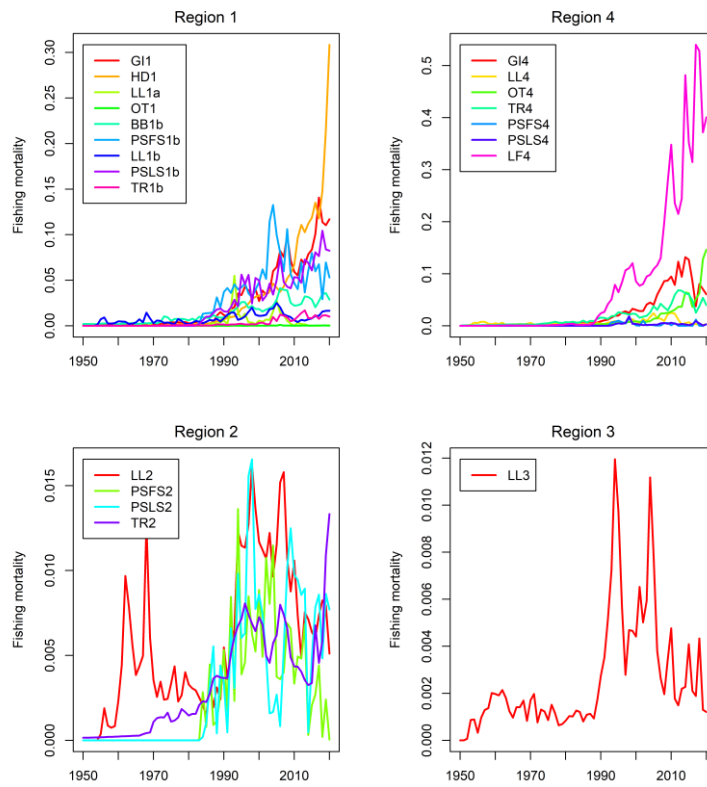
(Figure 31). Fishing mortality rates for the HD1a fishery increased sharply since 2010 corresponding to relatively high catches from the fishery (particularly in 2020). Estimates of fishing mortality for the PSL1b fishery appeared to be low by comparison, considering that this fishery mostly catches juvenile fish (in contrast to the HD1a fishery which catches mostly adults). Estimation of the size selection of the PSL1b fishery may have been biased towards larger fish by the inclusion of the tagging data which show some conflicts to purse seine size composition data (see Figure B11). In Region 4, recent fishing

mortality rates from the LF4 fishery were high



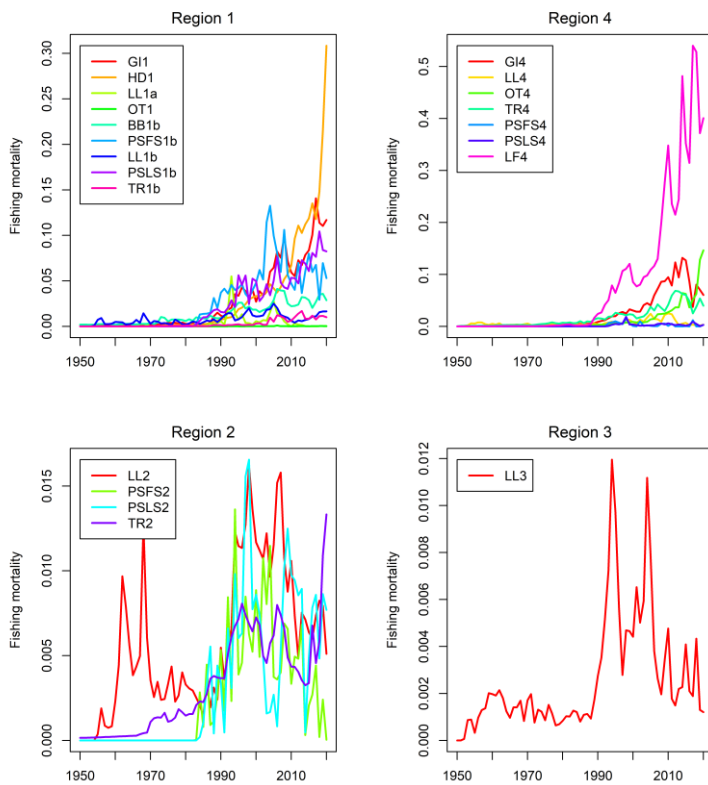
(Figure 31), although there remains great uncertainty in annual catches from the fishery during the last 10 years (Geehan & Setuadji 2018). The high fishing mortality rates correspond to the sharp decline in model biomass from the late 2000s and are also related to the selectivity of the fishery, with full

selection occurring at age 18 quarters. The GI4 and the TR4 fisheries represent the other main sources



of fishing mortality in Region 4 (

Figure 31). Fishing mortality rates are estimated to be very low in both Region 2 and Region 3



(

Figure 31).

Spatially aggregated, age-specific fishing mortality rates are derived for each model time period (Methot & Wetzel 2013). Average total fishing mortality rates were derived for the last two years of the assessment model (2019 and 2020) and the resulting age specific mortality schedule was applied in the computation of the *MSY* reference points. Aggregated fishing mortality rates were highest for age classes 18-24 (Figure 32).

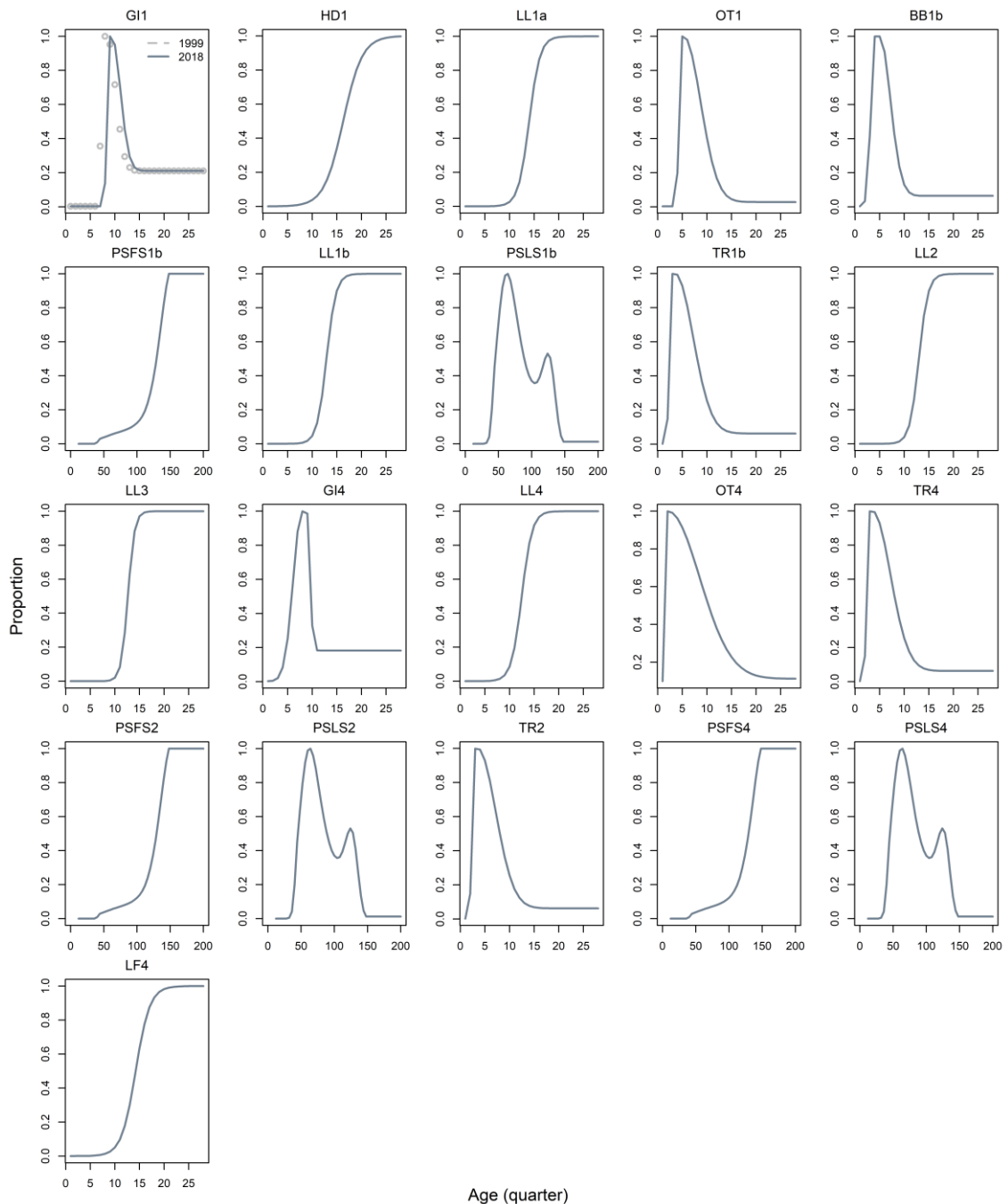


Figure 25: Age specific selectivity by fishery from the basic model (length-based selectivity for purse seine fisheries).

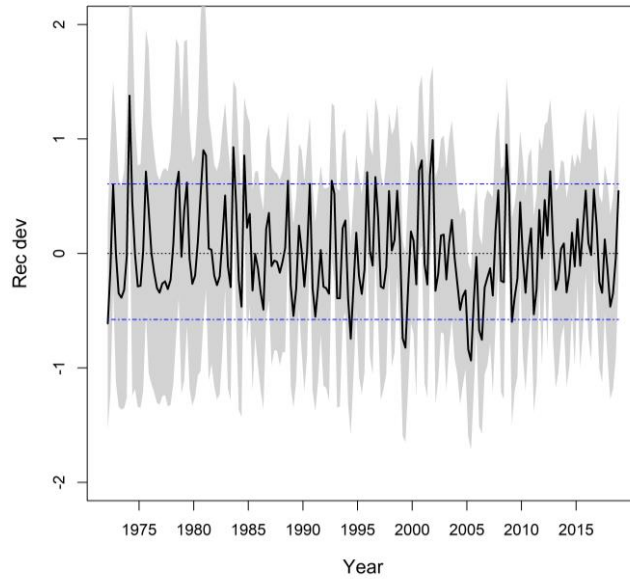


Figure 26: Recruitment deviates from the SRR with 95% confidence interval from the basic model

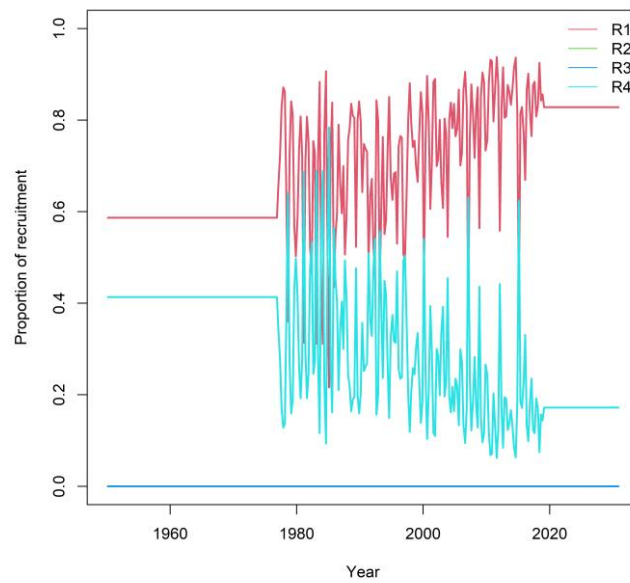


Figure 27: Proportion of the total quarterly recruitment assigned to region 1 (red) and region 4 (blue).

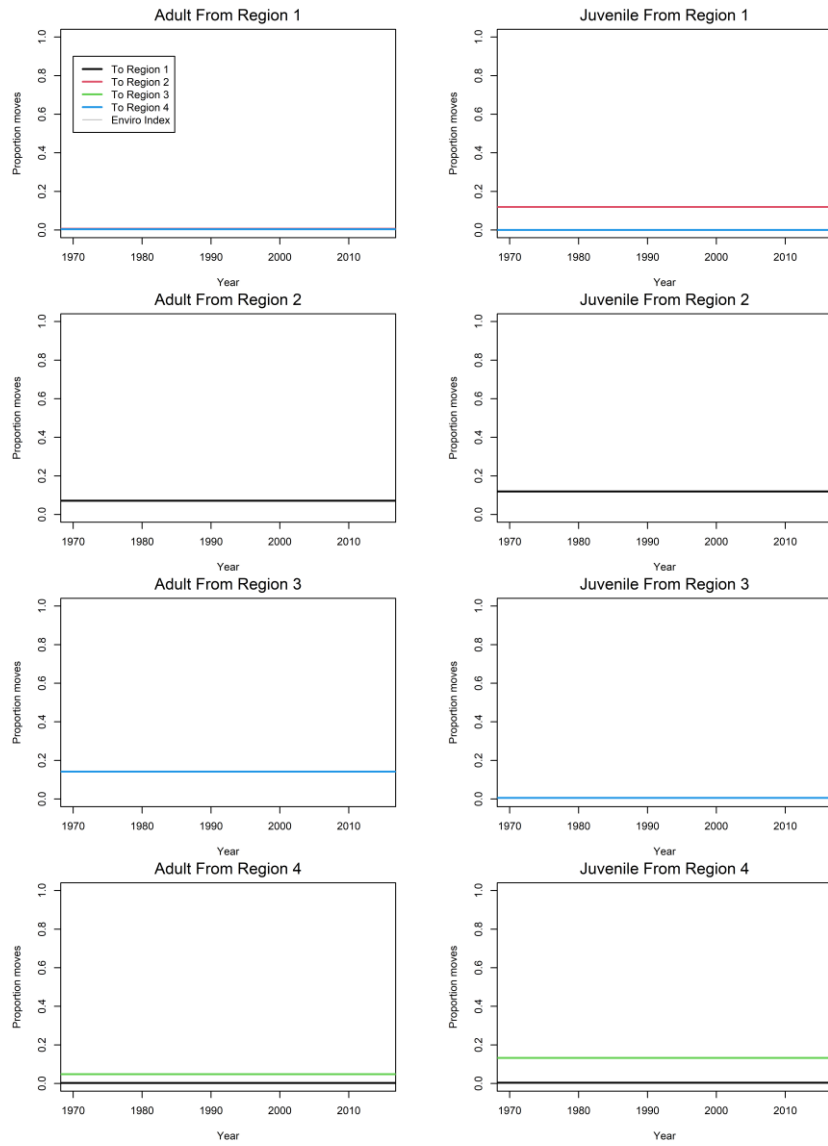


Figure 28: Estimated age specific movement parameters for the basic model.

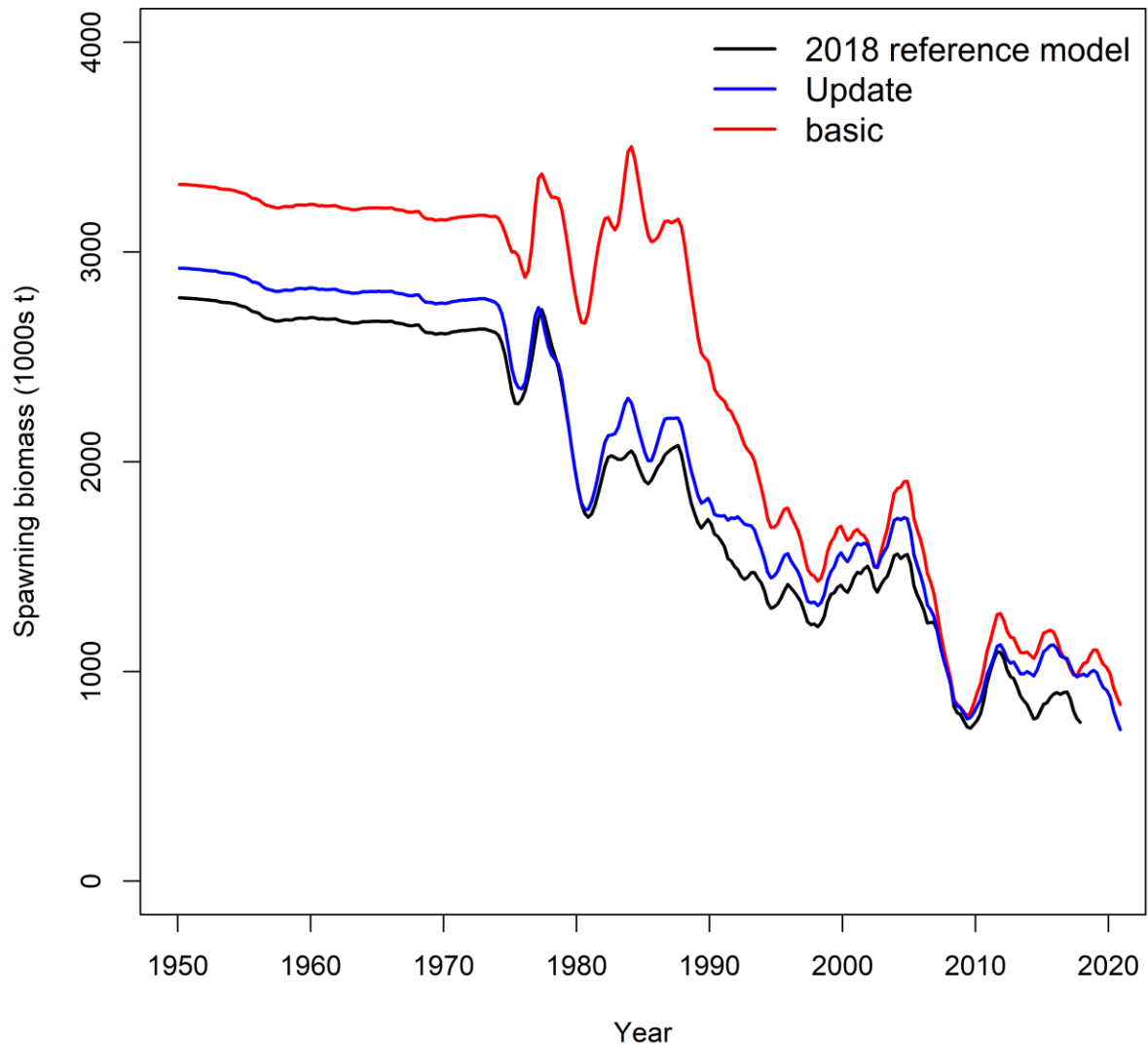


Figure 29: A comparison of estimated spawning biomass from the basic model and the reference model from the 2018 assessment, and model “update” which is the same as the basic model except it used CPUE indices from the 2018 assessment.

