Indian Ocean Tuna Commission
Commission des Thons de l'Ocean Indien

IOTC–2022–WPTmT08-09

Stock assessment of albacore tuna (*Thunnus alalunga*) in the Indian Ocean using Stock Synthesis.

8th IOTC Working Party on Temperate Tuna
25 – 29 July 2022
Virtual Meeting

Joel Rice[1]

[1] Joel Rice Consulting. (ricemarineanalytics@gmail.com)

Executive summary

This paper presents a stock assessment of albacore tuna in the Indian Ocean using Stock Synthesis (version 3.30.19.01 http://nft.nefsc.noaa.gov/Download.html). The albacore tuna assessment model is an age structured (14 years), spatially aggregated (1 region) and two sex model. The catch, effort, and size composition of catch, are grouped into 23 fisheries covering the time period from 1950 through 2020. Fifteen indices of abundance, fourteen of which are from longline fisheries were considered for this analysis. The estimated abundance trend is decreasing throughout the time frame of the model, and spawning stock abundance has decreased to approximately 2 times SSBMSY The fishing mortality has increased over the model time frame with F2020/FMSY= 0.6.

Albacore tuna are most often caught in long line fisheries in the Indian Ocean tuna fisheries, though some bycatch occurs in the purse seine fisheries as well as other mixed gear fisheries. This analysis was developed based on the 2019 assessment along with updates to the data and parameterization. A the diagnostic case, is referred to in the main text when presenting the model parametrization and diagnostics. The upcoming 8th meeting of the Indian Ocean Tuna Commission Working Party on Temperate Tuna and Bycatch (WPTmT08) will recommend the final parameterization as a base case model for the provision of stock status. Initial analysis based on the sensitivity analysis done with SS3 indicated that the stock is not over fished nor experiencing overfishing.

1 Introduction

Commercial fisheries for albacore tunas have operated in the Indian Ocean since the early 1950s. The earliest known exploitation was by the Japanese longline fishery in the 1950s, followed by the Korean and Taiwanese longline fisheries in the mid and late 1950s (Kim et al. 2010; Chen 2009). Drift nets were employed in the albacore fishery from the early-1980s until 1992 when an international ban on drift net fishing came into force. Taiwanese and Indonesian longline catch has recently accounted for around 70% of the total catch. Between 2008 and 2011, following the onset of piracy in waters off Somalia, part of the longline fleets that had traditionally targeted tropical tunas or swordfish in those waters moved towards albacore fishing grounds in the southern part of the Eastern Indian Ocean.

Like albacore fisheries in other oceans, the Indian Ocean fishery is characterized by smaller fish at higher latitudes. Unlike other oceans however, there is no significant troll or pole and line fishery for albacore, and since the ban on the drift net fishery there has been no large-scale targeting of small fish.

Stock assessments of the Indian Ocean albacore stock have been conducted in the past using several different methods, including the non-equilibrium production model ASPIC (Chang et al. 2012, Matsumoto et al. 2012, Matsumoto et al. 2014, Matsumoto 2016), the age-structured production model ASPM (Nishida et al. 2012, Nishida et al. 2016), a Bayesian biomass dynamic model (Guan et al. 2016), and Stock Synthesis (SS) (Kitakado et al. 2012, Hoyle et al. 2014, Langley & Hoyle 2016, Langley 2019).

The most recent (2019) assessment using Stock Synthesis incorporated data to 2017 (Langley 2019). The assessment incorporated longline CPUE indices for albacore from a collaborative project between Japan, Taiwan, Korea, and the IOTC, similar to the work done in 2016 (Hoyle et al 2019. Hoyle et al. 2015, Hoyle et al. 2016). The standardized CPUE indices were derived from operational-level longline data from the three fleets, incorporating cluster analyses to address

the effects of target change and standardization models including vessel effects and spatial effects. The project built on a study applying similar methods to Indian Ocean bigeye and yellowfin tuna (Hoyle et al. 2015). The 2019 assessment investigated the interactions between the key data sets (CPUE indices and length composition data) and the uncertainty associated with the key biological parameters (natural mortality and SRR steepness). The 2019 model used the fleets as areas approach with four longline fleets, and four corresponding CPUE series (LL1 – LL4). The CPUE in the southwest area are mostly likely to represent the abundance of albacore tuna in the Indian Ocean, as the indices were based on a main target fishery with more consistent fishing operations. The southwest area also represents a significant proportion of the albacore biomass in the Indian Ocean. The WPTmT 07 assessment meeting noted that regarding the assessment results:

i. There are conflicts between LL3 CPUE indices and LF data in the southern area (LL3 and LL4).

ii. Reducing weight on the LL3 and LL4 LF data results in a shift in the selectivity to improve fits to LL3 and LL4 CPUE.

iii. The PS LF data are still influential in determining stock status when LL LF data are removed.

iv. Conflict between LL1 CPUE and LL3 CPUE not resolved but separate model fits to LL1 and LL3 CPUE yield similar estimates of stock status when LL LF data are substantially down weighted

Based on the assessment results the WPTmT 07 adopted a subset of the SS model scenarios for the determination of the stock status of Indian Ocean albacore and the formulation of management advice for 2019. These are

i. Model 1 - CPUE-Northwest, LL and PS LF included

ii. Model 2 - CPUE-Southwest, LL and PS LF included

iii. Model 3 - CPUE-Northwest, LL and PS LF excluded (selectivity fixed at values from initial fit)

WPTmT7 concluded that " catches in 2017 were marginally above the MSY level of the SS3 model. Fishing mortality represented as F_{2017}/F_{MSY} is 1.346 (0.588–2.171). Biomass is estimated to be above the SBMSY level (1.281 (0.574–2.071)). These changes in stock status since the previous assessment are possibly due to decreases in the CPUE in recent years. the stock status in relation to the Commission's BMSY and FMSY target reference points indicates that the stock is not overfished but is subject to overfishing.

In advance of the current assessment, a preparatory meeting was held in April 2022 to compile the data sets for the assessment (IOTC 2022). The meeting reviewed the updated CPUE indices derived from operational data from the Indian Ocean longline fleets (Kitakado, et al. 2022), recent biological studies and updated, catch data. The meeting also specified the details of the assessment modelling requested during the WPTmT 08 assessment meeting. This assessment has been built around Model 2 - CPUE-Southwest, with longline and purse seine length frequency included.

This report presents the preliminary results of the 2022 stock assessment modelling of Indian Ocean albacore tuna using Stock Synthesis (Version 3.30.19_01). The assessment results will be finalized at the WPTmT8 meeting.

2 Methods

Data

There are many different fleets catching albacore tuna in the Indian Ocean, with the main fleets consisting of longline fleets from distant water fishing nations. The data used in the albacore tuna assessment consist of fishery specific catch and length composition data along with standardised longline CPUE indices. The details of the configuration of the fishery specific data sets are described below.

There is enough uncertainty about the selectivity assumptions with respect to time, and the interannual variability with respect to the numbers of size composition data, that the size

composition data are not expected to be very informative about year-class strength. Hence, in the assessment presented here, the length-composition data are down weighted so as to inform the selectivity and recruitment but not alter the model fit to the abundance trend.

2.1 Spatial stratification

The 2016 assessment partitioned the Indian Ocean into four quadrants demarcated at 25°S latitude and 75°E longitude (Langley & Hoyle 2016). The spatial stratification was primarily adopted to partition the longline fisheries by the size of fish caught; the longline fisheries in the southern area tend to catch albacore that are considerably smaller than the fisheries in the northern area (Chen et al. 2004, Geehan & Hoyle 2013, Nikolic et al 2013). Higher longline CPUE for albacore has been associated with the North Subtropical Front in the southern Indian Ocean (30–35°S latitude) where SST was 15–19°C (Lan et al 2011). There is no indication of a longitudinal trend in the size of albacore caught by the longline fisheries (Geehan & Hoyle 2013) and the overall distribution of the southern longline fishery is continuous throughout the southern region (Figure 1). However, during the history of the fishery there were periods when there was an appreciable separation between the operation of the longline fishery in the southwestern and south-eastern quadrants of the Indian Ocean, most notably during the 1960s and 1970s (Figure 2). To account for potential longitudinal variation in the key fishery data sets, the northern and southern areas of the Indian Ocean were partitioned at 75°E longitude. An investigation of the stock structure of Indian Ocean albacore by morphometric and DNA sequence methods categorised samples into two major groups partitioned by 90°E longitude suggesting that there may be two albacore stocks in the Indian Ocean (Yeh et al. 1995). Thus, the spatial stratification of the assessment data sets can be applied to approximate the more complex stock structure within the Indian Ocean.

The four regions of the Indian Ocean were used to define the spatial domain of the model fisheries and define the region specific longline data sets for the CPUE analysis (Hoyle et al 2016 and 2019, Kitado et al 2022). There are apparent differences in the trends in albacore CPUE indices between the two southern areas over the last decade.

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. The model was further stratified by quarter of the calendar year (Jan-Mar, Apr-Jun, Jul-Sep, Oct-Dec) and the various data sets were compiled accordingly.