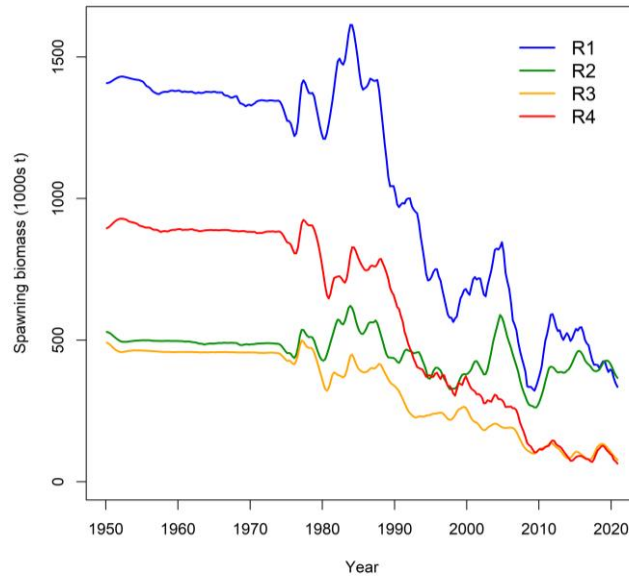


Figure 30: Estimated spawning biomass trajectories for the individual model regions from the basic model.

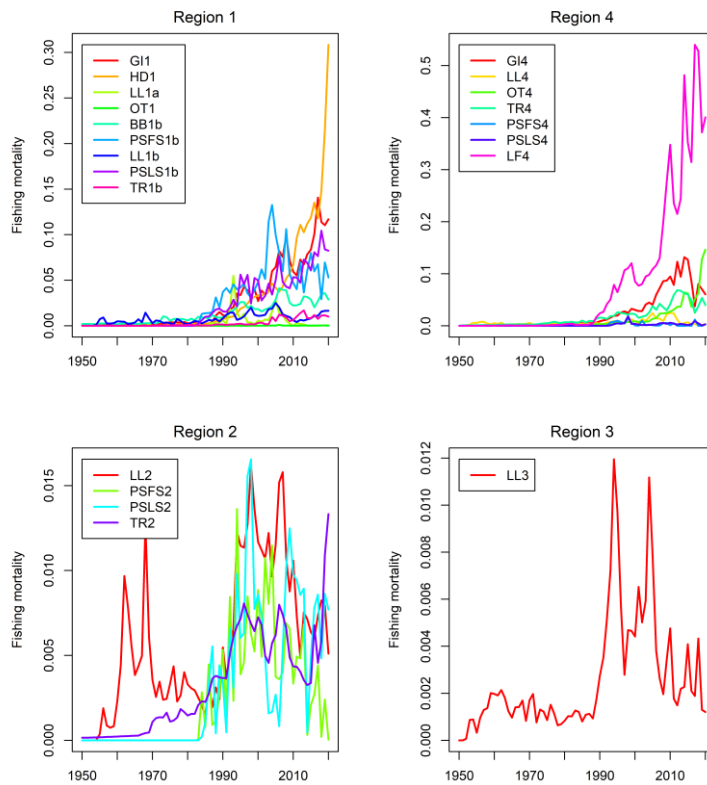


Figure 31: Trends in fishing mortality (quarterly) by fleet.

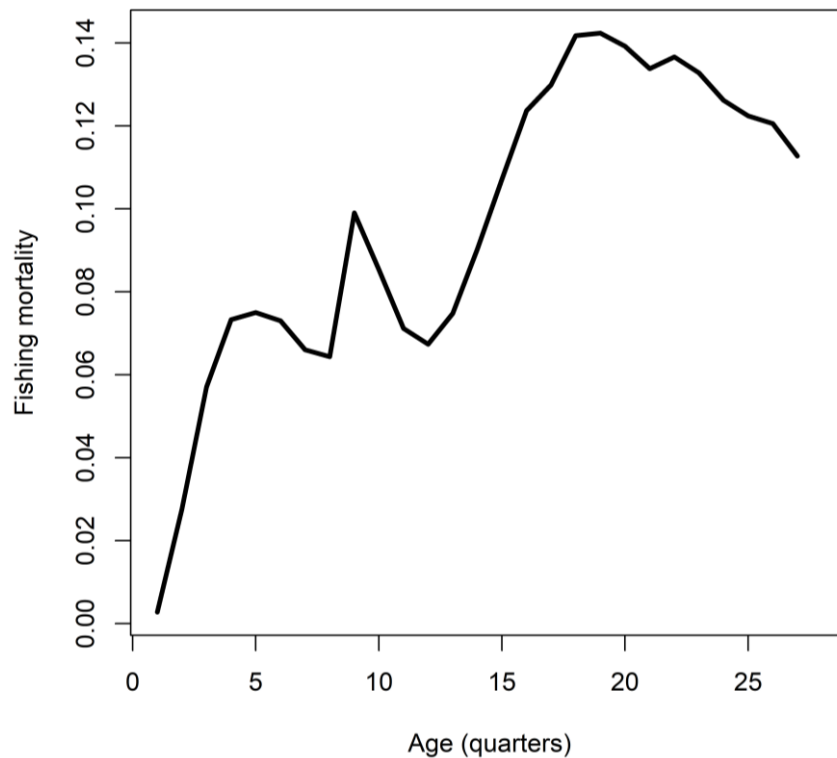


Figure 32. Fishing mortality (quarterly, average) by age class for the period used to determine the total F-at-age included in the calculation of MSY based reference points (2019 and 2020) for the basic model.

5.2 Exploratory analysis

The exploratory model examines alternatives to relevant parameters and structural assumptions and aims to determine possible improvement to the assessment model (see Table 4). The selected diagnoses of these models are listed in Appendix B and discussed briefly below.

Spatial structure

The revised model adopted the alternative 4-region spatial structure (see Figure 2 of Section 2.1), and the associated fishery definition (Table 1, Appendix A). The alternative regional structure better accommodates the distribution of the EU purse seine fishery and is consistent with the spatial coverage of the purse seine CPUE indices. The simplification of fleet definition helps to divide the purse seine fishery (both free and associated school) into the small (≤ 80 cm) and large (> 80 cm) mode, which makes it possible to include the purse seine CPUE, and better account for the variable composition of juvenile and adult fish in the length frequency time series of the purse seine fishery. Further, the modified regional structure contains the main range of the tag dispersion, thereby minimising the influence of the tagging observations on the movement estimation (the tagging design is shown not suitable to inform the movement of yellowfin tuna on the regional scale). For comparison, the revised model retained only the longline CPUE, and the purse seine CPUE were included in additional models. The main limitation is that the regional LL CPUE was derived for the original spatial structure and so is the regional scaling factor. However, this bias is unlikely to be large because the changes in the revised regional boundaries are relatively small.

Another revision includes excluding from the model tag recoveries from fisheries other than the purse seine fisheries. The reason for this is that recoveries from the purse seine fisheries accounted over 94% of observations, which is sufficient for the estimation of regional abundance and fishing mortality. On the other hand, a constant reporting rate must be estimated for each of the other fisheries when the reporting rate is more likely to have changed over time, which a source of bias. The reporting rates for all other fisheries are fixed to zero in the revised model.

The revised model estimates that stock productivity is higher, and SB0 is estimated to be about 20% higher than the basic model (Table B1, Figure B1). A more thorough diagnosis shows that the model performance has not significantly improved or deteriorated – overall the fits to the CPUE indices and tagging data (PS only) are very similar to the basic model. By estimating separate yet constant selectivity for the two length modes (Figure B2), the residuals from the fits to the purse seine length compositions are improved considerably as the length distributions under the two length modes are modelled independently (Figure B3). By dividing the annual/quarter catches into the small and large length components on basis of the length distribution, the model implicitly assumed non-stationary selectivity, as reflected in the variable catch compositions of large vs small length modes. The model estimated even lower movement rates between regions as the tag distribution is limited to (revised) region 1 therefore it is not expected to influence the estimation of movement. However, the estimation of movement still appears to be sensitive to starting parameter values.

On another front, it is suggested it is not necessary to use the hypothetical spawning locations to define where recruitment occurs in the model. As such, *bRecruit4* let the model determine the distribution of fish when they enter the fishery rather than limit them to the 2 tropical regions. The model estimated an increasing recruitment trend in R2 (similar to R1), and a decreasing trend for R3 (similar to R4) (Figure B4). By given additional freedom in distributing the recruitment, the model fits to both CPUE and length compositions were improved with noticeable reduction in (negative) likelihood values (Table B2). As the juvenile fish were distributed through recruitment, the model estimated very low movement rates (Figure B4). Overall, the SB0 estimated by the model is about 8% higher than the basic model.

Seasonal effects

The longline CPUE and length composition showed strong seasonality (see Figure 9) which was not explicitly considered in the assessment despite the seasonal resolution of the model. For example, if the catchability remains constant between seasons, the model interprets seasonal fluctuations in CPUE as changes in abundance. However, for tropical tuna, the seasonal changes in the longline CPUE and length distribution often reflect changes in the availability and vulnerability of the fish population to the fishing gear due to changes in environmental conditions. Yet the seasonal movement, catchability and selectivity may be confounded, which is a potential source of bias. Two sets of models were thus implemented to assess the influence of seasonality on model estimates. In the first model (*bSeasonX*), only longline CPUE and length composition from one quarter in each region were retained in the model. This assumes that the data from one season is sufficient to show long-term population trend and fishing selectivity. The model was repeated for each season (ideally, it is best to focus on the season with the largest spatial coverage). In the second model (*bSeason*), all seasons were included, but each regional longline fishery is divided into four seasonal fisheries with independent selectivity, and the regional longline CPUE is divided into four seasonal time series with a separate catchability (altogether the model has 16 LL fisheries and 16 LL CPUE indices).

Model *bSeason* better reflects the seasonal changes of catchability in the CPUE index (Figure B5). The model also shows seasonal differences in longline fishing selectivity: it seems that smaller fish were caught in the second quarter and larger fish were caught in the fourth quarter in both tropical regions, while the situation in the western temperate region is the opposite (Figure B5). However, despite the complexity of the configuration, the improvement in the fitting of the length composition data appears to be very small (Table B2), and the biomass estimates are very similar to the basic model (Table B1).

The models containing only one season's longline fishing data also estimated similar population biomass, but the model based on the first quarter data is more optimistic (Figure B1), because the LL 1b CPUE index in the first quarter tended to be flatter than the other quarters over the past ten years.

Length composition data

Several model runs were conducted to explain the changes in the longline length compositions in the early 2000s as the data were increasingly dominated by CPCs other than the Japan (see Section 2.6). These models make different assumptions about the cause of this change. Model *bLFDw* assumes that the quality of the recent length composition samples was not as good as the early years and therefore downweighted the length data after 2003. However, this model did not improve the fits to the early length composition as expected (in terms of residual patterns). Model *bSelrw* assumes the longline selectivity has changed since 2003 and estimated a random walk on the longline selectivity parameters. The model accounted for the large variations in the average fish size in recent longline catches better and showed the pattern in the residuals from the fits to the recent length compositions has improved (e.g., LL 4, see Figure B6). However, allowing the random variations on the selectivity parameters seemed to have introduced an artefact of noise into the predicted CPUE (Figure B6), although the magnitude of the noise can be reduced by assuming a smaller standard deviation of the random walk. It seems that a more practical solution to the increase in average fish size after 2003 is to separate the longline fishery before and after 2003. Model *bLFSplit* indicated that selectivity for LL 1b and 4 fisheries before 2003 are relatively dome shaped (Figure B7). The LL 2 fishery also showed a shift of selection towards larger fish in recent years (Figure B7). Model *bLFSplit* showed a slight improvement in the residuals of the fits to length compositions (e.g., LL1b, Figure B7).

CPUE indices

Model *bLL1bSplitQ* considered a scenario where the catchability for the LL1b CPUE before and after the piracy period have changed and were estimated independently in the model. The model estimated a 42% decrease in LL1b catchability for the period 2012 – 2020 compared to the early period (prior to 2007) assuming the decline in the index after the piracy period reflects the change in catchability rather than abundance. Accordingly, the model estimated a higher recent biomass level than the basic model which assumed a constant catchability (Figure B1).

Several model runs were conducted to examine the inclusion of purse seine CPUE and other abundance indices. The models are based on the revised region structure in which the purse-seine fisheries are divided into the small (≤ 80 cm) and large (> 80 cm) fish components. We focused more on the adult PSFS index, which is more comparable to the longline fishing index. Two approaches were considered. Model *rCPUEFSq* tries to match the PSFS index with the longline index in the same region. Firstly, the lambda of the PSFS index was set to zero so that it is uninformative, and the departure of index from the expected population trend was assumed to reflect the change in catchability (Kolody 2018), which was estimated to be around 1.315% per annum (Figure B8). Secondly, the PSFS index adjusted for the

catchability change was included in the model with full weight. On the other hand, Model *rCPUEFS_{pe}* included the PSFS index as it is, but estimated additional variance components for the indices. This model estimated that an additional standard error of 0.15 and 0.35 must be added to the LL1b and PSFS index, to offset their differences within the model. Both models estimated a recent biomass trend that is similar to the revised model (Figure B1).

Model *rCPUEFSLSBB_{pe}* included all available indices and estimated the process error of each index.

The PSFS Juvenile, PSFS adult, ABBI, and BB index require an additional standard error of 0.57, 0.30, 0.14 and 0.22, respectively. These indices were generally fitted well in the model except for the BB index showing an overall upward trend from 1990 to 2020 (Figure B9).

Biological parameters

Model *bDortel* produced much lower biomass estimates and a more pessimistic stock status (Figure B1). As the change in mean size of a fished population relative to the unfished state was usually interpreted by the stock assessment model as fishery-induced depletion, the lack of large fish in the catch, relative to the higher asymptotic length in the model would imply a higher level of fishing mortality, hence a large fishery depletion effect (McKechnie et al 2017). But models had a poorer fit the overall length composition compared to the basic model (Figure B10, Table B2).

The natural mortality assumed in model *bMlow* is about 60% of the base level and at one time was considered a credible lower bound for the range of values of natural mortality (Langley 2016), although more recent evidence gives support to lower values (Hoyle 2021). Models with lower values of natural mortality produced a considerably more pessimistic stock status than the base level natural mortality (Figure B1).

Model *Mlow* had a better fit to the overall tag returns, mainly for the older age classes (Figure B10). Model *Mmax18* predicted a much higher tag returns for the older age classes (Figure B10), suggesting the maximum of 18 is not consistent with the IO tagging data. However, this may be because the purse seine fishery is primarily targeting juvenile fish and the reporting rates for other fisheries (e.g., LL) are very low. There are also some conflicts between the tag recover and length composition data from the purse seine associated sets (Figure B11). The profile likelihood on natural mortality showed that the tag returns support a lower natural mortality, whereas the length composition data support a higher natural mortality (Figure B12). Both observations are relatively indifferent to natural mortality less than 70% of the base level. Overall, the profile likelihood indicated that the range of natural mortality consistent with the data and the structure of the base model is about 70% to 100% of the base value (Figure B12).

5.3 Final model options

The revised model provides an alternative yet plausible spatial representation of the IO yellowfin population. The revised model setting has yielded different estimates of stock productivity to the basic model. The spatial structure contributes to a source of uncertainty and therefore is included in the final model options.

A key conclusion from the previous assessment is that the tagging data conflicts with the CPUE and the model estimates (particular the population scaling parameter R_0) are sensitive to the weighting of the tagging data. Thus, the final model options included two different weighting assumptions on the tagging data set. The weightings were applied by the values assigned to the proportion (λ) of the two components of the tag likelihood included in the total model likelihood.

The final model options also included three alternative values of steepness of the BH SRR (h 0.7, 0.8 and 0.9). These values are considered to encompass the plausible range of steepness values for tropical tuna.

The effect of piracy (2007–2011) on the CPUE in region 1b has been extensively discussed in the WPTT meetings. Piracy severely restricted the operation of the longline fishing fleet, resulting in severe reduction (about 60%) in the number of vessels in the Indian Ocean, leading to lower coverage and potential biases in the CPUE estimates for this time period. It was suggested that the decrease in CPUE during these years may be due to changes in catchability from the spatial contraction of the fleet rather than changes in abundance. As such the 2018 assessment final models considered a scenario where the catchability for the LL1b CPUE before and after the piracy period have changed and were estimated independently in the model. The catchability option is therefore retained in the final model options.

Growth and natural mortality are also important sources of uncertainty. The empirical estimates of Fonteneau (2008) used in the basic model has a low Linf parameter value compared to many regional studies in the IO. Estimates of Dortel et al. (2014) integrates tag and size data and has a higher Linf value. The inclusion of both growth options can address the uncertainty in the estimation of growth-related stock depletion to a certain extent. The natural mortality rate assumed in Model *Mlow* model is based on aging studies in the Atlantic Ocean. Studies in the Pacific and Indian Oceans also indicate that the life span of yellowfin tuna is much longer than previously thought. However, the *Mlow* options tend to produce an estimate of stock productivity that may be too pessimistic (models indicate that the population will collapse very quickly under even low catches). Instead, the final model options included a natural mortality level set at 70% of the base value, as suggested by the profile likelihood analysis (see Section 5.2).

The final model ensemble corresponds to a full factorial combination of two spatial structures, three steepness values, two levels of tag weighting, and two growth option, and two options on the LL CPUE catchability assumption, with a total of 96 models (see Table 5 **Error! Reference source not found.**). The model ensemble includes a wide range of stock trajectories, and the initial spawning biomass (SB_0) ranges from 2 850 560 to 6 045 250, current depletion (SB_{2020}/SB_0) ranges from 19% to 41% (across the model grid (Table 6 **Error! Reference source not found.**). In general, higher stock biomass estimates are associated with the revised spatial structure, low weighting of tagging data, and a lower steepness value. We also note that the combination of low natural mortality and the Dortel et al. (2008) growth options tend to lead to more pessimistic stock estimates and struggle to support short-term catches (at the current level).

Table 6: Estimates of key reference quantities for each model from the final model ensemble .

Option	SB_0	SB_{MYS}	SB_{2020}	$\frac{SB_{2020}}{SB_{MYS}}$	$\frac{F_{2020}}{F_{MYS}}$	MSY
io_h70_q1_Gbase_Mbase_tlambda01	4363160	1703440	1299745	0.76	1.00	451396
io_h70_q1_Gbase_Mbase_tlambda1	3481140	1339050	914318	0.68	1.32	412840
io_h70_q1_Gbase_Mlow_tlambda01	6005010	2286720	1530335	0.67	1.49	394163
io_h70_q1_Gbase_Mlow_tlambda1	4832400	1816950	1060831	0.58	1.93	366388
io_h70_q1_GDortel_Mbase_tlambda01	3798520	1488860	969479	0.65	1.49	407280
io_h70_q1_GDortel_Mbase_tlambda1	3267920	1240520	761559	0.61	1.77	381730
io_h70_q2_Gbase_Mbase_tlambda01	4634820	1712140	1826568	1.07	0.67	446140
io_h70_q2_Gbase_Mbase_tlambda1	3486140	1306180	1143190	0.88	1.01	383648
io_h70_q2_Gbase_Mlow_tlambda01	6045250	2243200	1959890	0.87	1.14	364626
io_h70_q2_Gbase_Mlow_tlambda1	4717010	1755970	1272080	0.72	1.53	336935

io_h70_q2_GDortel_Mbase_tlambda01	3594880	1381190	1138930	0.82	1.23	364247
io_h70_q2_GDortel_Mbase_tlambda1	3144040	1191690	943759	0.79	1.37	343994
io_h80_q1_Gbase_Mbase_tlambda01	4170460	1561460	1284090	0.82	0.87	469428
io_h80_q1_Gbase_Mbase_tlambda1	3219450	1181190	880850	0.75	1.27	416364
io_h80_q1_Gbase_Mlow_tlambda01	5613250	2046380	1477320	0.72	1.35	402144
io_h80_q1_Gbase_Mlow_tlambda1	4523020	1631980	1032658	0.63	1.71	374498
io_h80_q1_GDortel_Mbase_tlambda01	3552240	1344140	944600	0.70	1.36	409436
io_h80_q1_GDortel_Mbase_tlambda1	3087420	1129500	750656	0.66	1.57	392292
io_h80_q2_Gbase_Mbase_tlambda01	4444080	1530210	1800423	1.18	0.57	468688
io_h80_q2_Gbase_Mbase_tlambda1	3318860	1176080	1129948	0.96	0.86	399402
io_h80_q2_Gbase_Mlow_tlambda01	5617900	1970970	1817638	0.92	1.00	380511
io_h80_q2_Gbase_Mlow_tlambda1	4414170	1558890	1239358	0.80	1.35	344149
io_h80_q2_GDortel_Mbase_tlambda01	3397470	1231660	1112808	0.90	1.06	371296
io_h80_q2_GDortel_Mbase_tlambda1	2974490	1066350	913637	0.86	1.22	354717
io_h90_q1_Gbase_Mbase_tlambda01	4033360	1444090	1273695	0.88	0.77	485208
io_h90_q1_Gbase_Mbase_tlambda1	3208800	1136430	902003	0.79	1.02	439400
io_h90_q1_Gbase_Mlow_tlambda01	5383340	1889040	1440423	0.76	1.17	419776
io_h90_q1_Gbase_Mlow_tlambda1	4244730	1473480	988775	0.67	1.52	387812
io_h90_q1_GDortel_Mbase_tlambda01	3414450	1250350	932027	0.75	1.19	421584
io_h90_q1_GDortel_Mbase_tlambda1	2947180	1038070	737899	0.71	1.39	402956
io_h90_q2_Gbase_Mbase_tlambda01	4313280	1367400	1786643	1.31	0.49	490856
io_h90_q2_Gbase_Mbase_tlambda1	3200560	1066200	1120388	1.05	0.75	414468
io_h90_q2_Gbase_Mlow_tlambda01	5330670	1708010	1765985	1.03	0.87	382600
io_h90_q2_Gbase_Mlow_tlambda1	4186690	1395300	1222353	0.88	1.20	351183
io_h90_q2_GDortel_Mbase_tlambda01	3288400	1125590	1106230	0.98	0.93	386362
io_h90_q2_GDortel_Mbase_tlambda1	2850560	970082	896347	0.92	1.07	365828
io_h70_q1_GDortel_Mlow_tlambda01	4695270	1763230	945056	0.54	2.39	351642
io_h70_q1_GDortel_Mlow_tlambda1	4151370	1533410	788338	0.51	2.82	322080
io_h70_q2_GDortel_Mlow_tlambda01	4418290	1653260	1121543	0.68	1.99	304082
io_h70_q2_GDortel_Mlow_tlambda1	3837430	1422630	902864	0.63	2.21	293483
io_h80_q1_GDortel_Mlow_tlambda01	4390210	1575460	898862	0.57	2.05	364273
io_h80_q1_GDortel_Mlow_tlambda1	3906390	1383330	752815	0.54	2.35	348073
io_h80_q2_GDortel_Mlow_tlambda01	4097950	1386670	944522	0.68	1.78	349571
io_h80_q2_GDortel_Mlow_tlambda1	3530930	1239870	831057	0.67	1.95	300047
io_h90_q1_GDortel_Mlow_tlambda01	4010030	1377950	815388	0.59	1.92	362444
io_h90_q1_GDortel_Mlow_tlambda1	3695780	1249120	733450	0.59	2.11	351396
io_h90_q2_GDortel_Mlow_tlambda01	3808540	1281060	963343	0.75	1.61	321126
io_h90_q2_GDortel_Mlow_tlambda1	3348050	1111930	814531	0.73	1.73	304943
sp_h70_q1_Gbase_Mbase_tlambda01	4637610	1836860	1450790	0.79	0.78	483292
sp_h70_q1_Gbase_Mbase_tlambda1	4050640	1609290	1125265	0.70	1.10	444392
sp_h70_q1_Gbase_Mlow_tlambda01	5887740	2202970	1470110	0.67	1.26	391463
sp_h70_q1_Gbase_Mlow_tlambda1	5180370	1972710	1188475	0.60	1.59	378549
sp_h70_q1_GDortel_Mbase_tlambda01	3869190	1517680	999718	0.66	1.30	414476
sp_h70_q1_GDortel_Mbase_tlambda1	3932690	1565800	1032926	0.66	1.28	425796
sp_h70_q2_Gbase_Mbase_tlambda01	4700830	1737940	1794285	1.03	0.59	471736

sp_h70_q2_Gbase_Mbase_tlambda1	3882700	1501760	1229098	0.82	0.78	422080
sp_h70_q2_Gbase_Mlow_tlambda01	5674260	2076170	1722608	0.83	1.01	363367
sp_h70_q2_Gbase_Mlow_tlambda1	5117870	1912200	1394578	0.73	1.26	360968
sp_h70_q2_GDortel_Mbase_tlambda01	3618110	1270170	1179873	0.93	0.96	348938
sp_h70_q2_GDortel_Mbase_tlambda1	3896880	1542010	1137758	0.74	1.22	415440
sp_h80_q1_Gbase_Mbase_tlambda01	4468570	1711730	1448773	0.85	0.68	504644
sp_h80_q1_Gbase_Mbase_tlambda1	3891480	1504120	1130868	0.75	0.91	469904
sp_h80_q1_Gbase_Mlow_tlambda01	5306750	1882200	1336463	0.71	1.13	394786
sp_h80_q1_Gbase_Mlow_tlambda1	4881470	1787970	1158388	0.65	1.37	392666
sp_h80_q1_GDortel_Mbase_tlambda01	3768480	1413290	1011780	0.72	1.08	439776
sp_h80_q1_GDortel_Mbase_tlambda1	3735950	1448270	1017848	0.70	1.06	436652
sp_h80_q2_Gbase_Mbase_tlambda01	4471130	1480830	1749800	1.18	0.49	489040
sp_h80_q2_Gbase_Mbase_tlambda1	3920830	1474410	1373548	0.93	0.71	455672
sp_h80_q2_Gbase_Mlow_tlambda01	5464380	1849740	1702325	0.92	0.85	383675
sp_h80_q2_Gbase_Mlow_tlambda1	4805190	1720530	1363963	0.79	1.10	368491
sp_h80_q2_GDortel_Mbase_tlambda01	3427240	1170090	1170708	1.00	0.85	371075
sp_h80_q2_GDortel_Mbase_tlambda1	3709540	1409350	1113923	0.79	0.97	433728
sp_h90_q1_Gbase_Mbase_tlambda01	4361770	1618650	1446563	0.89	0.59	525680
sp_h90_q1_Gbase_Mbase_tlambda1	3777690	1419250	1131660	0.80	0.79	490332
sp_h90_q1_Gbase_Mlow_tlambda01	5246050	1835150	1354660	0.74	0.99	426188
sp_h90_q1_Gbase_Mlow_tlambda1	4649500	1621480	1134768	0.70	1.19	395802
sp_h90_q1_GDortel_Mbase_tlambda01	3531800	1292880	978109	0.76	1.00	438620
sp_h90_q1_GDortel_Mbase_tlambda1	3601110	1356420	1009112	0.74	0.99	456492
sp_h90_q2_Gbase_Mbase_tlambda01	3627870	1341140	1138883	0.85	0.75	481344
sp_h90_q2_Gbase_Mbase_tlambda1	3789630	1366000	1369513	1.00	0.63	469108
sp_h90_q2_Gbase_Mlow_tlambda01	5222210	1694870	1676420	0.99	0.76	397671
sp_h90_q2_Gbase_Mlow_tlambda1	4574120	1545220	1321875	0.86	0.97	397904
sp_h90_q2_GDortel_Mbase_tlambda01	3268680	977924	1126008	1.15	0.73	375226
io_h90_q2_GDortel_Mlow_tlambda1	3268680	977924	1126008	1.15	0.73	375226
sp_h70_q1_GDortel_Mlow_tlambda01	3534200	1298690	1089213	0.84	0.84	435508
sp_h70_q1_GDortel_Mlow_tlambda1	4868050	1827130	954670	0.52	2.13	351826
sp_h70_q2_GDortel_Mlow_tlambda01	4849360	1845500	995185	0.54	2.12	359063
sp_h70_q2_GDortel_Mlow_tlambda1	4529580	1665200	1203930	0.72	1.49	308706
sp_h80_q1_GDortel_Mlow_tlambda01	4151220	1541230	1028487	0.67	1.67	301134
sp_h80_q1_GDortel_Mlow_tlambda1	4536520	1639230	930091	0.57	1.72	379436
sp_h80_q2_GDortel_Mlow_tlambda01	4473950	1602010	938955	0.59	1.89	366455
sp_h80_q2_GDortel_Mlow_tlambda1	4362420	1546760	1052797	0.68	1.43	346660
sp_h90_q1_GDortel_Mlow_tlambda01	4365040	1587110	1039687	0.66	1.65	341484
sp_h90_q1_GDortel_Mlow_tlambda1	4217730	1452980	864160	0.59	1.64	367902
sp_h90_q2_GDortel_Mlow_tlambda01	4321170	1526760	954213	0.62	1.52	392914
sp_h90_q2_GDortel_Mlow_tlambda1	4061540	1344250	1009770	0.75	1.28	343474

5.4 Diagnostics

Several diagnostic tools were run on the basic model to evaluate the model stability and performance, including Jitter analysis, ASPM analysis, retrospective analysis, and hindcasting.

5.4.1 Jitter analysis

The Jitter analysis examines the effect of varying the starting values of parameters on model convergence. A well-behaved model should converge on a global likelihood minimum across a reasonable range of input parameter values. The analysis is run using an automated routine in Stock Synthesis (Taylor et al. 2021), and involves repeatedly running the model after adding a small random jitter to the initial parameter values. Across the 20 model runs conducted, most converged on likelihood values close to the global minimum, but in 4 runs (~20%) the likelihood deviated from the minimum by more than 100 likelihood units (Figure 33). Another 7 runs differed from the global minimum by less than 50 units. All runs that converged to different likelihood values also generated SSB0 estimates that differed, mostly by small amounts but one by approximately 10%. The analysis suggested that there is some instability in the basic model.

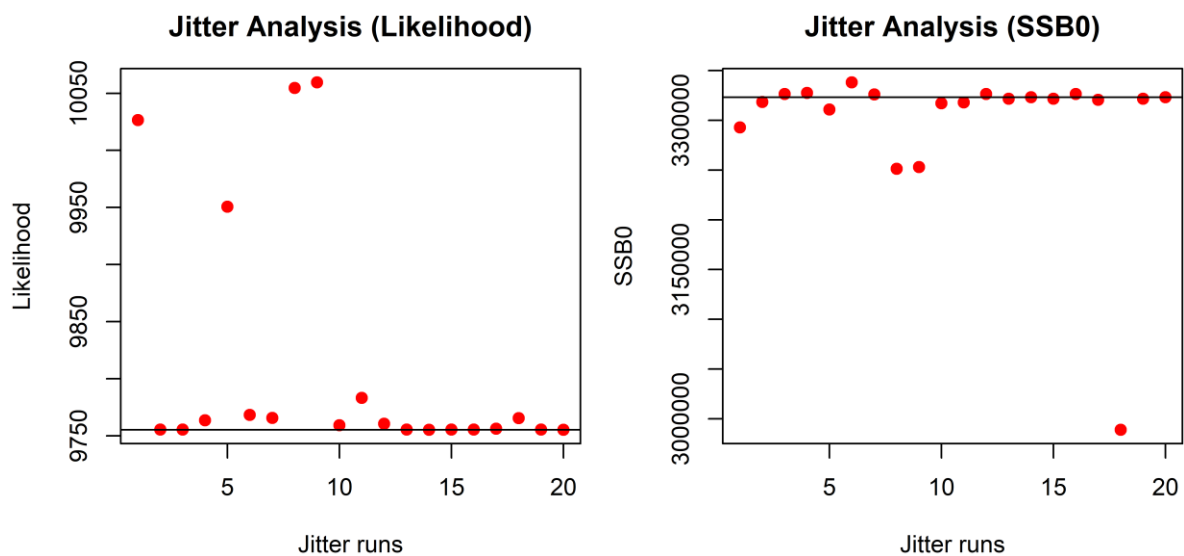


Figure 33: Jittering analysis performed to the basic model: likelihood values across the 20 Jittering model runs (left), and the associated estimates of SB0 (right). The line indicates the maximum likelihood estimate from the initial starting values.

5.4.2 ASPM analysis

The Age Structured Production Model (ASPM) analysis (Maunder & Piner 2015) was used to identify what is the main driver of the population trend, and how much influence the composition data has on the estimates of abundance. The ASPM analysis involved running the model with the length composition data removed from the model (selectivity parameters fixed) and recruitment deviates fixed to be zero. The stock biomass from the ASPM model runs is shown in Figure 34. Analysis shows that the catch and abundance index of yellowfin tuna populations are generally consistent, i.e., except for certain time periods, the catch alone can well explain the historical trend of population abundance from

a long-term perspective. The analysis also shows the length composition data has a certain influence on the estimation of the population scaling parameter, increasing the overall biomass level.

In particular, the increased catch in the 1980s (see Figure 2) conflicted with the increased CPUE, so the estimated number during this period was mostly driven by recruitment. If standardization did not adequately account for the increase in catchability due to the changes in fishing practice by the longline fleets in the 1980s, then the CPUE index may be biased. (The exploratory model run b1990 that excluded the CPUE indices prior to 1990 estimated a stock biomass that was more in line with the ASPM model. See Figure 2) The large fluctuations in the stock between 2005 and 2010 were related to the abnormal catch of yellowfin tuna in the mid-2000s (usually thought to be related to favourable oceanic conditions) and the subsequent decline in longline fishing operations due to piracy in 2007–2011. The catch has been high in recent years and the ASPM shows that if the recruitment since 2012 had been at the long term average, the reduction in biomass would have been greater.

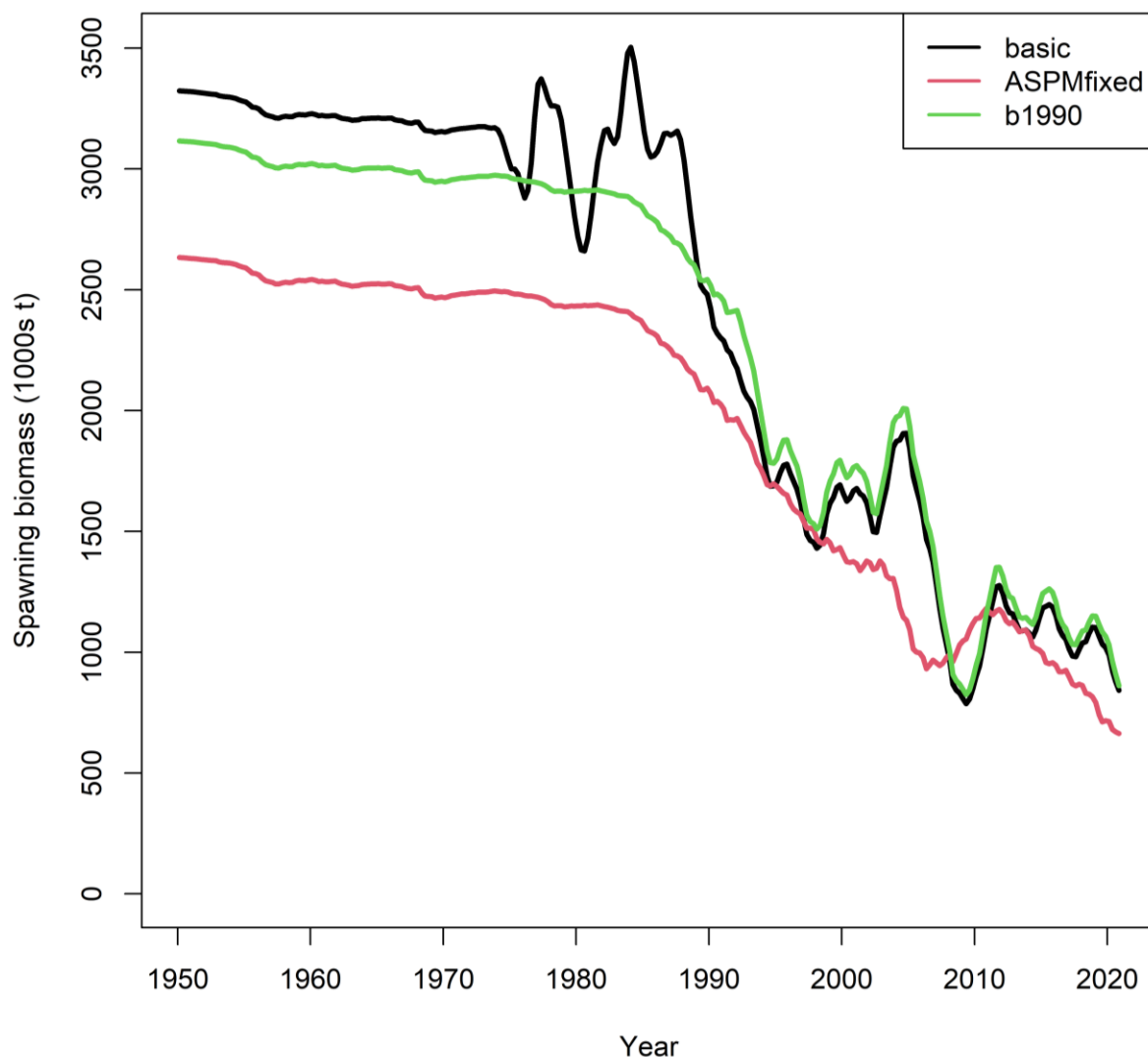


Figure 34: ASPM run for the basic model: recruitment deviations were fixed to zero and the length composition data excluded.