2.3 Definition of fisheries

The spatial stratification was applied to define model fisheries based on region and fishing gear. These "fisheries" are considered to represent relatively homogeneous fishing units, with similar selectivity and catchability characteristics. As an update for this assessment the single aggregate longline fisheries by region have been separated into quarterly fisheries, thus there are sixteen longline fisheries, as opposed the previous assessment which used four.

A total of twenty-three fisheries were defined, including 16 longline fisheries (one per quarter per region), two driftnet fisheries, one purse seine fishery and four 'other' fisheries that account for the troll, sport fish, and miscellaneous fisheries by region (Table 2).

Longline northwest (LL1). The composite longline fishery in the north-western region was initially developed by the Japanese fleet in the mid 1950s. Data for the Taiwanese fleet are available from the late 1960s and the fleet has operated continuously since then. Japanese fishing effort declined in the early 1970s and remained relatively low until the mid 1990s. Fishing effort by the Japanese fleet recovered to a moderate level during the late 1990s and 2000s but was at a relatively low level from 2010. The composite longline fishery included effort by the Korean longline fleet during the late 1970s and 1980s. Limited albacore catches have been taken by other fleets, most notably the Chinese longline fleet operating during the last decade.

Longline northeast (LL2). The composite longline fishery in the north-eastern region has a similar composition to the LL1 fishery. The Japanese fleet was the dominant component of the fishery during 1953-1970 and continued to operate at a lower level over the subsequent years.

Data for the Taiwanese fleet's operations in the area are available from the late 1960s, while the Korean fleet primarily operated in the fishery during 1976-1987. Longline southwest (LL3) . The composite longline fishery in the south-western region was dominated by the Japanese fleet from the inception of the fishery in the late 1950s until the introduction of the Taiwanese fleet in the mid-1960s. Since the early 1970s, most of the catch was taken by the Taiwanese fleet. There was a short period of higher Japanese catch during 2004-2008.

Longline southeast (LL4) . The composite longline fishery in the south-eastern region was dominated by the Japanese fleet from the inception of the fishery in the late 1950s until the introduction of the Taiwanese fleet in the early 1970s. During 1974–2006, most of the catch was taken by the Taiwanese fleet. Japanese longline catches increased from 2006 and in recent years considerable catches have also been taken by China and Korea.

Drift net fisheries. Drift net fisheries were defined for the south-western (DN3) and south-eastern (DN4) regions. These fisheries were comprised exclusively of drift net vessels flagged to Taiwan China, which operated in the southern waters of the Indian Ocean from 1982 until 1992 when the UN adopted a worldwide ban on drift nets.

Purse seine. A single purse seine fishery (PS1) was defined as virtually all purse seine catch was taken within the north-western region. The purse seine fishery is made up of various fleets although the majority of the catches of albacore are reported by purse seiners flagged to the European Union and other fleets under EU ownership, including the Seychelles (86% of the total catches of albacore over the time series). The purse seine catches of Iran, Japan, Mauritius, Thailand, and the Republic of Korea are also included in the fishery.

Other. A miscellaneous ("Other") fishery was defined for each of the regions. The "Other" fisheries include various coastal longline, gillnet, trolling, hand lines and other minor artisanal gears, which are used in coastal countries of the Indian Ocean. Collectively, the "Other" fisheries account for a small proportion of the Indian Ocean albacore catch (2% of the entire

catch). Most of the catch was reported by Indonesia with the remainder of the catches reported by Mauritius, Reunion and Mayotte (EU), Comoros, Australia, South Africa, and East Timor.

2.4 Total catch

Catch data were compiled by IOTC secretariat based on the fishery definitions (IOTC-2022-WPTmT08(AS)-DATA12-SA_ALB_02_SS3.xlsx). All catches were expressed in metric tonnes (mt). There were minor changes in the gear specific catch histories from the previous assessment (Langley 2019). The most notable change was a redistribution of the annual longline catches amongst the model regions due to a revision of the spatial catches from the longline fleet

The longline fisheries for albacore developed from the mid 1950s and total annual catches averaged about 17,000 t during 1960-1980 (Figure 3). Catches increased during the period of drift net fishing in the late 1980s and early 1990s and continued to increase with the expansion of the longline fishery in the late 1990s to reach a peak in catch of 40-45,000 mt in 1998-2001. Total catches declined to about 30,000 t in 2003-2006 and have since fluctuated between 33-43,000 t per annum (Figure 3). During the last decade, the longline fisheries have accounted for 95% of the total albacore catch apportioned amongst the regions as follows: northwest 23%, northeast 13%, southwest 34% and southeast 30%. Most of the catch from the southern longline fisheries occurs during the 2nd and 3rd quarter of the year, while catches from the north-western longline fishery are predominantly taken in the 4th quarter.

Longline in the northwest region (LL1). Fishing Fleets 1-4. The longline catch was relatively low from the north-western region during the late 1980s and early 1990s immediately following the reduction in fishing by the Korean fleet and during a period of low catch by the Japanese fleet. The high catches in 2000 and 2001 were predominantly attributable to higher catches by the Taiwanese fleet in those years (Figure 3). Annual catches averaged about 8,000 t per annum over approximately the last decade (2010-2020), with the average catch by quarter being 2000, 287, 600, and 5030 MT for quarters 1-4 respectively.

Longline in the northeast region (LL2) (Fishing fleets 5-8) The higher catches from the fishery in the late 2000s (2007 and 2008) are primarily attributable to higher catches by the Taiwanese fleet in 2007 and 2008 and the Indonesian longline fleet during 2009 and 2010. Catches were generally declining from the early 2000's to the approximately 2015, after which they increased in all quarters (Figure 2), and averaged 1235, 1257, 1912, and 1000 MT .by quarter (1-4 respectively) over 2010-2020.

Longline in the southwest region (LL3) Fishing fleets 9-12. The high catches from the fishery during 1996-2002 are primarily attributable to a considerable increase in catch by the Taiwanese fleet during that period (Figure 2). Annual catches dropped markedly in 2003 and the catch from the Taiwanese fleet remained relatively low during 2003-2005. Annual catches increased steadily since about 2006 , most notably in region 2. Quarterly catches averaged 854,7068, 5735 and 692 by quarter (1-4 respectively), with highest level of catch of about 19,000 MT in 2019, followed by 17,000 MT in 2020.

Longline southeast region (LL4), Fishing Fleets 13-16. Longline catches from the south-eastern region reached an historically high level of about 10,000 t in 2014. The higher level of catch around this time was primarily attributable to an increase in catch by the Japanese fleet since 2006. Annual catches have declined in more recent years and were about 4,000 t in 2019 and 2020. Average quarterly catches are about 300, 4,000, 2360, and 193 MT for quarters 1-4 respectively.

2.5 Relative abundance indices

Similar to the 2019 assessment, for each of the four regions, standardised CPUE indices for albacore tuna were derived using generalized linear models (GLM) from operational longline catch and effort data provided by Japan, Korea, and Taiwan, China (Kitakado et al 2022). Contrary to the 2019 assessment data from the Seychelles was not used. Cluster analyses of species composition data by vessel-month for each fleet were used to separate datasets into

fisheries that are believed to target different species or species compositions. Selected clusters were then combined and standardized using generalized linear models. In addition to the year-quarter, models included covariates for vessel identity, 5 square location, effort and cluster. The analysis is a refinement of the approach used to derive the CPUE indices included in the 2016 and 2019 albacore stock assessment (Hoyle et al 2016, Hoyle et al. 2019).
Four sets of CPUE indices by year-quarter were derived based on different treatment of the fishing vessel variable in the CPUE modelling, resulting in one LL fleet by region (Kitakado et al 2020). These four regional CPUE series were then dis-aggregated by quarter, resulting in 14 CPUE series pertaining to quarters 1-4 for regions 1-3 and quarters 2 & 3 for region 4. This Models were run for the period 1975-2020, the early time period included in the previous assessment was not considered as previous assessments have highlighted that the large decline in the CPUE indices during the early period (1960- 1970) was not consistent with the relatively low catches taken from each of these regions during the period (Langley & Hoyle 2016, IOTC 2019).

CPUE indices for regions 1-3 incorporated the effects of target and effort in the delta-component and all the effects including the quarter-space interactions in the positive lognormal component. Due to instability with same model with data from region 4, the log-normal model was used an alternative approach. The study authors note that the diagnostics plots show some deviation from the normal distribution.

The CPUE indices from region 1 are characterized by a decline in the magnitude and variability of the indices through the timeseries (Figure 4). The CPUE indices from region 2 are characterized by higher and more variable CPUE in quarter 2, compared to the other quarters, of which quarters 3 and 4 are the most stable. The CPUE series from region 3 are similar to region 2 with higher and more variable CPUE in quarter two, with lower and less variable standardized values in quarter 3, 4, and to an extent in region 1. The high variability in the region 3 CPUE indices during the later part of the 1970s may have been primarily attributable to a shift in targeting behavior (Figure 4).