5.4.3 Retrospective analysis

Retrospective analysis is a diagnostic approach to evaluate the reliability of parameter and reference point estimates and to reveal systematic bias in the model estimation. It involves fitting a stock assessment model to the full dataset. The same model is then fitted to truncated datasets where the data for the most recent years are sequentially removed. The retrospective analysis was conducted for the last 5 years of the assessment time horizon to evaluate whether there were any strong changes in model results. The analysis involves sequentially removing 4 quarters of data at each trial.

The analysis of the basic model found that when the data is deleted for up to 2 years, the spawning population biomass is estimated to be stable, and when the data is deleted for 3-5 years, the estimated SB is lower (Figure 35). Ratio-based estimates (i.e., SB/SB_{msy}, SB/SB₀, and F) are generally very stable. The reasons for the retrospective pattern have been extensively investigated, but it is not clear what the main driving factors are. In particular, the following hypotheses have been examined:

- (1) Conflicts between observational datasets
- (2) Potential instability of the model due to the confounding between the stationary catchability, selectivity, and movement parameters
- (3) Misspecification of key productivity parameters, such as M and growth.
- (4) The complexity of spatial structure

Based on the above, the following model runs were performed:

- Removing or down-weighting the tagging dataset
- Down-weighting the length composition dataset or the longline length data after 2003
- Trying to balance the potential conflict between recent catches and LL CPUE by either halving (hypothetically) the catches since 2012 or allowing random walks on the LL catchability
- Model *bSeason* & *bSeasonX*
- An annual model without seasonal segmentation, which uses the observational data aggregated yearly.
- The revised model
- The revised model with no movement (recruitment occurred in all regions)
- A single area model fitting only to data in region 1

In general, these additional model runs were not able to eliminate the retrospective pattern although in some configurations, the pattern is much alleviated (e.g., a single area model with no spatial structure fitting only to data in R1 with recent catch halved) and it is likely that the retrospective pattern is related to a series of factors that affect the estimation of the stock productivity.

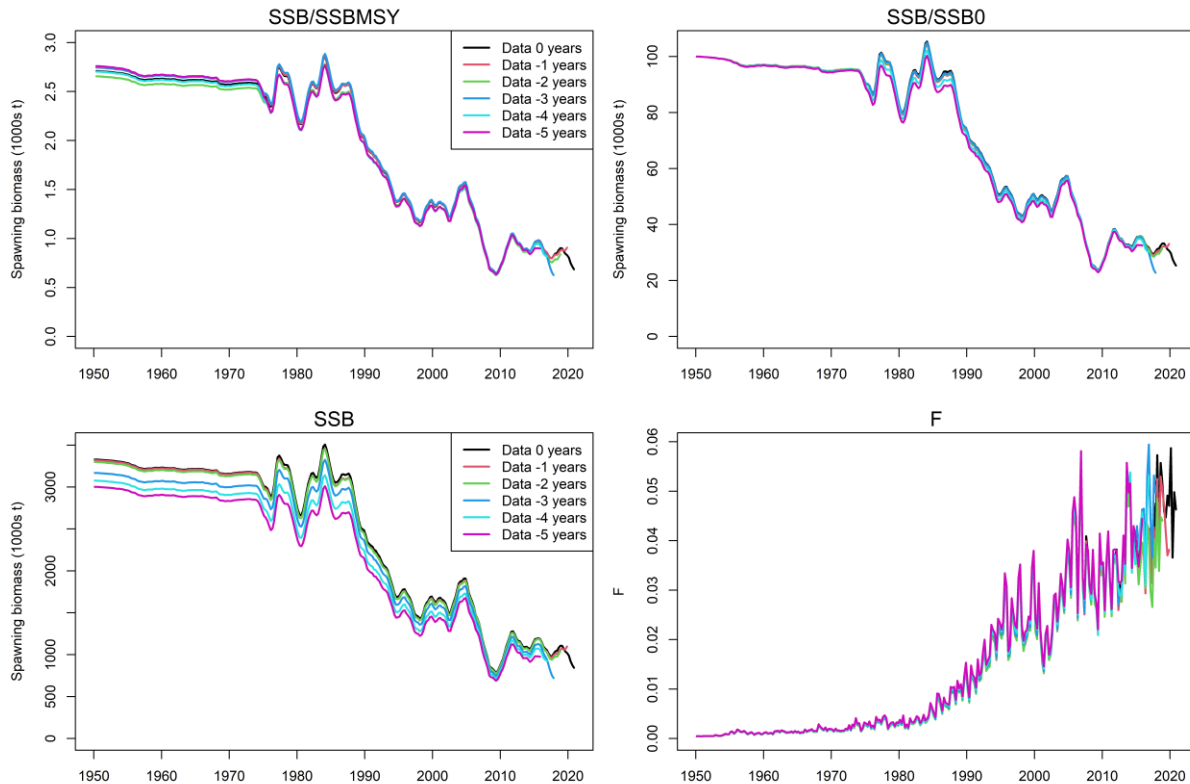


Figure 35: Retrospective analysis summary for the basic model.

5.4.4 Hindcasting analysis

Retrospective analysis evaluates the model’s stability with respect to recent data. The Hindcasting analysis (Kell et al. 2016) further assesses the model’s predictive power by making forward projections of the CPUE index using truncated models (i.e., models were fitted with data sequentially removed and were projected forward with catches added back in). The Hindcasting diagnostics were provided for the basic model using the ‘ss3diags’ package (Winker et al. 2020). The results show that the predictive ability of the model is reasonably stable, as the vulnerable biomass predicted in the truncated model is quite close to the prediction in the full model, and is within the error range of the observed CPUE, except for the LL1b CPUE in quarter 3.

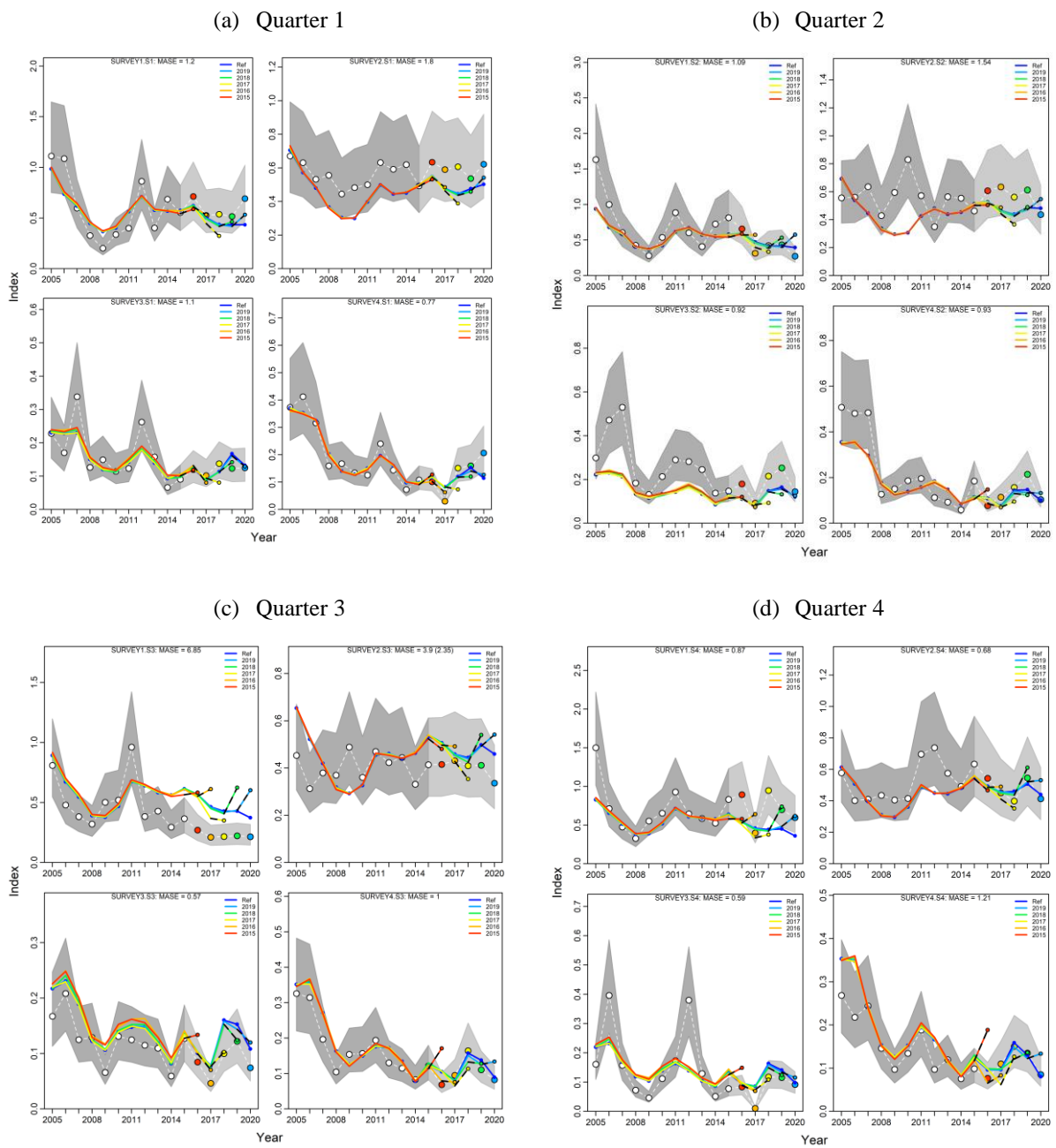


Figure 36: The Hindcasting analysis summary for the basic model: each panel shows the predicted quarterly longline CPUE index from models with data sequentially removed for 1, 2, 3, 4, 5 years.

6. STOCK STATUS

6.1 Current status and yields

MSY based estimates of stock status were determined for the final model options, including alternative assumptions on spatial structure, CPUE catchability, SRR steepness, the weighting of tag dataset, growth, and natural mortality. Stock status was determined for individual models (see Table 6), as well as for all (72) models combined including uncertainty from individual models based on estimated variance-covariance matrix of parameters (Table 7). The incorporation of both model and estimation uncertainty into management advice is necessary to accurately capture the current state of stock status as model uncertainty is not always greater than estimation uncertainty (Ducharme & Vincent 2021). MSY based reference points were derived for the model options based on the average F-at-age matrix for 2019–20 (Figure 32). The period was considered representative of the recent average pattern of exploitation from the fishery. However, variation in the proportion of catches between the main fishing gears are likely to influence the F-at-age matrix and, hence, MSY based indicators.

Amongst all 96 models, 55 models have achieved a maximum gradient at the ML estimate low than the pre-specified threshold of 0.0001, indicating full convergency. 14 models have a final maximum gradient greater than 0.01, and 12 of them were associated the model options of low M and Dortel et al. (2015) growth combination. All models have achieved Positive Definite Hessian (PDH) at the ML estimate, from which the estimation uncertainty can be quantified. A simple diagnostic based on Mean Squared Error (MSE) indicated there is very little variability in the fits to the CPUE indices amongst the model options (Figure B13).

For the final model options, point estimates of MSY ranged 293 482– 525 680mt, with an average of 393 827 mt, slightly lower than the most recent annual catches of about 432 600 mt is (Table 7). Model options with higher steepness and lower weighting of the tag dataset generally yielded comparatively higher estimates of MSY. The annual trends in F_t/F_{MSY} and SB_t/SB_{MSY} for 1950–2020 were computed across the final model options (Figure 37). Prior to 1990, exploitation rates were low and adult biomass remained well above SB_{MSY} . In the early 1990s, F_t/F_{MSY} increased and biomass levels declined before stabilizing during the mid-1990s–early 2000s. Overall fishing mortality rates increased sharply in 2005 with the large increase in catches during 2004/2005. Adult biomass declined considerably in the subsequent years, attributable to a period of very low recruitment during 2004–2006 and declined below the SB_{MSY} level in 2008. The stock rebounded during 2009-2012 before declining below the SB_{MSY} level

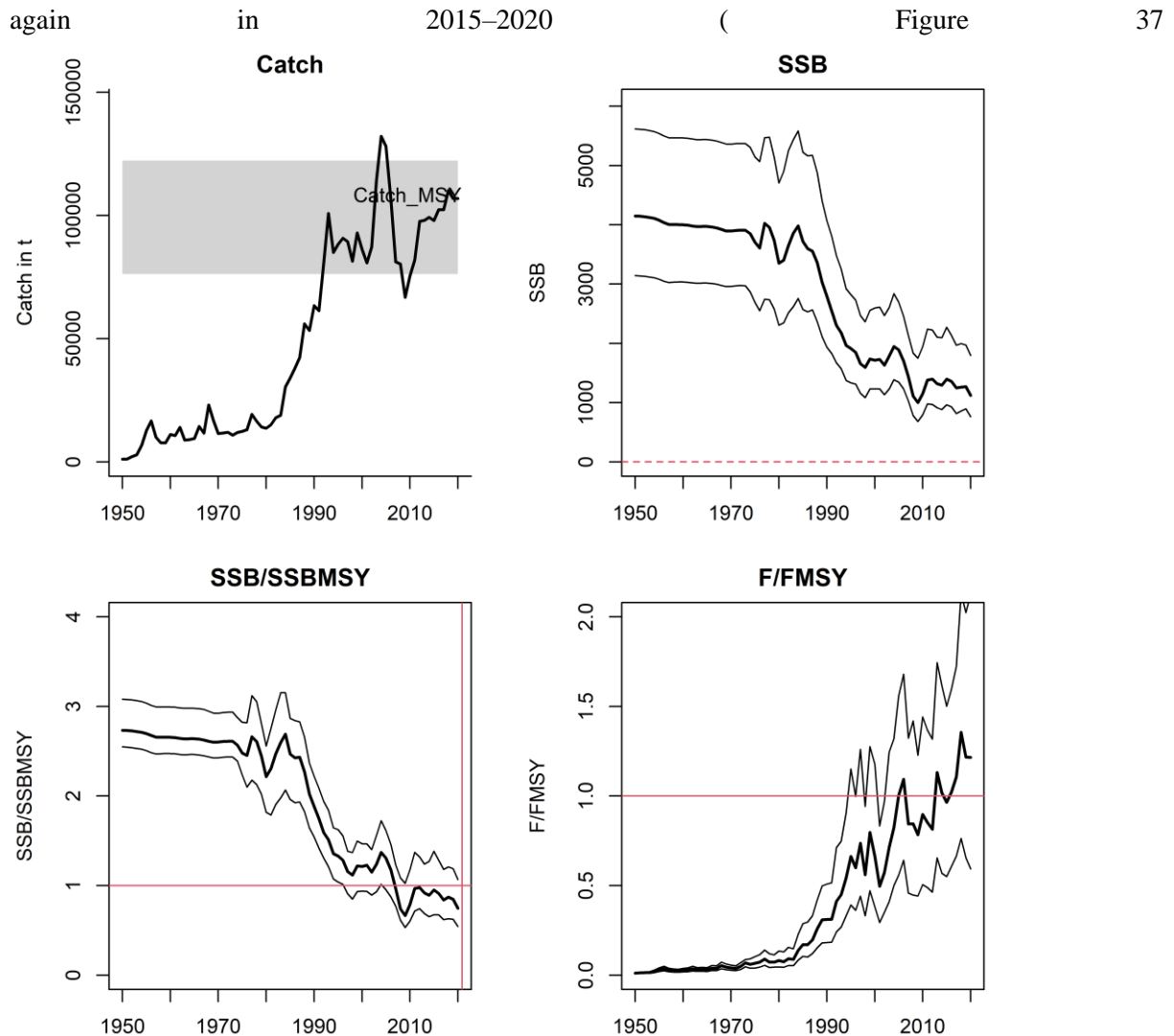


Figure 37).

In general, current stock status relative to the MSY based benchmarks is not fundamentally different between the two alternative spatial structures (Figure 38), although the models with the revised spatial structure yielded higher estimates of biomass, depending on other model options (see Table 6). The levels of stock depletion were sensitive to the Longline CPUE catchability options, i.e., depletion levels were higher assuming the memory of piracy did not affect longline catchability for the next eight years (Figure 38). The levels of current fishing mortality were more sensitive to the growth option, i.e., current fishing mortality (relative to F_{MSY}) is much higher for models assuming the growth estimates of Dortel et al. (2015) or the lower natural mortality option (Figure 39). The levels of stock depletion and fishing mortality are relatively insensitive to the relative weighting of the tagging data, nor the options of steepness value, although the estimate of absolute stock abundance are obviously dependent on these alternative options (Figure 39). Amongst 96 models, 86 models (90%) estimated current (2020) spawning biomass to be below SB_{MSY} level ($SB_{2020}/SB_{MSY} < 1.0$), 63 models (66%) estimated current fishing mortality to be above the F_{MSY} level ($F_{2020}/F_{MSY} > 1.0$) (see Table 6).

Estimates were combined across from the 72 models to generate the KOBE stock status (Table 7, Figure 39). For individual models, the uncertainty is characterised using the multivariate lognormal Monte-Carlo approach (Walter et al. 2019, Walter & Winker 2019, Winker et al. 2019), based on the maximum

likelihood estimates and variance-covariance of F/F_{MSY} and SB/SB_{MSY} . Thus, estimates of stock status included both within and across model uncertainty. Combined across the model ensemble, SB_{2020} was estimated to be of $0.78 SB_{MSY}$ (0.57– 0.98), and F_{2020} was estimated $1.27 F_{MSY}$ (0.64–1.91) (Table 7) The probability of the stock being in the red Kobe quadrant in 2020 is estimated to be about 67%, and the probability in the green Kobe quadrant is estimated to be 10% (The probability in the yellow and orange quadrant is 24% and <1%, respectively). The stock is therefore considered to be overfished and subject to overfishing in 2020.

Table 7: Estimated Status of yellowfin tuna in the Indian Ocean from the model ensemble.