A short time series of annual Taiwanese drift net CPUE indices is available from 1985–1992 (Chang & Liu 1995) (Figure 5). The indices exhibit a high degree of interannual variability with high CPUE indices for 1987 and 1990. The reliability of the indices as an index of albacore abundance is unknown, although the high variability may provide some indication of variation in year class strength during the period given that the drift net fisheries typically catch a relatively narrow length range of albacore (corresponding to fish of about 1–2 years old). The annual indices were assumed to represent the relative abundance in the first quarter of the year (corresponding to the peak season of DN catch).

2.6 Size composition data

Longline fishery

Size frequency data are available for the Japan longline fishery from 1965. Length and weight data were collected from sampling aboard Japanese commercial, research and training vessels. Weight frequency data collected from the fleet (as live weight) have been converted to length frequency data via a weight-length key. Levels of sampling aboard the Japanese composite longline fleet over time were uneven in terms of both the sampling platform (commercial and non-commercial vessels) and sampling source (fishermen, scientists, observers). While in recent years the majority of the samples available have come from scientific observers on commercial vessels, in the past samples came from training and research vessels (scientists), and commercial vessels (fishermen).

Length frequency data from the Taiwanese longline fleet are also available from 1980. In recent years, length data are also available from other fleets and periods (e.g. Indonesia fresh-tuna longline, Seychelles, etc.).. Prior to the mid-2000s the length frequency data set is dominated by sampling from the Taiwanese deep-freezing longline fleet. Length samples from this component come from commercial vessels and include lengths recorded by fishermen and, to a lesser extent, lengths measured by scientific observers on some of those vessels, in recent years. A review of the Taiwanese length frequency data identified major differences in the

length frequencies of albacore recorded before and after 2003, with the majority of the smaller albacore missing from the length distributions since that year (Geehan and Hoyle 2013 and Hoyle et al 2021). Following the previous assessment concerns regarding the reliability of these data, all length samples collected from the Taiwanese
longline fleet via logbooks from 2003–2020 were excluded from the length composition data sets.

In the assessment modelling, the individual longline length frequency observations were assigned a relative weighting based on the number of fish measured, up to a maximum of 2000 fish (nfish). The Effective Sample Size (ESS) was determined from the number of fish divided by 400 (nfish/400) giving a maximum ESS of 5. The samples were assigned the relatively low ESS due to concerns regarding the reliability of the length samples from some key data sets.

Purse seine fishery

Purse seine fisheries catch adult albacore, as a bycatch (approximately 90-120cm), in the western central Indian Ocean. Albacore lengths are measured in port, by enumerators, during the unloading of purse seiners flagged in the EU and Seychelles. Length samples are available from the fishery from 1990-2020. The ESS of the individual samples was determined in the same manner as described for the longline fisheries.

Drift net fishery

Drift nets catch juvenile or sub-adult albacore (62-75 cm) (Figure 15). The 2019 assessment is the first time length composition data have been available from the Indian Ocean fishery. The data are from sampling of the Taiwanese drift net catch during 1985-1991. These data were recently provided to IOTC by T.
Nishida (IOTC 2019).

3 Model Assumptions

The most important model assumptions are described in the following sections. Standard population dynamics and statistical terms are described below, while equations can be found in Methot (2000, 2009). Attachment 1 is the template specification file for all of the models, and includes additional information on secondary elements of model formulation which may be

omitted in the description below. All of the specification files are archived with the IOTC Secretariat. Table 2 lists the assumptions for the sensitivity runs.

3.1 Software

The analysis was undertaken with Stock synthesis SS V3.30.19.01, 64 bit version (Methot 2000, 2009, executable available from http://nft.nefsc.noaa.gov/SS3.html), running on MS Windows™ 10). Typical function minimization of the fully disaggregated model on a 3.0 GHz personal computer required about 25 minutes. Additional simplifications and aggregations could probably reduce the minimization time further, without significant loss to the stock status inferences.

3.2 Population Dynamics

The assessment model was structured by sex and age with age classes of 0-13 years and an aggregate age class of 14+ fish. The model commences in 1950 at the start of the available catch history and continues to 2020. The initial population age structure was assumed to be in an unexploited, equilibrium state. Model years are partitioned into four quarterly seasons. A single spatial structures for the Indian Ocean albacore stock was used in conjunction with the quarterly disaggregated survey and fisheries, based on the partitioning of the Indian Ocean into four regions (Figure 1).

3.3 Biological inputs and assumptions

Sex Ratio

Across all oceans there are documented differences in the patterns of sex ratio at recruitment and at older ages and larger sizes, these patterns are likely to be caused by features of albacore biology. Males have been shown to grow considerably larger than females in the north Pacific (Chen et al. 2012), south Pacific (Williams et al. 2012), north Atlantic (Santiago and Arrizabalaga 2005), and Mediterranean (Megalofonou 2000). However there is no evidence for unbalanced sex ratio at the age of recruitment. Sex ratio at recruitment was assumed to be equivalent (1:1).

3.4 Growth

The standard assumptions made concerning age and growth in the SS model are (i) the lengths-at-age are assumed to be normally distributed for each age-class; (ii) the mean lengths-at-age are assumed to follow a von Bertalanffy growth curve. Following the previous assessment this study used the Farley et al. 2019 estimation of the age and growth of Indian Ocean albacore. The study sampled albacore from the western Indian Ocean, primarily from the longline fishery with smaller fish also sampled from pole-an-line and purse seine fisheries. The sampled fish ranged in length from 74 to 108 cm FL for females and 67 to 115 cm FL for males (Farley et al. 2019).

Reproductive potential of female albacore was assumed to be equivalent to south Pacific albacore maturity-at length (Farley et al. 2014) (Figure 7). This ogive takes into account sex ratio, sexual maturity, spawning fraction, and fecundity and so represents female reproductive output at length. The ogive has fish attaining sexual maturity at about 85 cm and 50% reproductive potential is reached at about 92 cm (Figure 7). When converted to age using the new Indian Ocean albacore growth curve (Farley et al. 2019), sexual maturity is attained at about age 4 years and 50% reproductive potential is reached at about 5 years. As an alternative, the maturity ogive derived for western Indian Ocean albacore (Dhurmeea et al. 2016) was also considered, although the ogive is considered less reliable due to the limited number of fish sampled from the smaller length classes (below L50)..

3.5 Natural mortality

Natural mortality was assumed to be 0.3 for both sexes, based on the previous assessment and the value applied in the north Pacific and the north Atlantic albacore stock assessments. Age specific natural mortality at age was investigated in a sensitivity analysis

3.6 Recruitment

The model partitions the population into 15 year-quarter age-classes in one region (Figure 1). The last age-class comprises a “plus group” in which mortality and other characteristics are assumed to be constant. The population is “monitored” in the model at year quarter time steps,

extending through a time window of 1950-2020. The main population dynamics processes are as follows:

Recruitment in the model occurs in the fourth quarter of each year, reflecting the summer spawning season (and the ageing protocol that assumed a birthday of 1 December Farley et al. 2019). Recruitment was based on a BH stock recruitment relationship (SRR) and annual deviates were estimated for the period of the model where there was the most data (1975−2020). Deviates from the SRR were given a small penalty, so that recruitment estimates in periods with less data were estimated closer to the mean. The applied penalty was based on the assumption that the true standard deviation of recruitment deviates (σR) is 0.3, reflecting the upper range of the magnitude in the. variation of recruitment deviates estimated during preliminary modelling. Imperfections in models and lack of full information in the data cause models to underestimate recruitment variability. Recruitment variability is assumed to be lognormally distributed, therefore mean recruitment is higher than median recruitment. Equilibrium recruitment is meant to represent the average recruitment through time, so the median value in the recruitment function must be bias-corrected upwards. Following Methot and Taylor (2011), the bias correction was adjusted across the time series according to the relationship between the assumed and estimated recruitment variability. recruitment deviates were estimated for 1975−2018.

The final model options included three (fixed) 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 tuna species such as albacore tuna and are routinely adopted in tuna assessments conducted by other tuna RFMOs.

3.7 Initial population state

In the previous assessment it was assumed that the albacore tuna population was at an unfished state of equilibrium at the start of the model (1950) with the beginning of longline fishing occurring in the following years (at least from the 1950s onwards), this assessment follows the same methodology.

3.8 Selectivity Curves and Fishing Mortality

Selectivity is fishery-specific and was assumed to be time-invariant. A double- normal functional (Method 2015) form was assumed for all selectivity curves. No sex-specific length data was available, all length data were aggregated. Length composition data was split by quarter and region, fishery F16_LL4_Q4 was mirrored to F8_LL2_Q4 due to poor fits in the initial modelling. A double normal selectivity was estimated for the purse seine fishery (PS1). The availability of length composition data from the drift net fishery enabled the estimation of a selectivity function for this method. The selectivity was parameterised as a double normal function and assumed to be equivalent for the two fisheries (DN3 & DN4).

No reliable length frequency data are available for the four "Other" fisheries. The selectivity of the "Other" fisheries was assumed to be equivalent to the selectivity of the drift net fishery. Initial modelling indicated that the model results were not sensitive to the selectivity assumed for the four Other fisheries due to the small magnitude of the catch associated with each of these fisheries. Length composition data was down weighted so that the data can inform removals by fisheries from the correct age class, and inform recruitment but not determine the scale or trend of the population.

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).

3.9 Parameter estimation and uncertainty

Model parameters were estimated by maximizing the log-likelihoods of the data plus the log of the probability density functions of the priors, and the normalized sum of the recruitment deviates estimated in the model. For the catch and the CPUE series we assumed lognormal likelihood functions while a multinomial was assumed for the size data. The maximization was performed by an efficient optimization using exact numerical derivatives with respect to the model parameters (Fournier et al. 2012). The Hessian matrix computed at the mode of the posterior distribution was used to obtain estimates of the covariance matrix. This was used in

combination with the Delta method to compute approximate confidence intervals for parameters of interest.

3.10 Profile Likelihood

An investigation of the information content in the data components was undertaken via the use of profile likelihood on the global scaling parameter (R0) (Lee et al 2014). The negative log likelihood of a specific parameter or data component should, in theory, decline to an obvious minimum. In situations where this does not happen, at least from one side, there may be insufficient information within the data to estimate other parameters. Virgin recruitment (R0) is an ideal scaling parameter because it is proportional to the unfished biomass. Profiles were run with the natural log of virgin recruitment, ln(R0), fixed at various values above and below the model estimated value; the corresponding likelihood profile quantified how much loss of fit was contributed by each data source. One of the primary uses of the likelihood profile is to identify conflicting data and provide a rationale for down weighting or excluding any data.

3.11 Hierarchical cluster analysis

A hierarchical cluster analysis (HCA) was used to identify groupings of CPUE series that represented similar, or same states of nature. The goal of this analysis was to develop a framework for identifying groupings of CPUE series that were similar, so that the model did not include trends that implied conflicting states of nature (i.e. increasing and decreasing). The methods were adapted from those recently implemented in an Atlantic shortfin mako assessment conducted by the International Commission for the Conservation of Atlantic Tunas (ICCAT 2017). As noted in the Atlantic shortfin mako assessment (ICCAT 2017), “it is not uncommon for CPUE indices to contain conflicting information. However, when CPUE indices are conflicting, including them in a single assessment (either explicitly or after combining them into a single index) tends to result in parameter estimates intermediate to what would be obtained from the data sets individually. Schnute and Hilborn (1993) showed the most likely parameter values are usually not intermediate but occur at one of the apparent extremes. Including conflicting indices in a stock assessment scenario may also result in residuals not

being identically and independently distributed (IID) and so procedures such as the bootstrap cannot be used to estimate parameter uncertainty. Consequently, when CPUEs with conflicting information are identified, an alternative is to assume that indices reflect hypotheses about states of nature and to run scenarios for single or sets of indices that represent a common hypothesis."

The HCA used methods conducted in R using FLR (http://www.flr-project.org/). and the *diags* package. FLR provides a set of common methods for reading these data into R, plotting and summarizing them to assess the consistency in the CPUE trends. The CPUE time series along with a lowess smoother fitted to CPUE each year using a general additive model (GAM) to compare trends for the CPUEs. It is important to note that the hierarchical cluster analysis is sensitive to the overlap in the time series as well as specific trends in the indices. The HCA use years in analysis of the CPUE series so that the trends would be comparable across indices. The results should be interpreted carefully, nevertheless the HCA identified possible groupings of time-series.

The first group identified by the HCA was characterized by time-series which were positively correlated with each other, LLCPUE4_Q2, LLCPUE4_Q3, LL_CPUE3_Q2, LLCPUE3_Q3. Which corresponds to the Southern areas in the austral winter. The second group was region 3 where the CPUEs for quarters 1-4 were positively correlated, as were the four CPUE for region1, though two regions were not positively correlated.

Because CPUEs with conflicting information were identified, it may be reasonable to assume that the indices reflect alternative hypotheses about states of nature and to run separate scenarios for each group. As previously, the diagnostic model option incorporates a single set of longline CPUE indices (from the south-western fishery).

3.12 Selection of a diagnostic case.

During the data prep meeting (April 2022) the WPTmT noted that the previous assessments investigated a wide range of structural assumptions and those results informed to structure of the most recent assessment. For the current assessment, there are no additional data to enable any further investigation of some of the main structural assumptions related to spatial structure, fishery selectivities, initial conditions and recruitment estimation.

Instead, this assessment focussed on the development of the individual fleets, and surveys by quarter primarily to accommodate the differences in the length composition and CPUE trends amongst the four regional longline fisheries. This is consistent with the 'fleets as areas' approach. The initial model was based around the assessment model that used the -Southwest - CPUE(Region 3), with longline and and Purse Seine length frequency.
Sensitivity trials were run using the other CPUE time series and combinations of CPUE. Groupings of CPUE series will be chosen by the WPTmT which will seek to use the results of the HCA, other diagnostics as well as expert opinion to selected the most appropriate parameterization and CPUE series groupings.

3.13 Benchmark and Reference Point Methods

Benchmarks included estimates of absolute population levels and fishing mortality for the terminal year, 2020 (F_{2020}, SSB_{2020}, B_{2020}). These values are reported against reference points relative to MSY levels, and depletion estimates (relative to virgin levels).

3.14 Diagnostics and additional model runs

Additional model diagnostics which were carried out includes expanded analysis on the residual and hierarchal cluster analysis, runs tests and joint residual plots, likelihood profiles, and age structured production models.

4 Results

In this section we focus on the results from the diagnostic case model and the key results and diagnostics for this model. We then comment on any important differences in both outputs and model diagnostics for the sensitivity analyses, and present preliminary results. Stock Synthesis 3(v3.30.19.01) was implemented here as a length-based age-structured stock assessment model (Methot and Wetzel 2013; e.g., Wetzel and Punt 2011a, 2011b). Stock synthesis utilizes an integrated modeling approach (Maunder and Punt 2013) to take advantage of the many data sources available for the Indian Ocean stock of albacore tuna. An advantage of the integrated modeling approach is that the development of statistical methods that combine several sources of information into a single analysis allows for consistency in assumptions and permits the uncertainty associated with each data source to be propagated to final model outputs (Maunder and Punt 2013).

4.1 Diagnostic case model

The choice of model parameters and data inputs reflected the input of the WPTmT 08 data prep meeting and the available updated data for biology and life history to the extent possible. This will be further refined by input at the WPTmT assessment meeting. This case was built around the SW CPUE run.

Model Fits to Abundance Indices

The model was able to fit the general trends of the indices of abundance (Figure 10). Although the CPUE series for quarter 2 underestimates the CPUE in the recent years of the model, while the index from quarter 4 over estimates the final period of the model with respect to the index estimates. Indices 1 and 3 exhibit annual variation in addition to a decline starting around 2010 and lasting until 2016 or 2016 before increasing. The indices for quarter 2 and 4 also show this trend to a lesser extent. The model does not quite fit this trend. The spawning output was estimated to increase slightly in the late 1990s to the early 2000s followed by a period of decline coincident with the increase in catch (Figure 2) and decline in the CPUE series.

Fits to the Length composition

The overall fit to the length data was generally good (Figure 13). Fleet specific annual length samples were often quite different, i.e. left skewed one year and bimodal the next, which accounts for the small amount of misfit in the aggregated samples. Pearson residuals of the fit to the length compositions were generally small – on the order of 2 to -2 and did not show any temporal trend (Figures 14-16).

Stock-recruitment Parameters

The predicted virgin recruitment (R0; number of age 0) was approximately 17,000,000 animals and the number of recruits was relatively constant from the early 1950s Through the early 1980s, after which estimated recruitment experienced large fluctuations from about 1980-2020 (Figure 16). The bias correction estimated in the model is shown in Figure 17. The corresponding estimated stock recruitment relationship and annual deviations are also shown in Figure 18.

Fishing Mortality

Estimated F/F_{MSY} and fleet-specific instantaneous fishing mortality rates are presented in Figures 20 and 21 respectively. Fishing mortality was relatively low from the 1950 to the mid 1990s, which is in accordance with low catches and effort during that period. In the late 1980s fishing mortality increased with the advent of driftnet fishery. Starting in the late-1990s overall fishing mortality increased, with large fluctuations in the individual fisheries contribution to the overall fishing mortality. Since the early 2000s the overall fishing mortality has been increasing but below F_{MSY} (i.e. overfishing is not occurring) for the entire time series.