Catch in 2020 :	432623
Average catch 2016–2020:	434568
MSY (1000 mt) (plausible range):	394 (325 –463)
F_{MSY}	0.18 (0.14–0.21)
SB₀(1000 mt) (80% CI):	4192 (3228–5156)
SB₂₀₂₀ (1000 mt) (80% CI):	1162 (773–1550)
SB_{MSY}	1515 (1146 – 1885)
SB₂₀₂₀/SB₀ (80% CI):	0.28 (0.21–0.34)
SB₂₀₂₀ / SB_{MSY}	0.78 (0.57–0.98)
F₂₀₂₀ / F_{MSY}	1.27 (0.64–1.91)

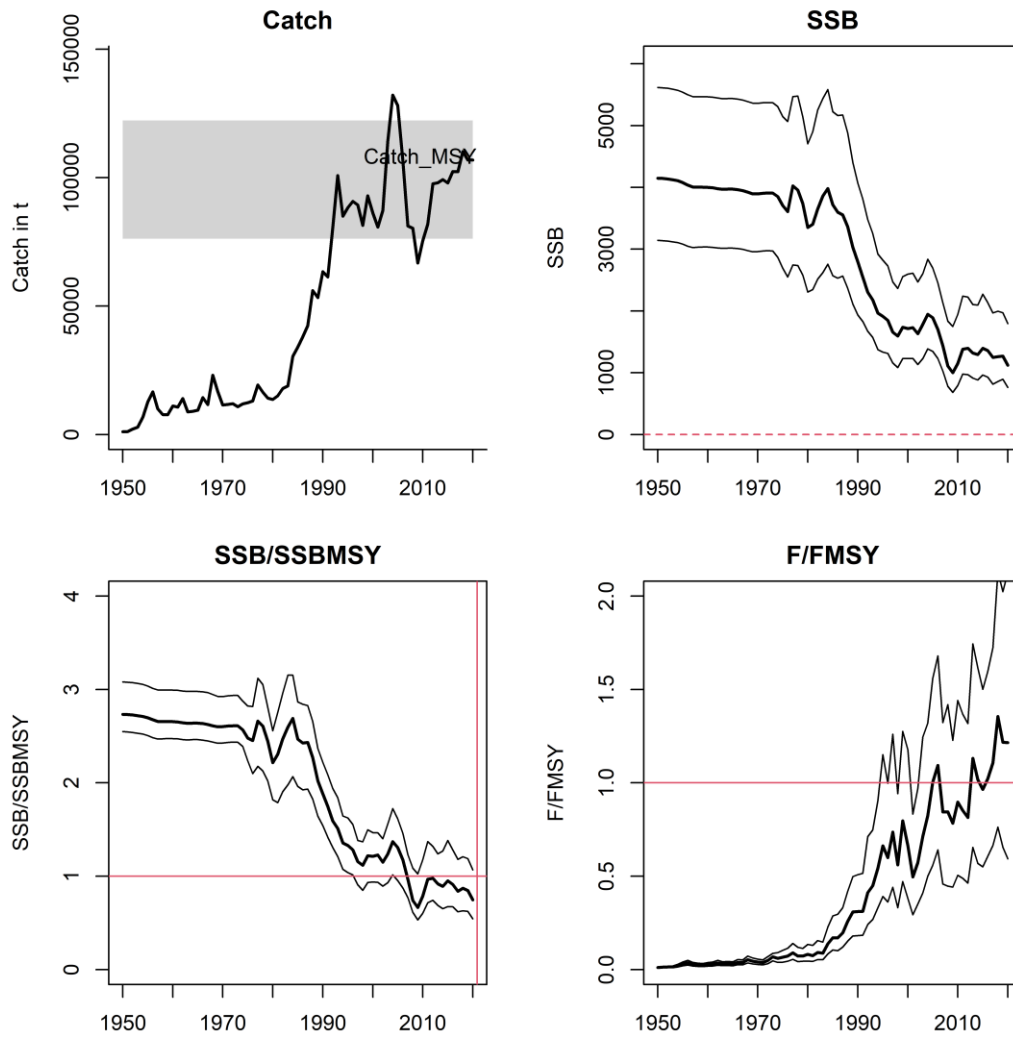


Figure 37: Estimated stock trajectories for the Indian Ocean yellowfin tuna from the final model grid. Thick black lines shaded areas represent 5th and 95th percentiles across all models. In the catch plot, dotted lines represent estimate of MSY (quarterly), the shaded area represents 5th and 95th percentiles.

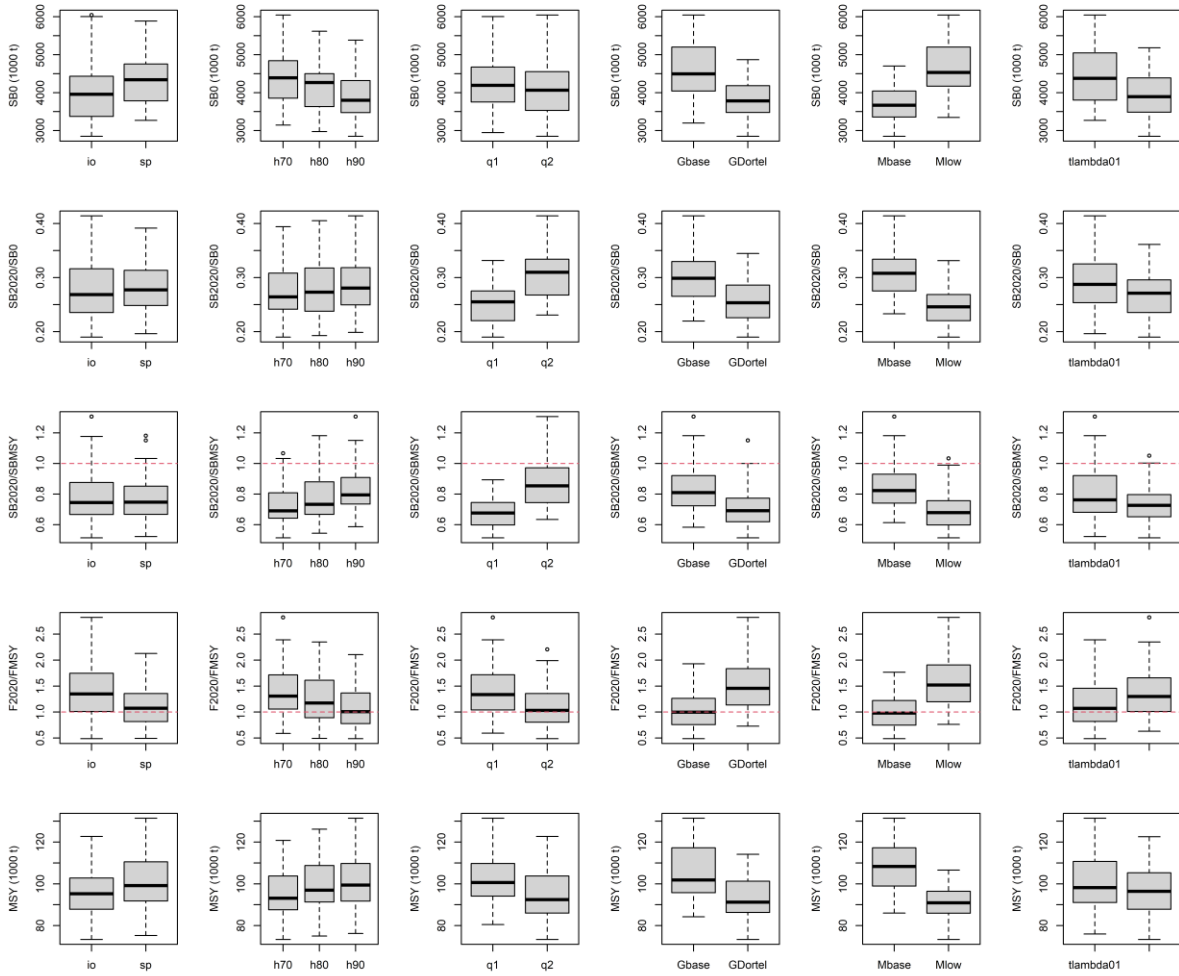


Figure 38: Distribution of management quantities estimates: SB0, SB2020/SB0, SB2020/SBMSY, F2020/FMSY, and MSY for the 96 models, partitioned by assessment options (see Table 5).

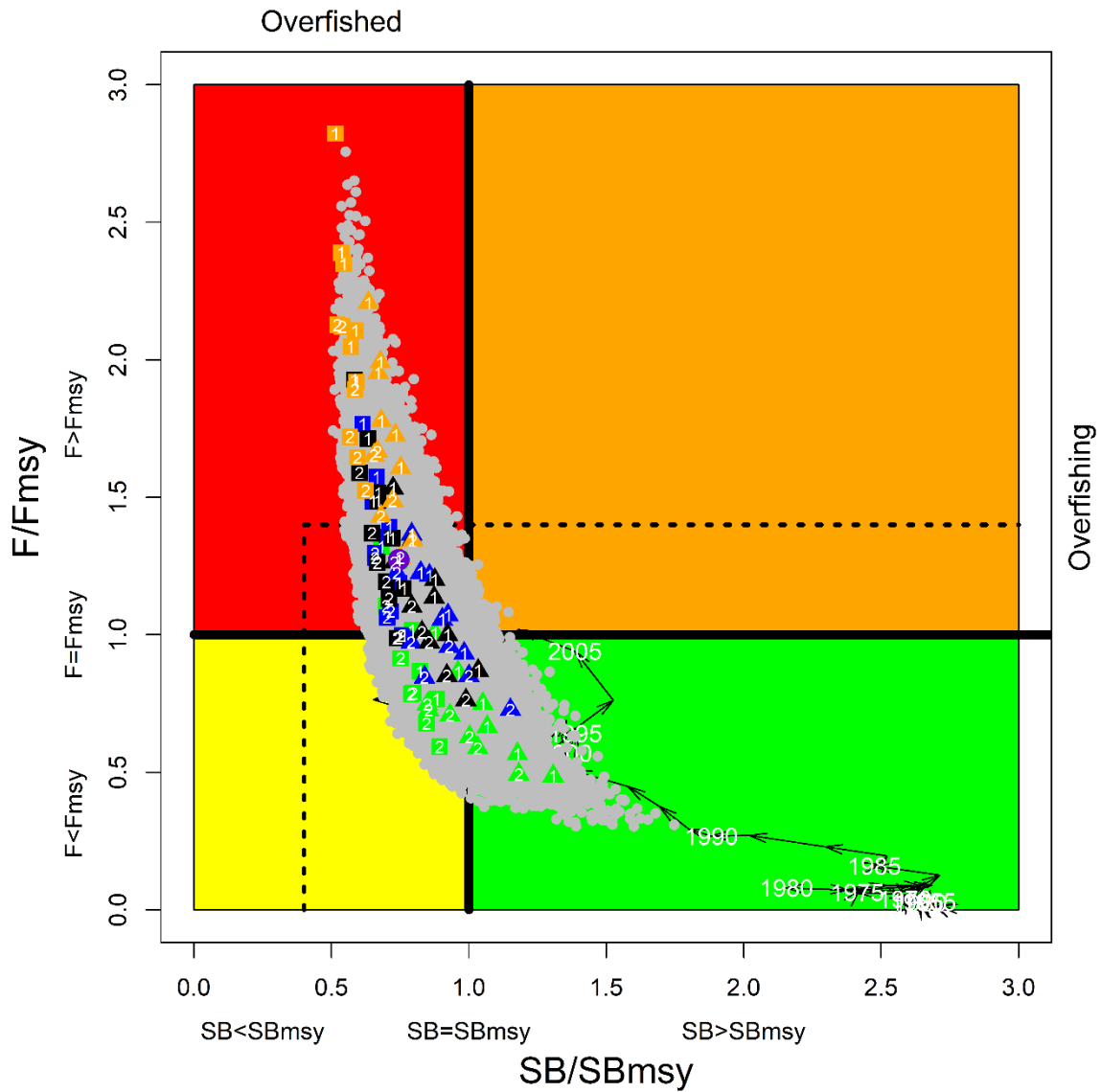


Figure 39: current stock status, relative to SB_{MSY} (x-axis) and F_{MSY} (y-axis) reference points for the final model options. Coloured symbols represent MPD estimates from individual models: square and Triangles and represents LL CPUE catchability options q_1 and q_2 respectively; green, blue, black, and orange represents growth and natural mortality option combination G_{base_Mbase} , $GDortel_Mbase$, G_{base_Mlow} , and $GDortel_Mlow$ respectively; 1,2, represents spatial structure option io and sp respectively). The purple dot and arrowed line represent estimates of the basic model. Grey dots represent uncertainty from individual models. The dashed lines represent limit reference points for IO yellowfin tuna ($SB_{lim} = 0.5 SB_{MSY}$ and $F_{lim} = 1.4 F_{MSY}$).

7. DISCUSSION

This report presents a preliminary stock assessment for Indian Ocean yellowfin tuna using a sex-aggregated, age-structured, spatially explicit Stock Synthesis model. It represents an update and revision of the 2018 assessment model with newly available information, including updated and revised CPUE indices and length composition data. There is no fundamental change in the structure of the assessment model compared to the previous assessment (Fu et al. 2018a) but progress and improvements made through additional analysis since the previous assessment has been included. A basic model was configured to assess the model performance. A series of exploratory model runs were performed to address issues in observational datasets and improve the stability of the assessment model. The final model options involved a combination of model settings related to the spatial structure, stock-recruitment steepness, tag data weighting, CPUE catchability, and key biological parameters. These model options have contributed the main source of uncertainty associated with estimates of stock dynamics and productivity. The final estimates of stock status are based on a model grid of 72 models, incorporating uncertainty estimates from both within and across the model ensemble. The overall stock status estimates obtained from the range of model options do not differ substantially from the previous assessment. Biomass was estimated to have continued to decline in more recent years since the last assessment. Spawning biomass in 2020 was estimated to be 78% of the level that supports the maximum sustain yield ($SB_{2020}/SB_{MSY} = 0.78$). Current fishing mortality was estimated to be 27% higher than FMSY ($F_{2020}/F_{40\%SB} = 1.27$). The probability of the stock being currently in the red Kobe quadrant is estimated to be 67%. The catches in the last five years remains higher than the estimated MSY. The recent longline CPUE and purse seine indices are stable but remains at a historically low level. Considering the quantified uncertainty, the stock is considered to be overfished and is subject to overfishing in 2020.

As earlier assessments, the models presented here, while providing a reasonable fit to some key data sets (e.g., the CPUE indices), also show some signs of poor fit (e.g., LF data). There are conflicts amongst observational datasets, noticeably between the CPUE and tag data, and between the length composition and tag data, and the model estimates are sensitive to the relative weighting of these data. Estimates of movement rates were probably more influenced by model configurations than tag data. The nature and extent of the dispersal of tagged fish remains a key uncertainty in the assessment. The Longline CPUE is generally consistent with the catch history of yellowfin tuna over the long term except that in certain periods changes in the relative abundance index cannot be fully explained by the catches. The catches between 2013 and 2020 are close to an historically high level but both the longline and purse seine CPUE remained relative stable during recent years. If the CPUE is a reliable index of abundance, it may indicate that the stock has been fished down to a relatively low level.

Excluding the size data from the Taiwanese and Seychelles longline logbook helped eliminate some of the unstable trends in the observations that are mainly attributed to data reporting. However, there is still lack of fit to the longline size data, and the model does not predict well the larger fish that appear in the later stages of each time series (the timing may be different in each region, but generally from 2003 onwards). This may reflect changes in selectivity or sampling issues, and both possibilities have been explored. The longline length data since the early 2000s was sampled from a larger mix of fleets and are much noisier than the early years which are dominated by the Japanese data. A more diverse fleet mix is more likely to include data with poorer sampling and coverage characteristics. On the other hand, it can be argued that there was a fleet-level switch to larger fish due to changing gear characteristics. There is little information about possible selectivity changes, but there is more evidence that the quality of sampling and reporting varies significantly between fleets.

The previous yellowfin assessment model showed a certain degree of instability and the likelihood function seemed to have local minima. The model would converge at either low or high rates of movement, leading to estimation of inconsistent regional population trends. Although this problem did not reappear in the current assessment, the cause of the problem may still exist (as verified through a profile likelihood analysis on movement parameters). It is hypothesised that the problem may be related to the potential confounding of seasonal movement, catchability, and selectivity. A seasonally structured model that specifically considered seasonal catchability and selectivity changes for the main longline fisheries did not seem to fundamentally improve the model performance (in terms of improving stability and retrospectives). In a highly parameterized and spatially structured model, the estimation of movement is complex, in that movement can be affected by many factors, including dispersion of tags, regional biomass distribution, assumptions on recruitment dynamics, and the spatial and seasonal variability in the CPUE and length composition data. Given the limited release site, the tagging data is unlikely to be informative of regional migration, and movement estimates are expected to be affected by the locations of regional boundaries. For example, most tagging data are very close to the boundary between region 1 and 2, so the probability of crossing the boundary is much higher than the average for the population. Purse seine effort in region 2 also tends to be very close to the boundary, so they will be more likely to find tags from region 1. These factors can lead to overestimation of population movement rates. Therefore, the revised regional structure reduces the impact of the tag data on movement estimation and is a useful addition to the model ensemble.

The 2018 assessment model configuration was affected by a problem with the setup of the modelling software. A substantial number of models yielded non-sensical results, with the stock crashing within a few years into the projection period even under low catch scenarios. Kolody & Jumppanen (2021) also found that the model struggled to remove the catch that was reported in 2019 when the terminal year was extended. The problem seems to have been related to the parameterisation of recruitment distribution which constrained the regional recruitment during the projection period. In the current assessment, with the revision of recruitment parameters, this problem has not recurred. The current model exhibits a retrospective pattern, with a tendency to estimate lower productivity as data are sequentially removed. This pattern appears to be persistent among the explored models and may be related to a series of factors that affect the estimation of the stock productivity. We recommend further analysis to investigate this issue.

Growth is an important source of uncertainty. As the change in mean size of a fished population relative to the unfished state is usually interpreted by the model as linked to fishery-induced depletion, the lack of large fish in the catch, relative to a higher asymptotic length in the model would imply a higher level of fishing mortality and, hence a large fishery depletion effect. The empirical estimate of Fonteneau (2008) used in the assessment model has a L_{inf} parameter value of 145 cm. This value is based around studies of the early life history growth of yellowfin tuna, i.e., of immature / maturing fish. The estimate of Dortel et al. (2015) is higher (156 cm), but the study is still based on samples from purse seine fishery which catches mostly small fish. These L_{inf} estimates are low when compared to estimates from other oceans or estimates from many regional studies in the Indian Ocean (S. Creech, per. Comm.) IOTC's summary information on yellowfin tuna indicates that the species can grow up to 240 cm and individuals of more than 200 cm were regularly caught between 2010 and 2019. Age and growth study to better quantify the average size of the oldest cohort of yellowfin should be a research priority.

8. ACKNOWLEDGMENTS

We are grateful to the many people that contributed to the collection of this data historically, analysts involved in the CPUE standardization, and developers for providing the SS3 software, and in particular to Adam Langley who conducted the previous assessments, and to Henning Winker who provided the ssDiag package, and the yellowfin work group that provided many useful suggestions for the assessment.

9. REFERENCES

- Andrews, A.H., Pacicca, A., Allman, R., Falterman, B.J., Lang, E.T., Golet, W. (2020) Age validation of yellowfin (*Thunnus albacares*) and bigeye (*Thunnus obesus*) tuna of the northwestern Atlantic Ocean. *Canadian Journal of Fisheries and Aquatic Sciences*, 77(4): 637-643.
- Baidai, Y., Dagorn, L., Gaertner, D., Deneubourg, J.L., Duparc, A., Floch, L., Capello, M. 2021. Associative Behavior-Based abundance Index (ABBI) for yellowfin tuna (*Thunnus albacares*) in the Western Indian Ocean. IOTC-2021-WPTT23(DP)-15.
- Chassot, E. 2014. Are there some small yellowfin tunas caught in free-swimming schools? IOTC-2014-WPDCS-10-INF05.
- Castillo-Jordan, C., Hampton, J., Ducharme-Barth, N., Xu, H., Vidal, T., Williams, P., Scott, F., Pilling, G., Hamer, P. (2021) Stock assessment of South Pacific albacore, WCPFC-SC-2017/SA-WP-02. Western and Central Pacific Fisheries Commission Scientific Committee, Seventeenth Regular Session, Online meeting, 10 August 2021.
- Dammannagoda, S.T., Hurwood, D.A., Mather, P.B. 2008. Evidence for fine geographical scale heterogeneity in gene frequencies in yellowfin tuna (*Thunnus albacares*) from the north Indian Ocean around Sri Lanka. *Fisheries Research* 90 147–157.
- Dortel, E., Sardenne, F., Bousquet, N., Rivot, E., Million, J., Croizier, G. Le., Chassot, E. 2014. An integrated Bayesian modeling approach for the growth of Indian Ocean yellowfin tuna: *Fish. Res.* (2014), <http://dx.doi.org/10.1016/j.fishres.2014.07.00>.
- Ducharme-Barth, N., Vincent, M. 2021. Focusing on the front end: A framework for incorporating uncertainty in biological parameters in model ensembles of integrated stock assessments. WCPFC-SC17-2021/SA-WP-05.
- Eveson, P., Million, J., Sardenne, F., Le Croizier, G. 2012. Updated growth estimates for skipjack, yellowfin and bigeye tuna in the Indian Ocean using the most recent tag-recapture and otolith data. IOTC-2012-WPTT14-23.
- Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences* 68: 1124–1138.
- Fonteneau, A. 2008. A working proposal for a Yellowfin growth curve to be used during the 2008 yellowfin stock assessment. IOTC-2008-WPTT-4.
- Froese, R., Pauly, D.E. 2009. *FishBase*, version 02/2009, FishBaseConsortium.
- Fu, D. 2017. Indian ocean skipjack tuna stock assessment 1950-2016 (stock synthesis). IOTC–2017–WPTT19–47_rev1.
- Fu, D., Langley, A., Merino, G., Ijurco, A.U. 2018a. Preliminary Indian Ocean Yellowfin Tuna Stock Assessment 1950-2017 (Stock Synthesis). IOTC–2018–WPTT20–33.
- Fu, D., Langley, A., Merino, G., Ijurco, A.U. 2018b. Indian Ocean Yellowfin Tuna SS3 Model Projections. IOTC–2018–SC21–16.

- Fu, D. 2020. Tag Data Processing for IOTC Tropical Tuna Assessments. IOTC–2020–WPTT22(DP)–10.
- Gaertner, D., Hallier, J.P. 2015. Tag shedding by tropical tunas in the Indian Ocean and other factors affecting the shedding rate. *Fisheries Research*. 2015/163.
- Geehan, J., Setuadji, B. 2018. Revision to the IOTC scientific estimates of Indonesia’s fresh longline catches. OTC-2018-WPB16-22.
- Global Tuna Alliance. 2021. Sustainability of Yellowfin tuna (*Thunnus albacares*) fisheries in the Indian Ocean, with a special focus on juvenile catches.
- Greehan, J., Hoyle, S. 2013. Review of length frequency data of the Taiwanese distant water longline fleet. IOTC-2013-WPTT15-41.
- Guery, L., Kaplan, D., Grande, M., Merino, G., Marsac, F., Abascal, F., Baez, J.C., Gaertner, D. 2021. European Purse Seine CPUE Standardization: Methodology and Framework for the YFT Stock Assessment. IOTC-2021-WPTT23(DP)-16
- Harley, S.J. 2011. Preliminary examination of steepness in tunas based on stock assessment results. WCPFC SC7 SA IP-8, Pohnpei, Federated States of Micronesia, 9–17 August 2011.
- Hillary, R.M., Million, J., Anganuzzi, A., Areso, J.J. 2008a. Tag shedding and reporting rate estimates for Indian Ocean tuna using double-tagging and tag-seeding experiments. IOTC-2008-WPTDA-04.
- Hillary, R.M., Million, J., Anganuzzi, A., Areso, J.J. 2008b. Reporting rate analyses for recaptures from Seychelles port for yellowfin, bigeye and skipjack tuna. IOTC-2008-WPTT-18.
- Hoyle, S. 2012. Tagger effects – models to estimate tag loss and mortality for stock assessment. Indian Ocean Tuna Tagging Symposium, Grand Baie International Conference Centre, Royal Road, Grand Baie, MAURITIUS, 30 October – 2 November 2012.
- Hoyle, S.D.; Leroy, B.M.; Nicol, S.J.; Hampton, W.J., 2015. Covariates of release mortality and tag loss in large-scale tuna tagging experiments. *Fisheries Research*. 163:106-118.
- Hoyle, S.D., Langley, A. 2018. Indian Ocean tropical tuna regional scaling factors that allow for seasonality and cell areas. IOTC-2018-WPM09-13.
- Hoyle, S.D., Okamoto, H., Yeh, Y., Kim, Z., Lee, S.4 and Sharma, R. (2015) IOTC–CPUEWS–02 2015: Report of the Second IOTC CPUE Workshop on Longline Fisheries, April 30th–May 2nd, 2015. IOTC–2015–CPUEWS02–R[E]: 128pp.
- Hoyle, S.D., Satoh, K., Matsumoto, T., 2017. Causes of the historical discontinuity in Japanese longline CPUE. Indian Ocean Tuna Commission, WPM08. Victoria, Seychelles
- Hoyle, S., Chang, S.T., Fu, D., Geehan, J., Itoh, T., Lee, S.I., Matsumoto, T., Yeh, Y.M., Wu, R.F. 2021. IOTC-2021-WPTT(DP)23-08: Review of size data from Indian Ocean longline fleets, and its utility for stock assessment.
- Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystem and aquaculture. *Journal of Fish Biology*, 42:627–647.

- IOTC 2008a. Report of the First Session of the IOTC Working Party on Tagging Data Analysis, Seychelles, 30 June to 4 July 2008. IOTC-2008-WPTDA-R[E].
- IOTC 2008b. Report of the 10th session of the IOTC Working Party on Tropical Tunas, Bangkok, Thailand, 23 to 31 October 2008. IOTC-2008-WPTT-R[E].
- IOTC 2018a. Report of the 20th Session of the IOTC Working Party on Tropical Tunas. Seychelles, 29 October – 3 November 2018. IOTC–2018–WPTT20–R. 131 pp.
- IOTC 2018b. Report of the 21st Session of the IOTC Scientific Committee. Seychelles, 3 – 7 December 2018. IOTC–2018–SC21–R. 250 pp.
- IOTC 2020. Report of the 23rd Session of the IOTC Scientific Committee. Seychelles, 7 – 11 December 2020. IOTC–2020–SC23–R[E]: 211pp.
- IOTC 2021a. Review of Yellowfin Tuna Statistical Data. IOTC-2021-WPTT23(DP)-07_Rev1
- IOTC 2021b. Report of the 23rd Session of the IOTC Working Party on Tropical Tunas. Online, 22 - 24 June 2020. IOTC–2021–WPTT23(DP)–R[E]: 35 pp.
- Kitakado, T., Wang, S.P., Satoh, K., Lee, S.I. Tsai, W.P., Matsumoto, T., Yokoi, H., Okamoto, K., Lee, M.K., Lim, J.H., Kwon, Y.I., Su, N.J., Chang, S.J., Chang, F.C. 2021. Report of trilateral collaborative study among Japan, Korea and Taiwan for producing joint abundance indices for the yellowfin tunas in the Indian Ocean using longline fisheries data up to 2019. IOTC–2021-WPTT23(DP)-14.
- Kleiber, P., Hampton, J., Fournier, D.A. 2003. MULTIFAN-CL Users' Guide. <http://www.multifan-cl.org/userguide.pdf>.
- Kell, L.T., Kimoto, A., Kitakado, T., 2016. Evaluation of the prediction skill of stock assessment using hindcasting. *Fisheries research*, 183, pp.119-127.
- Kolody, D. 2011. Can length-based selectivity explain the two stage growth curve observed in Indian Ocean YFT and BET. IOTC–2011–WPTT13–33.
- Kolody, D., Herrera, M., Million, J. 2011. Indian Ocean Skipjack Tuna Stock Assessment 1950-2009 (Stock Synthesis). IOTC-2011-WPTT13-31_Rev1.
- Kolody, D. 2018. Estimation of Indian Ocean Skipjack Purse Seine Catchability Trends from Bigeye and Yellowfin Assessments. IOTC–2018–WPTT20–xx.
- Kolody, D.; Hoyle, S.D., 2013. Evaluation of tag mixing assumptions for skipjack, yellowfin and bigeye tuna stock assessments in the Western Pacific and Indian Oceans. WCPFC Scientific Committee, 9th Regular Session. Pohnpei, Federated States of Micronesia: Western and Central Pacific Fisheries Commission.
- Kolody, D., Jumpanen, Paul. 2021. Indian Ocean Yellowfin Tuna Management Procedure Evaluation Update March 2021. IOTC-2021-WPM12(MSE)-03
- Kunal, S.P., Kumar, G., Menezes, M.R., Meena, R.M. (2013). Mitochondrial DNA analysis reveals three stocks of yellowfin tuna *Thunnus albacares* (Bonnaterre, 1788) in Indian waters. *Conserv Genet* (2013) 14:205–213.

- Langley, A. 2012. An investigation of the sensitivity of the Indian Ocean MFCL yellowfin tuna stock assessment to key model assumptions. IOTC-2012-WPTT-14-37.
- Langley, A. 2015. Stock assessment of yellowfin tuna in the Indian Ocean using Stock Synthesis. IOTC-2012-WPTT-17-30.
- Langley, A. 2016a. An update of the 2015 Indian Ocean Yellowfin Tuna stock assessment for 2016. IOTC-2016-WPTT18-27.
- Langley, A. 2016b. Stock assessment of bigeye tuna in the Indian Ocean for 2016 — model development and evaluation. IOTC-2016-WPTT18-20.
- Langley, A., Hampton, J., Herrera, M., Million, J. 2008. Preliminary stock assessment of yellowfin tuna in the Indian Ocean using MULTIFAN-CL. IOTC-2008-WPTT-10.
- Langley, A., Herrera, M., Hallier, J.P., Million, J. 2009. Stock assessment of yellowfin tuna in the Indian Ocean using MULTIFAN-CL. IOTC-2009-WPTT-11.
- Langley, A., Herrera, M., Million, J. 2010. Stock assessment of yellowfin tuna in the Indian Ocean using MULTIFAN-CL. IOTC-2010-WPTT-12.
- Langley, A., Herrera, M., Million, J. 2011. Stock assessment of yellowfin tuna in the Indian Ocean using MULTIFAN-CL. IOTC-2011-WPTT-13.
- Langley, A., Herrera, M., Million, J. 2012a. DRAFT Stock assessment of yellowfin tuna in the Indian Ocean using MULTIFAN-CL. IOTC-2012-WPTT-14-38.
- Langley, A., Herrera, M., Million, J. 2012b. Stock assessment of yellowfin tuna in the Indian Ocean using MULTIFAN-CL. IOTC-2012-WPTT-14-38.
- Langley, A., Million, J. 2012. Determining an appropriate tag mixing period for the Indian Ocean yellowfin tuna stock assessment. IOTC-2012-WPTT-14-31.
- McKechnie, S., Pilling, G., Hampton, J. 2017. Stock assessment of bigeye tuna in the western and central Pacific Ocean. WCPFC-SC13-2017/SA-WP-05 Rev1.
- McKechnie, S. Hampton, J., Pilling, G. M., Davies N. 2016. Stock assessment of skipjack tuna in the western and central Pacific Ocean. WCPFC-SC12-2016/SA-WP-04.
- Maunder, M.N., Aires-da-Silva, A. 2012. A review and evaluation of natural mortality for the assessment and management of yellowfin tuna in the eastern Pacific Ocean. External review of IATTC yellowfin tuna assessment. La Jolla, California. 15-19 October 2012. Document YFT-01-07.
- Maunder, M. N., and Piner, K. R. 2015. Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science*, 72: 7–18.
- Medley, P., Ahusan, M. 2021. Bayesian Skipjack and Yellowfin Tuna CPUE Standardisation Model for Maldives Pole and Line 1970-2019. IOTC-2021-WPTT23(DP)-13.
- Merino, G., Adam, S., Murua, H., Fu, D., Bruyn, P. 2019. State of the development of the workplan to improve the current assessment of yellowfin tuna. IOTC-2019-SC22-INF01.

- Merino, G. Indian Ocean yellowfin workplan (2019-2020). IOTC-2020-WPTT22(AS)-INF01
- Methot, R.D., Wetzel, C.R., Taylor, I.G., Doering, Kathryn 2020. Stock Synthesis User Manual Version 3.30.16.
- Methot, R.D., Wetzel, C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research* 142 (2013) 86–99.
- Methot, R.D. 2019. Recommendations on the configuration of the Indian Ocean yellowfin tuna stock assessment model.
- Ménard, F., Lorrain, A., Potier, M., Marsac, F. 2007. Isotopic evidence of distinct feeding ecologies and movement patterns in two migratory predators (yellowfin tuna and swordfish) of the western Indian Ocean. *Marine Biology*. 153:141-152.
- Nishida, T., Shono, H. 2005. Stock assessment of yellowfin tuna (*Thunnus albacares*) resources in the Indian Ocean by the age structured production model (ASPM) analyses. IOTC-2005-WPTT-09.
- Nishida, T., Shono, H. 2007. Stock assessment of yellowfin tuna (*Thunnus albacares*) in the Indian Ocean by the age structured production model (ASPM) analyses. IOTC-2007-WPTT-12.
- Pacicco, A.E., Allman, R.J., Lang, E.T., Murie, D.J., Falterman, B.J., Ahrens, R., Walter, J.F. 2021. Age and Growth of Yellowfin Tuna in the U.S. Gulf of Mexico and Western Atlantic. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science* 13:345–361.
- Polacheck, T., Preece, A., Betlehem, A., & Klaer, N. (1997). . In *Int. Symp. on Fishery Stock Assessment Models for the 21st Century*, Anchorage, Alaska, EEUU. 8-11 Oct 1997
- Shih, C.-L., Hsu, C.-C., Chen, C.-Y. (2014) First attempt to age yellowfin tuna, *Thunnus albacares*, in the Indian Ocean, based on sectioned otoliths. *Fisheries Research*, 149: 19-23.
- Suzuki, Z. 1993. A review of the biology and fisheries for yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. In Shomura, R.S.; Majkowski, J.; Langi S. (eds). *Interactions of Pacific tuna fisheries. Proceedings of the first FAO Expert Consultation on Interactions of Pacific Tuna Fisheries. 3–11 December 1991. Noumea, New Caledonia. Volume 2: papers on biology and fisheries. FAO Fisheries Technical Paper. No. 336, Vol.2. Rome, FAO. 1993. 439p.*
- Taylor, I., Kathryn, G., Doering, L., Johnson, K.F., Wetzel, C.R., Stewart, I.J. 2021. Beyond visualizing catch-at-age models: Lessons learned from the r4ss package about software to support stock assessments, *Fisheries Research*, 239:105924. <https://doi.org/10.1016/j.fishres.2021.105924>.
- Urtizberea, A., Fu, D., Merino, Gorka., Methot, R., Cardinale, M., Winker, H., Walter, J., Murua, H. 2019. Preliminary Assessment of Indian Ocean Yellowfin Tuna 1950-2018 (Stock Synthesis, V3.30). IOTC-2019-WPTT21-50.
- Urtizberea, A., Cardinale, M., Winker, H., Methot, R., Fu, D., Kitakado, T., Fernández, C., Merino, G.. 2020. Towards Providing Scientific Advice for Indian Ocean Yellowfin in 2020. IOTC-2020-WPTT22(AS)-21
- Winker, H., Carvalho, F., Cardinale, M., Kell, K. 2020. ss3diags. R package version 1.0.2.
- Winker, H., Walter, J., Cardinale, M., Fu, D. 2019. A multivariate lognormal Monte-Carlo approach

for estimating structural uncertainty about the stock status and future projections for Indian Ocean Yellowfin tuna. IOTC–2019–WPTT21–xx.

Walter, J., Hiroki, Y., Satoh, K., Matsumoto, T., Winker, H., Ijurco, A.U., Schirripa, M., 2019. Atlantic bigeye tuna stock synthesis projections and kobe 2 matrices. Col. Vol. Sci. Pap. ICCAT 75, 2283–2300.

Walter, J., Winker, H., 2019. Projections to create Kobe 2 Strategy Matrices using the multivariate log-normal approximation for Atlantic yellowfin tuna. ICCAT-SCRS/2019/145 1–12.

Zudair, I., Murua, H., Grande, M., Bodin, N. 2013. Reproductive potential of yellowfin tuna (*Thunnus albacares*) in the western Indian Ocean. Fish. Bull. 111:252–264.

APPENDIX A: REVISEED MODEL FISHERY DEFINTION

Table A1: Definition of fisheries for yellowfin tuan assessment for the revised four-region spatial struture used in the revised model.

Fishery	Gear	Region
1. GI 1a	Gillnet	1a
2. HD 1a	Handline	1a
3. LL 1a	Longline (distant water)	1a
4. OT 1a	Other	1a
5. BB 1b	Baitboat	1b
6. PS FS 1b (small)	Purse seine, school sets (≤ 80 cm)	1b
7. LL 1b	Longline (distant water)	1b
8. PS LS 1b (small)	Purse seine, log/FAD sets (≤ 80 cm)	1b
9. OT 2	Other	2
10. LL 2	Longline (distant water)	2
11. LL 3	Longline (distant water)	3
12. GI 4	Gillnet	4
13. LL 4	Longline (distant water)	4
14. OT 4	Other	4
15. LF 4	Longline (fresh tuna)	4
16. PS FS 1b (large)	Purse seine, school sets (> 80 cm)	1b
17. PS LS 1b (large)	Purse seine, log/FAD sets (> 80 cm)	1b

APPENDIX B: RESULTS FROM THE EXPLORATORY MODELLING

Table B1. Maximum Posterior Density (MPD) estimates of the main stock status indicators from the exploratory model options.