Estimated stock status and other quantities

The estimated equilibrium yield curve for the diagnostic case model is shown in Figure 21. The estimated MSY is approximately 47,000 MT and this is predicted to occur at 21.5% of the unfished biomass (Figure 21), which is less than the standard Schaefer production model ($0.5SSB_0$). The diagnostic case model estimates that the total biomass of the stock was at approximately 100% of the unfished level at the start of the model period (Figure 11) and

steadily decreased to an estimate of $SB_{2020}/SB_{MSY} = 1.9$ that corresponds with $F_{2020}/F_{MSY} = 0.62$. Recruitment is fairly well estimated throughout the model time period (Figure 8), with recent recruitment estimated to be lower than the implied stock recruitment curve due to deviations implied by the length data. The estimates of recruitment were quite tightly constrained to the stock recruitment curve for the initial period of the model when there was no length information to inform the model. The main trends in the population dynamics can be explained through the estimated fishing mortality which was greatly increased in the 1990s and early 2000s due to the increase in catch (Figures 19 and 20). These changes in fishing mortality correspond to an overall stock status that is headed from a virgin state to the direction of overfished and overfishing (Figure 22).

Model Uncertainty

Stock status uncertainty was evaluated delta-Multivariate lognormal (MVLN)approximation to generate joint error distributions for SSB_{2020}/SSB_{MSY} and F_{2020}/F_{MSY}. Figure 22 shows the estimated stock status based on the MVLN analysis for the base case model and Figure 23 shows the estimated timeseries based on the MVLN approximation. Figure 24 shows the distribution of the MLE estimates of SSB_{2020}/SSB_{MSY} and F_{2020}/F_{MSY}.

Stock synthesis provides estimates of the MSY-related quantities and these and other quantities of interest for management are provided in Table 4.

Retrospective Analysis

As part of an analysis of model structure, retrospective analysis (sequentially deleting 1 year of data from the end of the model and re-running) was run using the diagnostic case formulation (the southwest CPUE series). The estimates of spawning depletion remain very similar across all the retrospective model runs considered (Figure 25) indicating that the changes in estimates of virgin spawning biomass are based on the total catch (Figure 25 right panel). The last retrospective run (-5 years) estimated a more depleted stock that corresponds to a slightly smaller virgin recruitment (Figure 25 right panel), this is associated with higher estimated total fishing mortalities in the last 4 years. In general the retrospective analysis shows no large

departures from the estimated scale, depletion, or overall trend based on the sequential deletion of the last 7 years of data.

4.2 Other model diagnostics

Annex 1 shows the results of the expanded analysis on the residuals, hierarchal cluster analysis, runs tests joint residual plots, likelihood profiles, age structured production model and hindcasting cross-validation are presented in Annex 1. Select details from Annex 1 are repeated here, the reader is encouraged to read Annex 1 in its entirety.

The runs test indicated that the CPUE from quarter 1 and two and length data did contain some patterns in the residuals, indicating a trend in the departure from the expected values the (Figures A7 and A8) the cause of this was the early time period in which the index had high contrast and little data. The joint residual plots (Figure A6) shows that in the latter part of the time series the model fits the Japanese data fairly well.

An age structured production model (ASPM) can help evaluate whether the catch and CPUE data give evidence for a production function within the model (Carvalho 2017). Overall the ASPM evaluates whether the effect of surplus production and observed catches alone could explain trends in the CPUE, in contrast to a more complex model (i.e. SS3) that incorporates annual recruitment deviations to improve the fit (Carvalho et al 2021). Maunder and Piner (2017) note that if the ASPM fits well to the indices of abundance with contrast the production function is likely to drive the stock dynamics and the indices will provide information about absolute abundance (Minte-Vera et al., 2017). Figure A13 shows that the biomass trajectories for both models (ASPM and the diagnostic) follow the same trend and that the estimates of *LN(R0)* are different. The fits to the indices are shown in Figure A14 and indicate an overall good fit, indicating that the information content in the data is sufficient.

5 Conclusion

The overall scale of the estimated stock biomass is similar to the previous assessment, though there is variation based on the model parameterization. The main change from the 2019 assessment is the separation of the longline CPUE series by quarter. The disaggregation of the CPUE series allowed the investigation of models based on seasonal quarters, and the exclusion of highly variable quarterly data.

The current assessment depends largely on the time series of southwestern longline CPUE indices and the catch history from the entire fishery. The LL3 CPUE indices showed declining trends during the late 1980s early 1990s for quarter 1 and 2. The CPUE trends in quarter 3 and 4 were fairly stable and lower in comparison. The declining trends in the CPUE indices during the early part of the model match the overall catch from the fisheries.

The main drivers of this assessment are the trend in the catch and CPUE series. In particular the large increase in recent years of catch has different interpretations (within the model). based on which the CPUE series are included.

6 Acknowledgements

The author would like to thank the IOTC staff for their help in preparing the data as well as all of the member countries that submitted data and analyses. This assessment benefited from the discussion and contribution of the WPTmT data prep attendees.

7 Reference

Bartoo, N., Holts. D. (1993). Estimated drift gillnet selectivity for albacore Thunnus alalunga. Fishery Bulletin92:371-378.

Bromhead, D., B., A. Williams and S. D. Hoyle (2009). Factors affecting size composition data from south Pacific albacore longline fisheries.

Cadrin, S.X., Vaughn, D.S., 1997. Retrospective analysis of virtual population estimates for Atlantic menhaden stock assessment. Fish. Bull. 95(3), 445-455.

Chang, F.-C., C.-Y. Chen, L.-K. Lee and S.-Y. Yeh (2012). Assessment on Indian albacore stock based mainly on Taiwanese longline data. Fourth Working Party on Temperate Tunas, Shanghai, China, 20–22 August 2012. . Indian Ocean Tuna Commission. IOTC-2012-WPTmT04-19.

Chang, S., H. Liu (1995). Adjusted Indian Ocean albacore CPUE series of Taiwanese longline and drift net fisheries. Sixth session of the Expert Consultations on Indian Ocean tunas, Colombo, Sri Lanka 25–29 September 1995. IOTC-1995-EC602-26.

Chang, S., H. Liu and C.-C. Hsu (1993). "Estimation of vital parameters for Indian albacore through length frequency data." Journal of the Fisheries Society of Taiwan 20(1): 1-17.

Chen, I. C., P. F. Lee and W. N. Tzeng (2005). "Distribution of albacore (Thunnus alalunga) in the Indian Ocean and its relation to environmental factors." Fisheries Oceanography 14(1): 71-80. Chen, K. S., T. Shimose, T. Tanabe, C. Y. Chen and C. C. Hsu (2012). "Age and growth of albacore Thunnus alalunga in the North Pacific Ocean." Journal of fish biology 80(6): 2328-2344.

Dhurmeea, Z.; Chassot, E.; Augustin, E.; Assan, C.; Nikolic, N.; Bourjea, J. et al. (2016). Morphometrics of albacore tuna (Thunnus alalunga) in the Western Indian Ocean. Seventh Working Party on Temperate Tunas: Data Preparatory Meeting, Kuala Lumpa, Malaysia, 14–17 January 2019. Indian Ocean Tuna Commission. IOTC-2019-WPTmT07(DP)-INF02.

Dhurmeea, Z.; Zudaire, I.; Chassot, E.; Cedras, M.; Nikolic, N.; Bourjea, J. et al. (2016) Reproductive Biology of Albacore Tuna (*Thunnus alalunga*) in the Western Indian Ocean. PLoS ONE 11(12): e0168605. doi:10.1371/journal.pone.0168605

Farley, J.H.; Eveson, J.P.; Bonhommeau, S.; Dhurmeea, Z.; West, W.; Bodin, N. (2019). Growth of albacore tuna (Thunnus alalunga) in the western Indian Ocean using direct age methods. Seventh Working Party on Temperate Tunas: Data Preparatory Meeting, Kuala Lumpa, Malaysia, 14–17 January 2019. Indian Ocean Tuna Commission. IOTC-2019-WPTmT07(DP)-21.

Farley, J. H., S. D. Hoyle, J. P. Eveson, A. J. Williams, C. R. Davies and S. J. Nicol (2014). "Maturity Ogives for South Pacific Albacore Tuna (Thunnus alalunga) That Account for Spatial and Seasonal Variation in the Distributions of Mature and Immature Fish." PloS one 9(1): e83017.

Farley, J., A. Williams, N. Clear, C. Davies and S. Nicol (2013). "Age estimation and validation for South Pacific albacore Thunnus alalunga." Journal of fish biology 82(5): 1523-1544.

Farley, J.H., A.J. Williams, S. D. Hoyle, C. R. Davies and S. J. Nicol (2013). "Reproductive dynamics and potential annual fecundity of South Pacific albacore tuna (Thunnus alalunga)." PLoS ONE 8(4): e60577.

Fonteneau, A. (2015). Indian Ocean albacore stock: review of its fishery, biological data and results of its 2014 stock assessment. IOTC/2015/XXX.

Fournier D A, Skaug HJ, Ancheta J, Ianelli J, Magnusson A, Maunder M, Nielsen A, Sibert J (2012) AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw. 27:233-249.

Francis, R. I. C. C. (2011). "Data weighting in statistical fisheries stock assessment models." Canadian Journal of Fisheries and Aquatic Sciences 68(6): 1124-1138.