Option	SB_0	SB_{MSY}	SB_{MSY}/SB_0	SB_{2020}	SB_{2020}/SB_0	SB_{2020}/SB_{MSY}	F_{2020}/F_{MSY}	MSY
Basic	3,323,090	1,229,350	0.37	908,149	0.27	0.74	1.15	427,332
bRecruit4	3,559,830	1,354,250	0.38	994,760	0.28	0.73	1.06	457,924
bLFdw	3,107,400	1,147,140	0.37	820,080	0.26	0.71	1.24	417,688
bLFSplit	3,238,060	1,190,030	0.37	874,748	0.27	0.74	1.19	418,464
bSelrw	2,879,610	1,046,010	0.36	1,329,565	0.46	1.27	0.80	391,415
bSeason	3,326,470	1,234,770	0.37	914,238	0.27	0.74	1.15	426,732
b1990	3,115,630	1,153,770	0.37	941,067	0.30	0.82	1.16	403,816
bLL1bSplitQ	3,312,670	1,177,490	0.36	1,121,743	0.34	0.95	0.86	400,296
revised	3,952,880	1,529,270	0.39	1,193,405	0.30	0.78	0.78	483,596
rCPUEFSq	3,848,370	1,511,790	0.39	1,135,360	0.30	0.75	0.94	481,736
rCPUEFSpe	3,842,650	1,472,640	0.38	1,115,803	0.29	0.76	0.92	471,004
rCPUEFSLSBBpe	3,808,060	1,444,200	0.38	1,180,283	0.31	0.82	0.86	448,248
bDortel	3,018,180	1,102,890	0.37	721,447	0.24	0.65	1.62	389,740
bMlow	4,846,930	1,758,630	0.36	1,065,512	0.22	0.61	1.89	358,372
bMAmax10	4,373,900	1,631,820	0.37	1,033,041	0.24	0.19	1.68	376,341
bMAmax18	6,643,360	2,386,150	0.36	1,376,545	0.21	0.17	2.23	342,048

Table B2: Details of objective function components for the exploratory model options.

	TOTAL	CPUE	Length_comp	Tag_comp	Tag_negbin	Recruitment	Parm_priors	Parm_devs	Catch
Basic	9639.71	-116.47	3996.11	4046.27	1818.03	-54.37	46.51	18.93	0.01
bLFdw	9191.93	-237.35	3946.65	4008.42	1830.74	-55.51	49.18	62.78	0.00
bRecruit4	9446.54	-132.60	3605.93	4039.16	1800.37	-50.17	48.16	17.66	0.07
bLFSplit	9559.37	-203.50	3997.76	4037.94	1825.52	-40.17	66.67	18.01	0.08
bSelrw	9174.95	-538.42	3932.05	4033.15	1776.31	-50.41	71.82	76.19	0.00
bSeason	9491.14	-244.91	3970.97	4052.12	1818.29	-52.29	58.11	16.12	0.00
b1990	9709.68	-53.77	4022.03	4048.22	1819.31	-27.15	44.63	9.90	0.00
bLL1bSplitQ	9603.95	-119.93	3975.96	4035.26	1841.99	-55.46	49.84	20.26	0.00
revised	8224.50	-148.18	3427.29	3040.33	1739.44	-50.29	310.88	23.16	0.03
rCPUEFSq	8344.05	-47.72	3459.11	3041.88	1745.14	-51.15	303.17	24.35	0.00
rCPUEFSpe	8158.00	-201.14	3475.46	3039.92	1742.14	-53.31	318.31	23.55	0.00
rCPUEFSLSBBpe	8164.06	-243.47	3424.15	3040.71	1750.65	-53.51	303.04	24.86	0.00
bDortel	9770.92	-96.74	4146.76	4029.98	1791.21	-37.70	51.44	16.47	0.35
bMlow	9700.95	-100.06	4104.89	4074.44	1735.72	-45.47	58.30	19.77	0.01
bMAMax10	9696.24	-99.15	3982.59	4056.69	1733.76	-52.57	55.86	18.83	0.00
bMAMax18	9975.75	-44.92	4155.56	4087.63	1743.80	-36.57	53.56	16.37	0.00

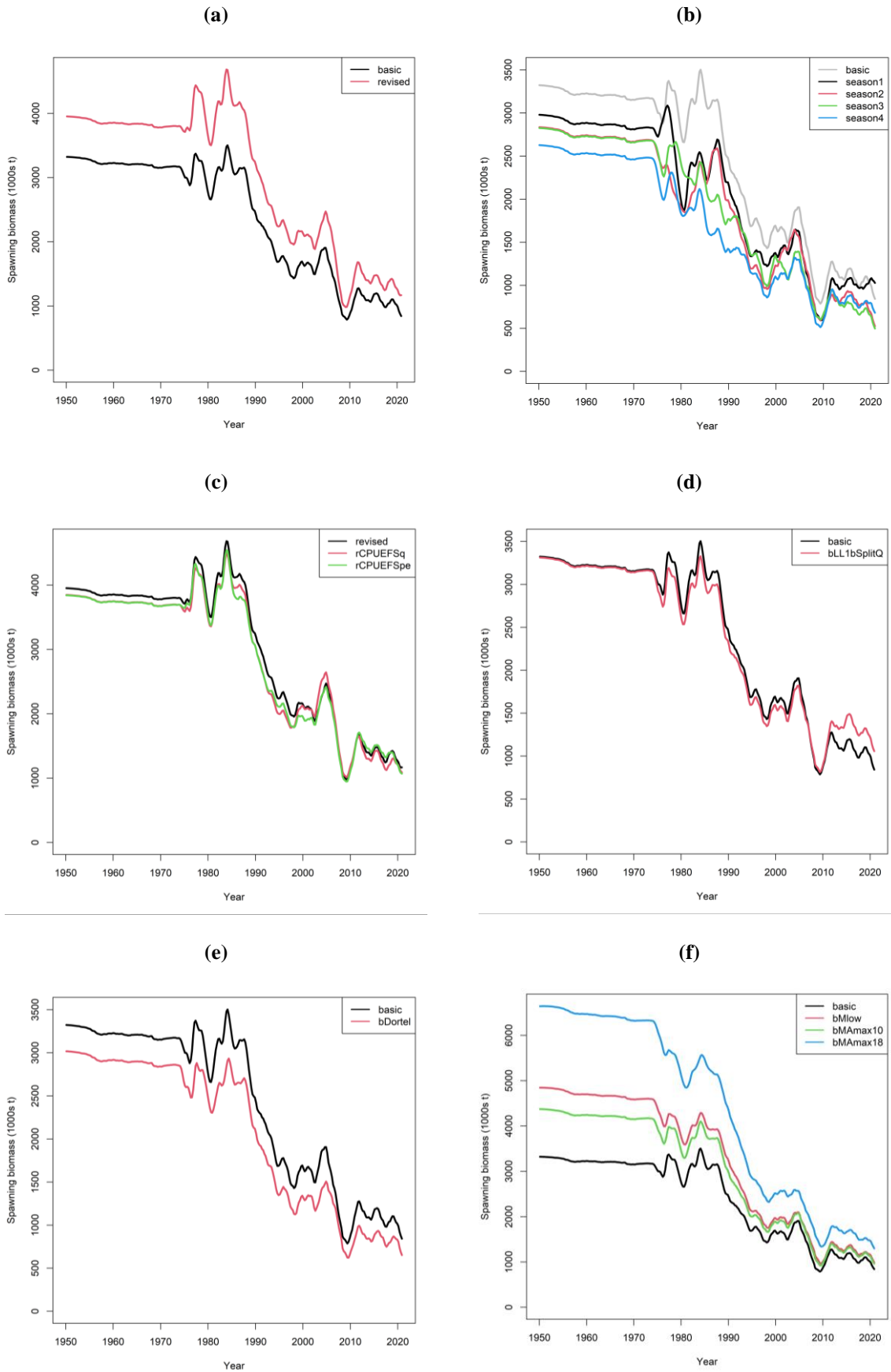


Figure B1: A comparison of estimated spawning biomass between the basic model and selected sensitivity models.

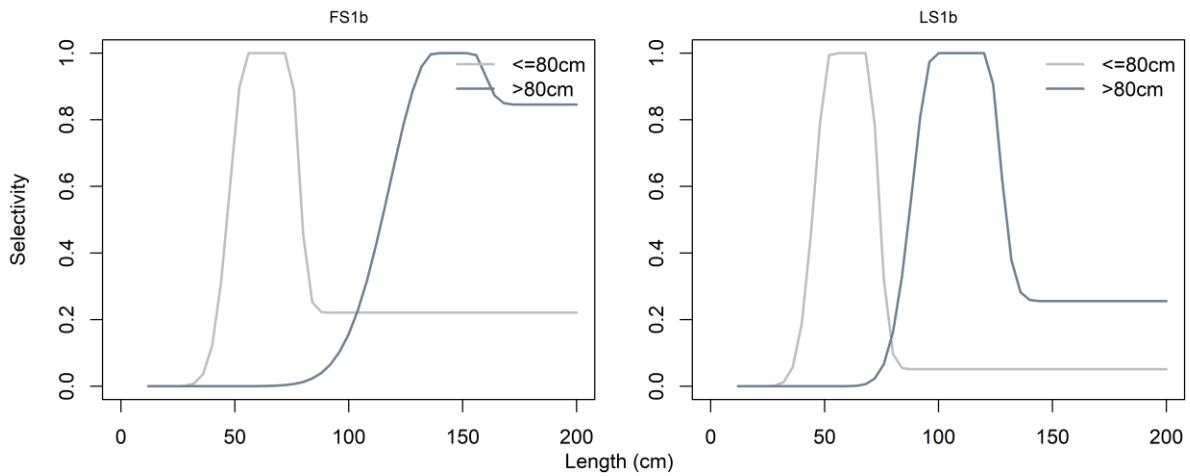


Figure B2: Estimated selectivity for PSFS fisheries 6 (≤ 80 cm) and 16 (> 80 cm), and PSLs fisheries 8 (≤ 80 cm) and 17 (> 80 cm) for the revised model.

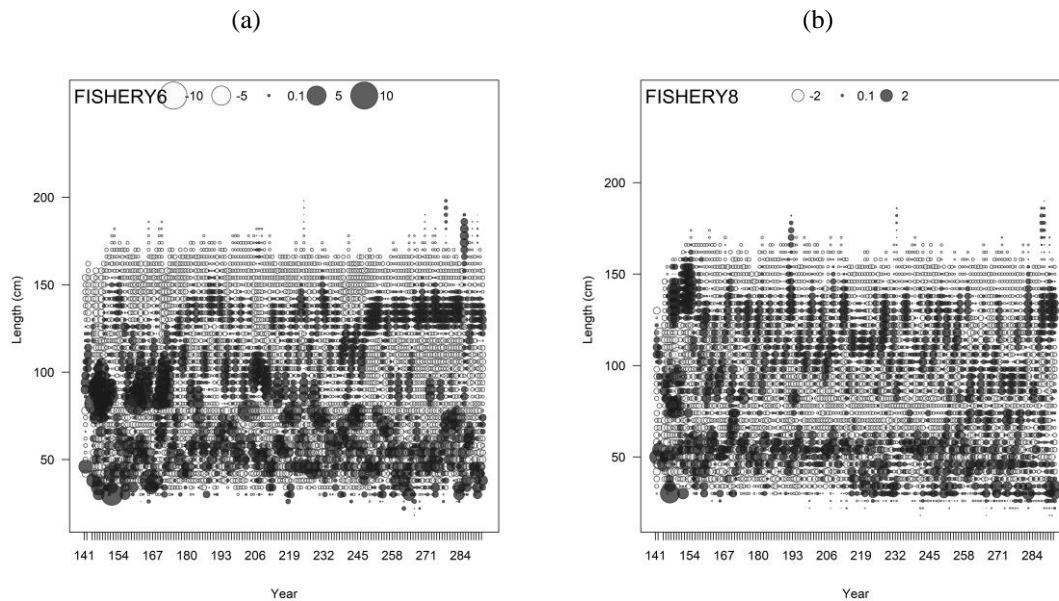


Figure B3: Relative residuals from the fits to the length compositions for (a) PSFS 6 and 16 (combined to show as Fishery 6), and (b) PSLs fisheries 8 and 17 (combined to shows as Fishery 8) for the revised model. Noting that the LF for the small length mode (≤ 80 cm, Fishery 6 and 8), and the large length mode (> 80 cm, Fishery 16 and 17) are fitted independently.

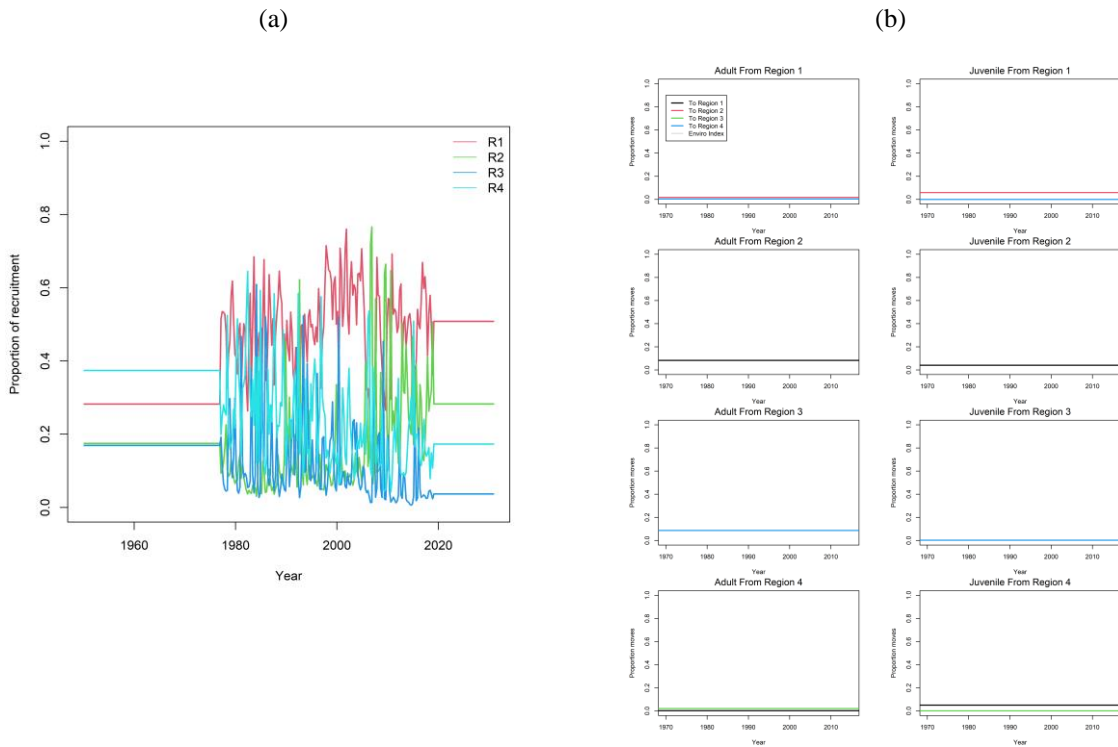


Figure B4: Diagnostics for model *rRecuit4* (a): proportion of the total quarterly recruitment assigned to each model region; (b) estimated age specific movement parameters.

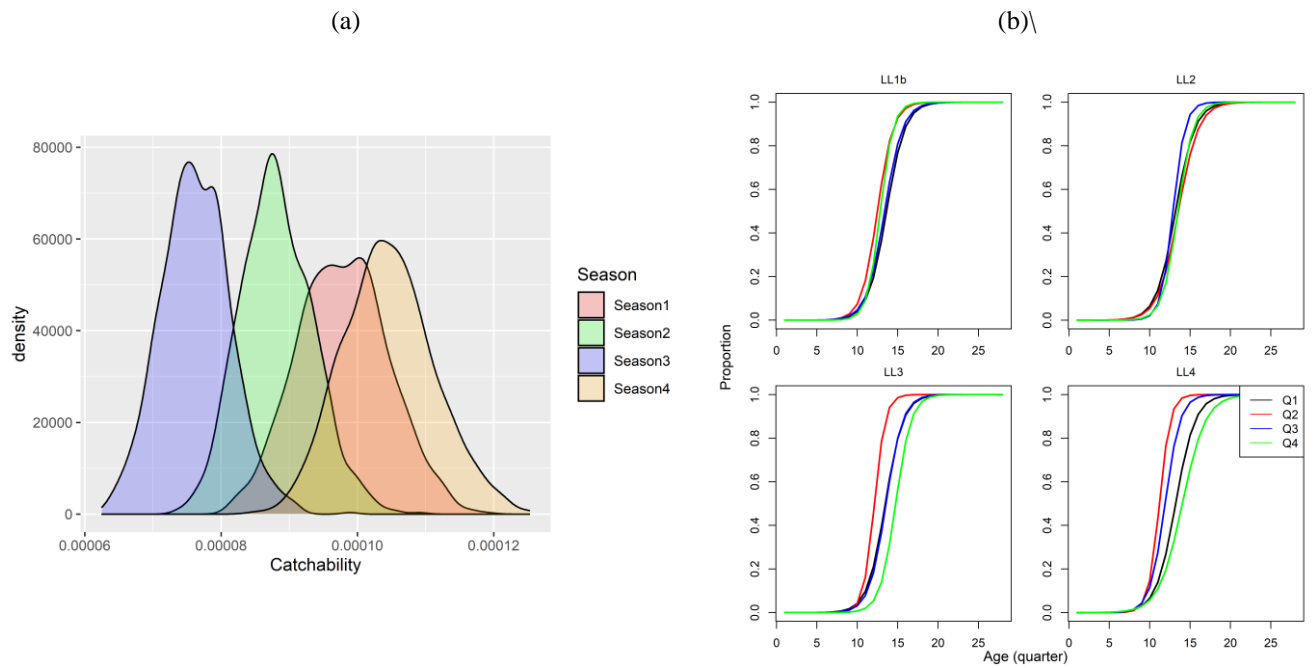


Figure B5: Diagnostics for model *bSeason*: (a) estimated seasonal LL CPUE catchabilities; (b) estimated selectivity for each quarter for each of the main longline fisheries LL1b, LL2, LL3, and LL 4.