Francis RICC (2014) Replacing the multinomial in stock assessment models: A first step. Fisheries Research, 151, (2014), 70-84

Geehan, J. and S. Hoyle (2013). Review of length frequency data of the Taiwan,China distant water longline fleet, IOTC-2013-WPTT15-41 Rev_1. Indian Ocean Tuna Commission Working Party on Tropical Tunas, San Sebastian, Spain, 23–28 October, 2013.

Guan, W.; Zhu, J.; Xu, L.; Dai, X, (2016). Using a Bayesian biomass dynamics model to assess Indian Ocean albacore. Sixth Working Party on Temperate Tunas, Shanghai, China, 18–21 July 2016. Indian Ocean Tuna Commission. IOTC-2014-WPTmT06-22.

Gunn, J. S., N. P. Clear, T. I. Carter, A. J. Rees, C. A. Stanley, J. H. Farley and J. M. Kalish (2008). "Age and growth in southern bluefin tuna, *Thunnus maccoyii* (Castelnau): Direct estimation from otoliths, scales and vertebrae." Fisheries Research 92(2): 207-220.

Hoyle, S.D. (2019). Regional scaling factors for Indian Ocean albacore tuna. Seventh Working Party on Temperate Tunas: Data Preparatory Meeting, Kuala Lumpa, Malaysia, 14–17 January 2019. Indian Ocean Tuna Commission. IOTC-2019-WPTmT07(DP)-13.

Hoyle, S., J. Hampton and N. Davies (2012). Stock assessment of albacore tuna in the South Pacific ocean. Scientific Committee, Eighth Regular Session, 7-15 August 2012, Busan, Republic of Korea: 90.

Hoyle, S. D. (2008). Adjusted biological parameters and spawning biomass calculations for south Pacific albacore tuna, and their implications for stock assessments. WCPFC Scientific Committee. Nouméa, New Caledonia: 20.

Hoyle, S. D., A. D. Langley and R. A. Campbell (2014). "Recommended approaches for standardizing CPUE data from pelagic fisheries."

Hoyle, S.D., R. Sharma, M. Herrera (2014). Indian Ocean albacore assessment. Fifth Working Party on Temperate Tunas, Busan, Korea, 28–31 July 2014. Indian Ocean Tuna Commission. IOTC-2014-WPTmT05-24.

Hoyle, S.D. Yin Chang, Doo Nam Kim, Sung Il Lee, Takayuki Matsumoto, Kaisuke Satoh, and Yu-Min Yeh (2016). Collaborative study of albacore tuna CPUE from multiple Indian Ocean longline fleets. Sixth Working Party on Temperate Tunas, Shanghai, China, 18–21 July 2016. Indian Ocean Tuna Commission. IOTC-WPTmT06-19a.

Hoyle, S.D.; Chassot, E.; Fu, D.; Kim, D.N.; Lee, S.; Matsumoto, T.; Satoh, K.; Wang, S.; Kitakado, T. (2019). Collaborative study of albacore tuna CPUE from multiple Indian Ocean longline fleets in 2019. Seventh Working Party on Temperate Tunas: Data Preparatory Meeting, Kuala Lumpa, Malaysia, 14–17 January 2019. Indian Ocean Tuna Commission. IOTC-2019-WPTmT07(DP)-19.

Hoyle, S.D.; Yeh, Y.M.; Chang, S.T.; Wu, R.F. (2015). Descriptive analysis of the Taiwanese Indian Ocean longline fishery, focussing on tropical areas. Seventeenth Working Party on Tropical Tunas, Montpellier, France, 23–28 October, 2015. Indian Ocean Tuna Commission. IOTC-2015-WPTT17-INF09.

Hoyle, S.D. Chang, S-t. et al. (2021) Draft review of size data from Indian Ocean longline fleets, and its utility for stock assessment. IOTC-2021-WPTT23(DP)-08.

Huang, H.; Liu, K. (2010). Bycatch and discards by the Taiwanese large-scale tuna longline fleets in the IndianOcean. Fisheries Research 106: 261-270

Huang, H., C.-L. Wu, C. Kuo and W. Su (1991). "Age and growth of Indian Ocean albacore (Thunnus alalunga) by scales."

Ianelli, J., M. Maunder and A. E. Punt (2012). Independent review of 2011 WCPO bigeye tuna assessment. IOTC Secretariat (2013). Report and documentation of the Indian Ocean Tuna Fisheries of Indonesia Albacore

Catch Estimation Workshop: Review of Issues and Considerations. Bogor-Jakarta, 21-25 June 2013. IOTC Technical Report. No. 2013/01. Bogor-Jakarta, IOTC. 2013. 40 pp.

IOTC (2014). Report of the Fifth Session of the IOTC Working Party on Temperate Tunas. Fifth Working Party on Temperate Tunas, Busan, Korea, 28–31 July 2014. Indian Ocean Tuna Commission. IOTC-2014-WPTmT05-R[E].

IOTC (2016). Report of the Sixth Session of the IOTC Working Party on Temperate Tunas. Sixth Working Party on Temperate Tunas, Shanghai, China, 18–21 July 2016. Indian Ocean Tuna Commission. IOTC-2016-WPTmT06-R[E].

IOTC (2019). Report of the Seventh Session of the IOTC Working Party on Temperate Tunas (Data Preparatory Session). Seventh Working Party on Temperate Tunas: Data Preparatory Meeting, Kuala Lumpa, Malaysia, 14–17 January 2019. Indian Ocean Tuna Commission. IOTC-2019-WPTmT07(DP)- RE.

IOTC–WPTmT07(AS) 2019. Report of the Seventh Session of the IOTC Working Party on Temperate Tunas: Assessment Meeting. Shizuoka, Japan, 23–27 July 2019. IOTC–2019–WPTmT07(AS)–R[E]: 36 pp.

IOTC–WPTmT08(DP) 2022. Report of the Eighth Session of the IOTC Working Party on Temperate Tunas. Online, 13–15 April 2022. *IOTC–2022–WPTmT08(DP)–R[E]: 36 pp.*

ISSF (2011). Report of the 2011 ISSF Stock Assessment Workshop: Rome, Italy, March 14-17, 2011. ISSF Technical Report 2011-02, http://iss-foundation.org/wp-content/uploads/downloads/2011/08/ISSF-2011-02- Report-2011-ISSF-WS.pdf. McLean, Virginia, USA, International Seafood Sustainability Foundation.

Karakulak, F. S., E. Özgür, M. Gökoğlu, I. T. Emecan and A. Başkaya (2011). "Age and growth of albacore (Thunnus alalunga Bonnaterre, 1788) from the eastern Mediterranean." Turkish Journal of Zoology 35(6): 801-810.

Kitakado, T.; Fiorellato, F.; De Bruyn, P. (2019). Allometric curve for the Indian Ocean albacore. Seventh Working Party on Temperate Tunas: Data Preparatory Meeting, Kuala Lumpa, Malaysia, 14–17 January 2019. Indian Ocean Tuna Commission. IOTC-2019-WPTmT07XX.

Kitakado, T., E. Takashima, T. Matsumoto, T. Ijima and T. Nishida (2012). First attempt of stock assessment using Stock Synthesis III (SS3) for the Indian Ocean albacore tuna (Thunnus alalunga). IOTC Working Party on Temperate Tunas, Shanghai, China, 20–22 August 2012, Indian Ocean Tuna Commission.

Kitakado, T et al. (2022) Joint CPUE indices for the albacore Thunnus alalunga in the Indian Ocean based on Japanese, Korean and Taiwanese longline fisheries data. IOTC–2022–WPTmT08(DP)–15

Lan, K.; Kawamura, H.; Lee, M.; Lu, H.; Shimada, T.; Hosoda, K.; Saikaida, F. (2011). Relationship between albacore (Thunnus alalunga) fishing grounds in the Indian Ocean and the thermal environment revealed by cloud-free microwave sea surface temperature. Fisheries Research 113: 1-7

Langley, A.D.; Hoyle, S.D. (2016). Stock assessment of albacore tuna in the Indian Ocean using Stock Synthesis. Sixth Working Party on Temperate Tunas, Shanghai, China, 18–21 July 2016. Indian Ocean Tuna Commission. IOTC-WPTmT06-25.

Lee, H.-H., Piner, K.R., Methot R.D., Maunder, M.N. 2014. Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: An example using blue marlin in the Pacific Ocean. Fish. Res.

Lee, H. H., M. N. Maunder, K. R. Piner and R. D. Methot (2012). "Can steepness of the stock-recruitment relationship be estimated in fishery stock assessment models?" Fisheries Research.

Lee, L.-K., C.-C. Hsu and F.-C. Chang (2014). Albacore (*Thunnus alalunga*) CPUE Trend from Indian Core Albacore Areas based on Taiwanese longline catch and effort statistics dating from 1980 to 2013.

Lee, Y. C. and H. C. Liu (1992). "Age determination, by vertebra reading, in Indian albacore, *Thunnus alalunga* (Bonnaterre)." Journal of the Fisheries Society of Taiwan 19(2): 89-102.

Lorenzen, K. (1996). "The relationship between body weight and natural mortality in juvenile and adult fish - a comparison of natural ecosystems and Aquaculture [Review]." Journal of Fish Biology 49(4): 627-647.

Matsumoto, T., T. Kitakado and T. Nishida (2014). Standardization of albacore CPUE by Japanese longline fishery in the Indian Ocean, IOTC–2014–WPTmT05–18 Indian Ocean Tuna Commission Working Party on Temperate Tuna (WPTmT): 16.

Matsumoto, T., T. Nishida and T. Kitakado (2012). Stock and risk assessments of albacore in the Indian Ocean based on ASPIC. Fourth Working Party on Temperate Tunas, Shanghai, China, 20–22 August 2012. Indian Ocean Tuna Commission. IOTC-2012-WPTmT04-20.

Matsumoto, T., T. Nishida and T. Kitakado (2014). Stock and risk assessments of albacore in the Indian Ocean based on ASPIC. Fifth Working Party on Temperate Tunas, Busan, Korea, 28–31 July 2014. Indian Ocean Tuna Commission. IOTC-2014-WPTmT05-22 Rev_1.

Matsumoto, T. (2016). Stock and risk assessments of albacore in the Indian Ocean based on ASPIC. Sixth Working Party on Temperate Tunas, Shanghai, China, 18–21 July 2016. Indian Ocean Tuna Commission. IOTC-2014-WPTmT06-20.

McAllister, M. K. and J. N. Ianelli (1997). "Bayesian stock assessment using catch-age data and the sampling importance resampling algorithm." Canadian Journal of Fisheries and Aquatic Sciences 54(2): 284-300.

Megalofonou, P. (2000). "Age and growth of Mediterranean albacore." Journal of Fish Biology 57(3): 700-715.

Methot, R.D. (2022). User Manual for Stock Synthesis model Version 330.19. 29 March2022.

Methot, R. D., Jr and I. G. Taylor (2011). "Adjusting for bias due to variability of estimated recruitments in fishery assessment models." Canadian Journal of Fisheries and Aquatic Sciences 68(10): 1744-1760.

Methot, R. D. and C. R. Wetzel (2013). "Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management." Fisheries Research 142: 86-99.

Methot, R. D. (2005) Technical description of the stock synthesis II assessment program: Version 1.17 (March, 2005), 54p.

Nikolic, N., Fonteneau, A., Hoarau, L., Morandeau, G., Puech, A., Bourjea, J. (2013). Short review on biology, structure, and migration of Thunnus alalunga in the Indian Ocean. Fifth Working Party on Temperate Tunas, Busan, Korea, 28–31 July 2014. Indian Ocean Tuna Commission. IOTC-2014-WPTmT05-13.

Nishida, T.; Dhurmeea, Z. (2019). Review of Indian Ocean albacore biological parameters for stock assessments. Seventh Working Party on Temperate Tunas: Data Preparatory Meeting, Kuala Lumpa, Malaysia, 14–17 January 2019. Indian Ocean Tuna Commission. IOTC-2019-WPTmT07(DP)-12

Nishida, T., T. Matsumoto and T. Kitakado (2012). Stock and risk assessments on albacore (*Thunnus alalunga*) in the Indian Ocean based on AD Model Builder implemented Age-Structured Production Model (ASPM).

Nishida, T.; Sato, K.; Chang, Y.; Matsumoto, T.; Lee, Y.C.; Kitakado, T. (2016). Stock assessments of albacore (*Thunnus alalunga*) in the Indian Ocean using Statistical-Catch-at-Age (SCAA). Age-Structured Production Model (ASPM). Sixth Working Party on Temperate Tunas, Shanghai, China, 18–21 July 2016. Indian Ocean Tuna Commission. IOTC-2014-WPTmT06-21.

Saji, N.H.; Yamagata, T. (2003). Possible impacts of Indian Ocean Dipole mode events on global climate. Climate Research 25(2) 151-169.

Santiago, J. and H. Arrizabalaga (2005). "An integrated growth study for North Atlantic albacore (*Thunnus alalunga*, Bonn. 1788)." ICES Journal of Marine Science 62(4): 740-749.

Schnute J.T., and R. Hilborn. 1993. Analysis of contradictory data sources in fish stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 50 (9): 1916-1923.

Wells, R. J. D., S. Kohin, S. L. H. Teo, O. E. Snodgrass and K. Uosaki (2013). "Age and growth of North Pacific albacore (*Thunnus alalunga*): Implications for stock assessment." Fisheries Research 147: 55-62.

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

Xu, Y., T. Sippel, S. L. H. Teo, K. Piner, K.-S. Chen and R. J. Wells (2014). A comparison study of North Pacific albacore (Thunnus alalunga) age and growth among various sources, ISC/14/ALBWG/04 ISC Albacore Working Group Meeting, 14-28 April 2014, La Jolla, CA, 92037, USA.

Yeh, S.; Hui, C.; Treng, T.; Kuo, C. (1995). Indian Ocean albacore stock structure studies by morphometric and DNA sequence methods. Sixth Session of the Expert Consultations on Indian Ocean Tunas, Columbo, Sri Lanka, 25–29 September 1995. Indian Ocean Tuna Commission. IOTC-1995-EC602-25.

Zhu, J., Y. Chen, X. Dai, S. J. Harley, S. D. Hoyle, M. N. Maunder and A. M. Aires-da-Silva (2012). "Implications of uncertainty in the spawner–recruitment relationship for fisheries management: An illustration using bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean." Fisheries Research 119-120: 89-93.

8 Tables

Table 1. Fishery and survey definitions for the Indian Ocean Albacore Assessment

ID Numbe	Fleet/ Survey Name	Gear(s)	Region	Quarter	Selectivty
1	F1_LL1_Q1	Longline	1	1	Estimated double normal
2	F2_LL1_Q2	Longline	1	2	Estimated double normal
3	F3_LL1_Q3	Longline	1	3	Estimated double normal
4	F4_LL1_Q4	Longline	1	4	Estimated double normal
5	F5_LL2_Q1	Longline	2	1	Estimated double normal
6	F6_LL2_Q2	Longline	2	2	Estimated double normal
7	F7_LL2_Q3	Longline	2	3	Estimated double normal
8	F8_LL2_Q4	Longline	2	4	Estimated double normal
9	F9_LL3_Q1	Longline	3	1	Estimated double normal
10	F10_LL3_Q2	Longline	3	2	Estimated double normal
11	F11_LL3_Q3	Longline	3	3	Estimated double normal
12	F12_LL3_Q4	Longline	3	4	Estimated double normal
13	F13_LL4_Q1	Longline	4	1	Estimated double normal
14	F14_LL4_Q2	Longline	4	2	Estimated double normal
15	F15_LL4_Q3	Longline	4	3	Estimated double normal
16	F16_LL4_Q4	Longline	4	4	Mirrored to Fleet 8
17	F17_DN3	Driftnet	3	NA	Fixed Double normal
18	F18_DN4	Driftnet	4	NA	Mirrored to fleet 17
19	F19_PS1	Purse Seine	1	NA	Fixed Double normal
20	F20_Other1	Other*	1	NA	Mirrored to fleet 17
21	F21_Other2	Other*	2	NA	Mirrored to fleet 17
22	F22_Other3	Other*	3	NA	Mirrored to fleet 17
23	F23_Other4	Other*	4	NA	Mirrored to fleet 17
24	LLCPUE1_Q1	Longline	1	1	Mirrored to fleet 1
25	LLCPUE1_Q2	Longline	1	2	Mirrored to fleet 2
26	LLCPUE1_Q3	Longline	1	3	Mirrored to fleet 3
27	LLCPUE1_Q4	Longline	1	4	Mirrored to fleet 4
28	LLCPUE2_Q1	Longline	2	1	Mirrored to fleet 5
29	LLCPUE2_Q2	Longline	2	2	Mirrored to fleet 6
30	LLCPUE2_Q3	Longline	2	3	Mirrored to fleet 7
31	LLCPUE2_Q4	Longline	2	4	Mirrored to fleet 8
32	LLCPUE3_Q1	Longline	3	1	Mirrored to fleet 9
33	LLCPUE3_Q2	Longline	3	2	Mirrored to fleet 10
34	LLCPUE3_Q3	Longline	3	3	Mirrored to fleet 11
35	LLCPUE3_Q4	Longline	3	4	Mirrored to fleet 12
36	LLCPUE4_Q2	Longline	4	2	Mirrored to fleet 14
37	LLCPUE4_Q3	Longline	4	3	Mirrored to fleet 15
38	DNCPUE4	Driftnet	4	NA	Mirroed to fleet 17

*Other inclues: Coastal Longline, gillnet, trolling, handlines, and artisanal gear.

Table 2. Recent catch data for Albacore in the Indian ocean.