(a)

(b)

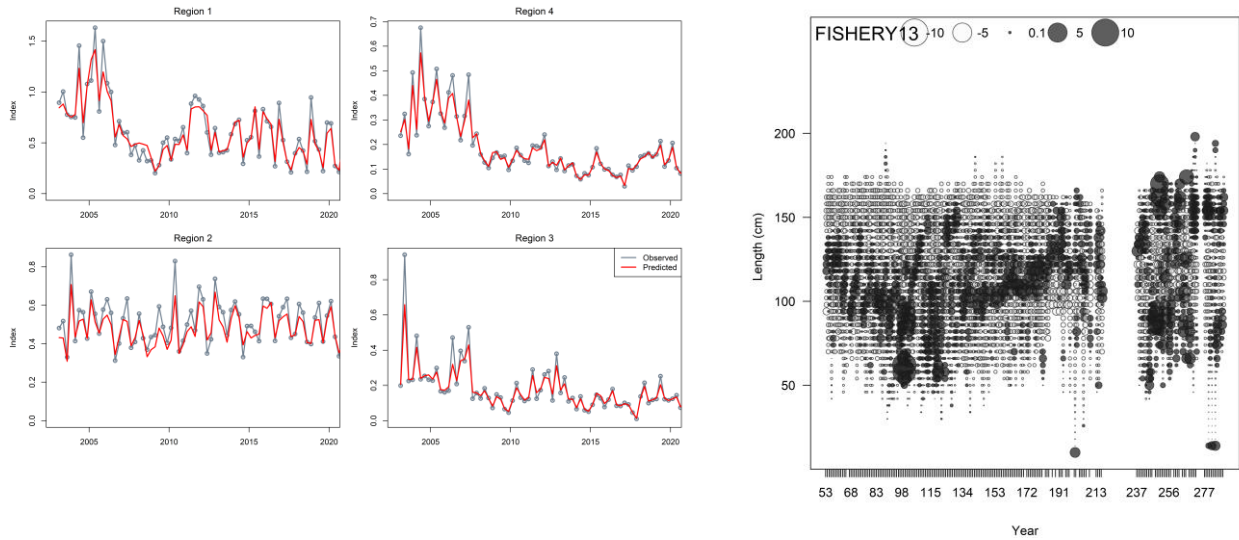


Figure B6: Diagnostics for model *bSelrw*: (a) fits to the LL CPUE indices 2003–2020; (b) relative residuals from the fits to the length compositions for LL4 (Fishery 13).

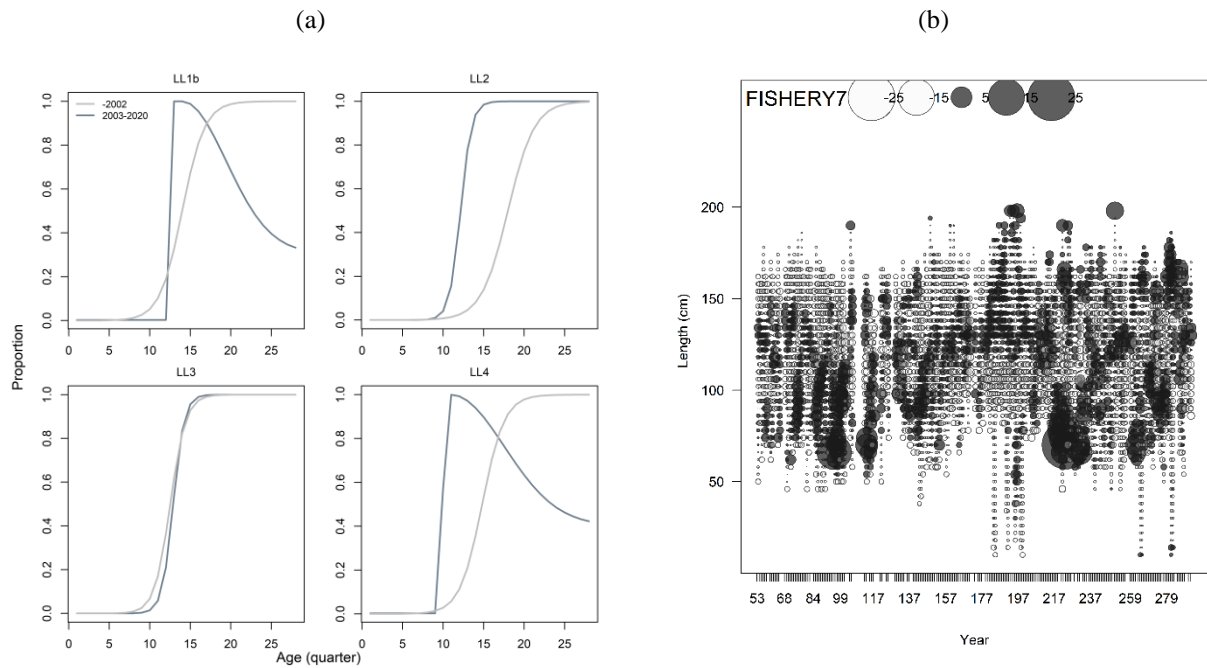


Figure B7: Diagnostics for model *bLFsplit*: (a) estimated selectivity for the main longline fisheries LL1b, LL2, LL3, and LL 4, where each fishery was divided into an early period 1950 – 2002, and 2003 – 2020; (b) relative residuals from the fits to the length compositions for LL1b (Fishery 7).

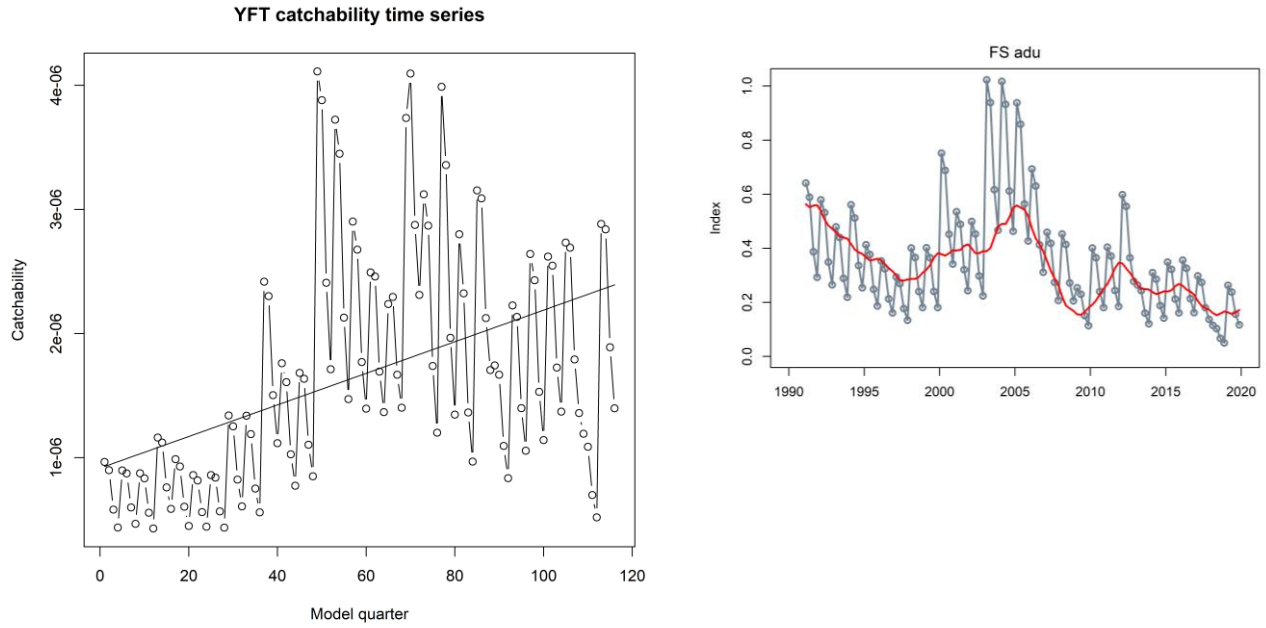


Figure B8: Time-varying catchability estimates for the yellowfin PSFS 1b adult CPUE indices (left - The line linear regression fit, and corresponds to a 1.305 % per year trend, compounded annually over a 32 year period), and the fits to the CPUE indices adjusted for catchability change (right).

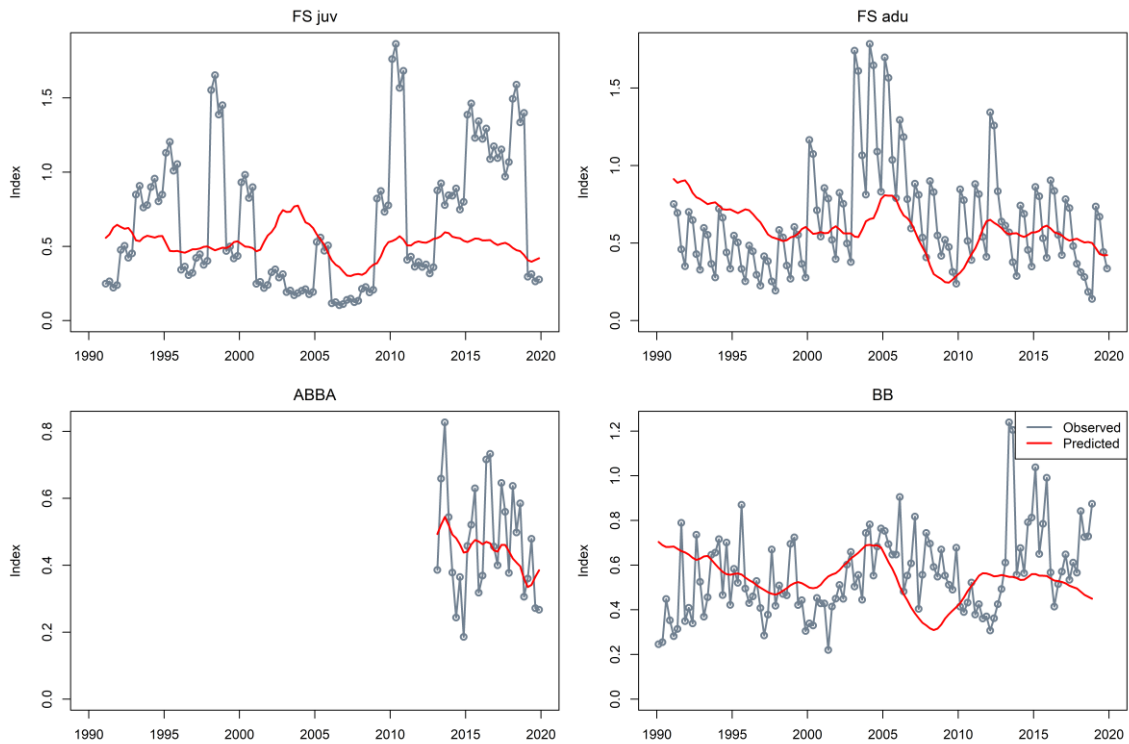


Figure B9: Fit to the CPUE indices from model rCPUEFSLBB, including the PSFS juvenile, PSFS adult, ABBA, and baitboat indices.

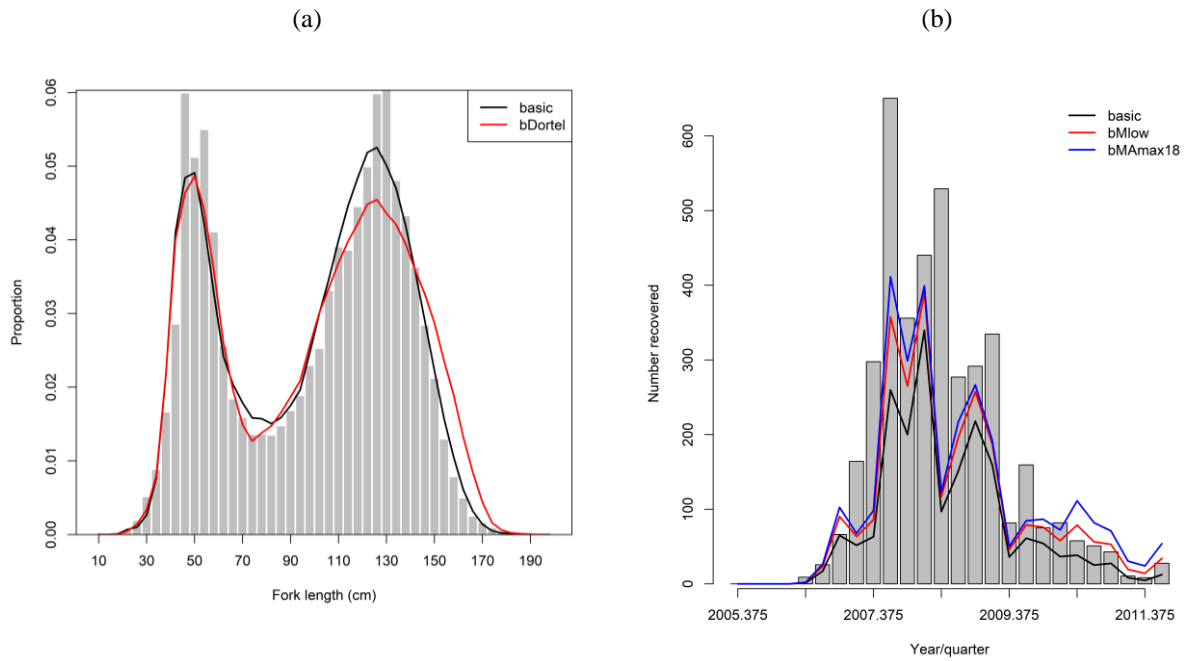


Figure B10: Diagnostics for models with alternative growth or M (a): fits to the LF aggregated for all fisheries and all years; (b) fits to the tag recoveries over time aggregated for all fisheries.

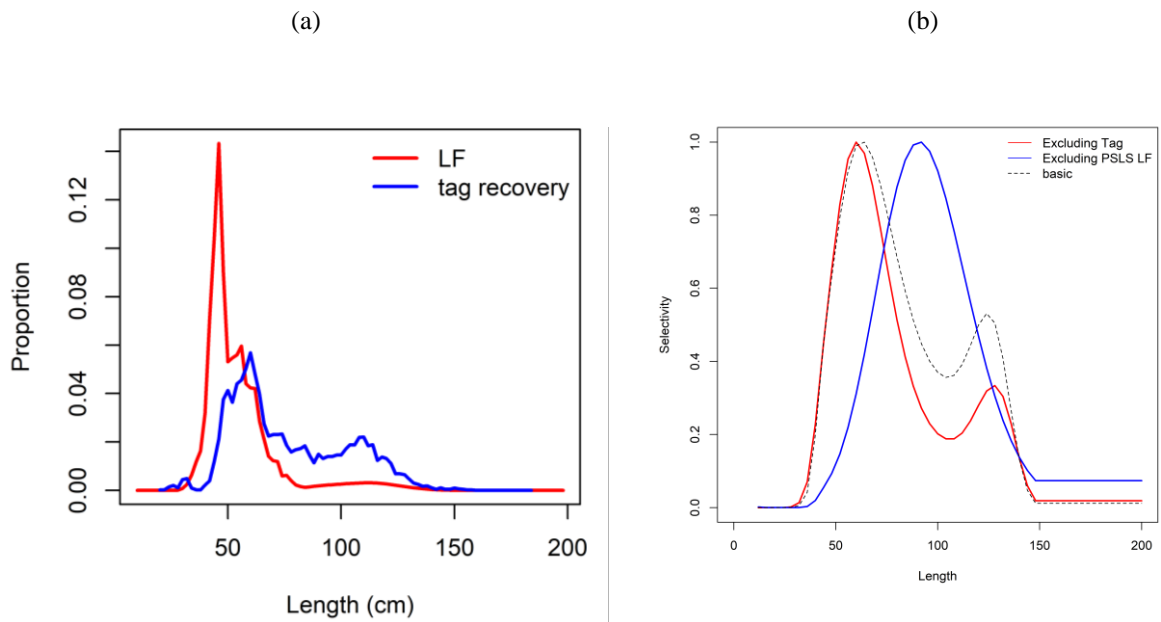


Figure B11: Diagnostics showing the conflict between tagging and LF data from the PSLs fishery: (a) aggregated length distribution from the catch samples and from tag recoveries by the PSLs fishery; (b) estimated PSLs selectivity when tagging data and LF data are excluded respectively

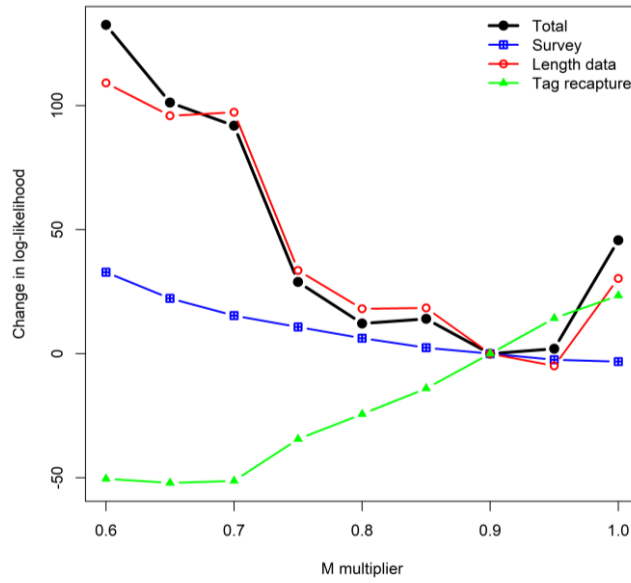


Figure B12: Likelihood profile on natural morality: Total and component likelihood. The natural morality is expressed as a multiplier of the base level natural mortality.

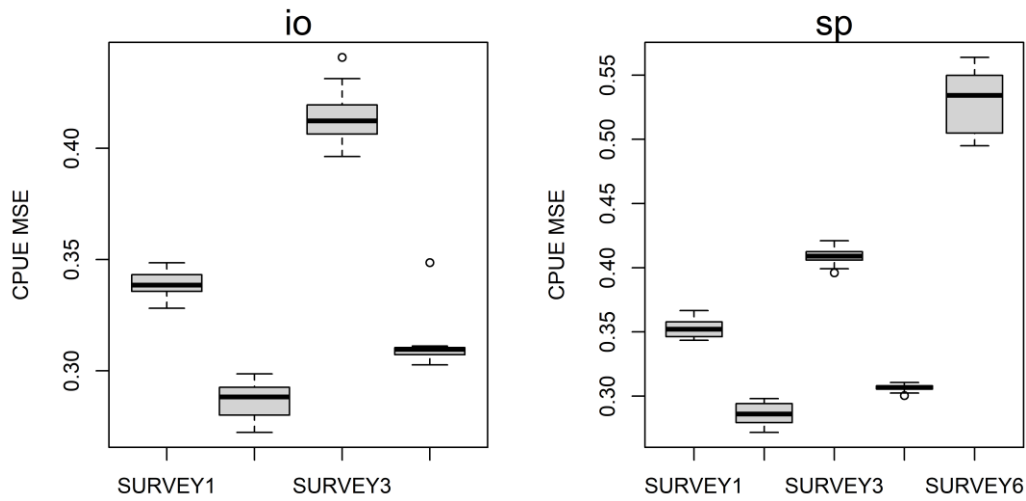


Figure B13: Summary distribution of the Mean Squared Error (MSE) from fits to CPUE indices for the final model ensemble: for models based on the basic spatial structure (left, LL CPUE 1 – 4), and for models based on the revised spatial structure (right, LL CPUE 1 – 4, and PS CPUE on adults).