			Catch (MT)	
Longline Fleet No.	Region	Quarter	Average (2010-2020)	Max (2010-2020)
1	1	1	2,001	2,322
2	1	2	287	629
3	1	3	601	1,619
4	1	4	5,030	7,035
5	2	1	1,235	2,289
6	2	2	1,357	2,345
7	2	3	1,912	4,129
8	2	4	1,006	1,549
9	3	1	854	1,957
10	3	2	7,068	10,165
11	3	3	5,735	7,142
12	3	4	692	1,092
13	4	1	344	542
14	4	2	4,072	6,360
15	4	3	2,363	3,873
16	4	4	193	469

Table 3. Summary of SS3 specification options for the Indian Ocean albacore tuna assessment models. Other assumptions were constant for all models. .

Model	Description	Base Parameters	Alternative Parameterizatoin
Steepness0.70	SRR steepness at 0.70	*h*=0.80	*h*=0.70
Steepness0.80	SRR steepness at 0.90	h=0.80	*h*=0.90
LengthAtAge	Increase variation in length-at-age for males and females.	CVs for length-at-age CV_young 0.06 CV_old 0.025	CV_young 0.10 CV_old 0.10
LengthWeight	Use length-weight parameters from Kitakado et al. (2019)	a = 1.3718e-05 b = 3.0973	a = 6.9e-06 b = 3.2263
Nat_Mort	Age Specific Natural mortality	M=0.30 all ages	Lorenzen M, average of 0.3 for age 4+
Maturity	Use maturity at length from recent western Indian Ocean study (rather than South Pacific reproductive potential at length)	Length based Reproductive Potential from South Pacific albacore.	Length based maturity ogive from Dhrumeea et al (2016)
Selectivity	Estimate separate length based selectivities for LL1 and LL2. Approx. logistic form	Each LL fleet has its own selectivity	LL fleets share selectivity by quarter and east-west region. i.e LL2_q1 mirrors LL1_q2.
CPUE South Q2&Q3		Use CPUE from Q1-Q4 in SW	Use CPUE from Q2 &Q3 in SW and SE
CPUE SW Q2&Q3		Use CPUE from Q1-Q4 in SW	Use CPUE from Q2&Q3 SW
CPUE Southwest & Northwest		Use CPUE from Q1-Q4 in SW	Use CPUE from Q1-Q4 in NW and SW

Table 4: Model estimates of stock status and derived quantities. .

.

Model	Diagnostic	Steepness= 0.7	Steepness= 0.9	All Southern CPUE	All Western CPUE	Q2&Q3 CPUE1 and Q1&Q4 CPUE 3	Southwestern CPUE Q2&Q3	Southern CPUE Q2&Q3	Higher Grow CV	Alternate Weight Len curve	Age Specific Natural M.	Alternate Maturity Schedule
C2020_msy	0.87	0.93	0.80	0.84	0.96	0.64	0.85	0.81	1.07	0.88	0.89	0.84
Y_MSY	47,108	44,355	51,375	48,729	42,843	64,127	48,317	50,429	38,304	46,655	46,224	49,089
B_zero	567,296	625,173	527,995	587,570	511,774	770,993	580,706	608,322	410,966	560,057	581,578	552,300
B_msy	122,384	163,700	83,631	126,744	110,642	165,583	124,999	130,655	92,499	121,923	122,315	127,435
B_cur	323,392	356,368	302,211	380,566	280,290	495,191	357,895	434,627	198,745	315,858	328,742	313,648
SB_zero	131,972	145,437	122,830	136,689	119,056	179,167	134,947	141,365	98,012	132,115	137,984	160,640
SB_msy	28,471	38,082	19,456	29,485	25,739	38,479	29,048	30,362	22,060	28,761	29,020	37,065
SB_cur	71,249	80,492	65,274	88,934	47,103	110,348	81,371	104,207	40,567	70,921	73,391	88,861
SB_2020/SB_msy	2.0	1.7	2.7	2.5	1.5	2.3	2.3	2.9	1.5	2.0	2.0	2.1
SB_cur/SB_msy	2.5	2.1	3.4	3.0	1.8	2.9	2.8	3.4	1.8	2.5	2.5	2.4
SB_cur_init	0.5	0.6	0.5	0.7	0.4	0.6	0.6	0.7	0.4	0.5	0.5	0.6
Fcur	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
F_msy	0.2	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.2
F_2020/msy	0.6	0.7	0.5	0.5	0.7	0.4	0.5	0.5	0.9	0.6	0.6	0.6
F_cur/msy	0.5	0.6	0.5	0.4	0.7	0.4	0.5	0.4	0.7	0.6	0.6	0.5
SB_2020	57143.600	65195.600	51952.000	74242.800	37604.800	89630.600	67356.300	89445.000	33715.000	56915.800	58649.700	76102.000
F_2020	0.091	0.083	0.097	0.077	0.106	0.060	0.081	0.068	0.140	0.092	0.090	0.093
gradient	2.53769E-05	1.78272E-05	3.22093E-05	8.66926E-05	6.00051E-05	9.94598E-05	7.64412E-05	4.54122E-05	9.09623E-05	9.02147E-05	9.24413E-05	8.83151E-05

9 Figures

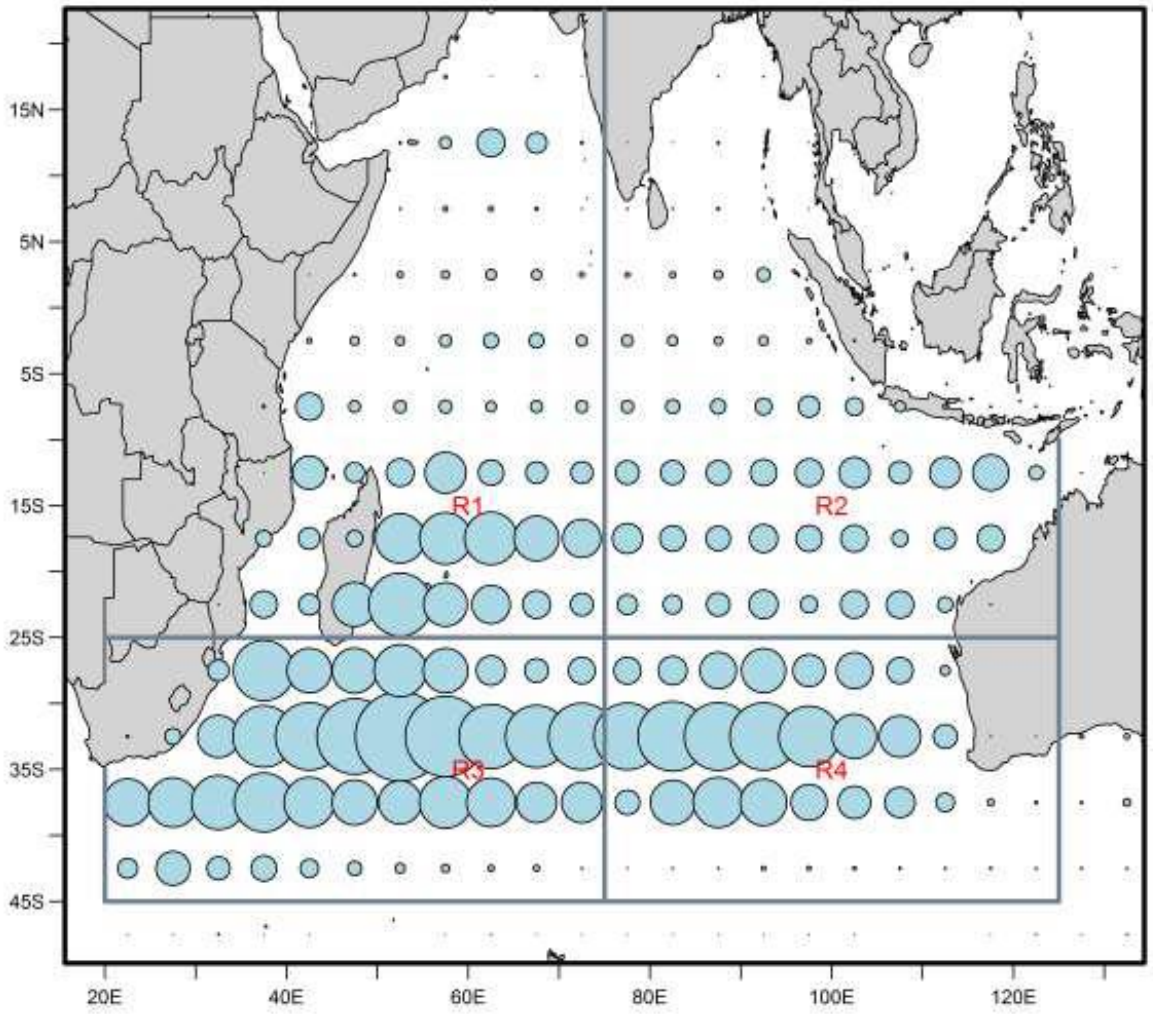

Figure 1. Spatial stratification of the Indian Ocean for the definition of the fisheries. The blue circles represent the aggregated Japanese and TW LL albacore catch (numbers of fish) by 5 degree cell from 1952-2017. The area of the circle is proportional to the magnitude of the catch (the largest circle represents a catch of 2.45 million fish).

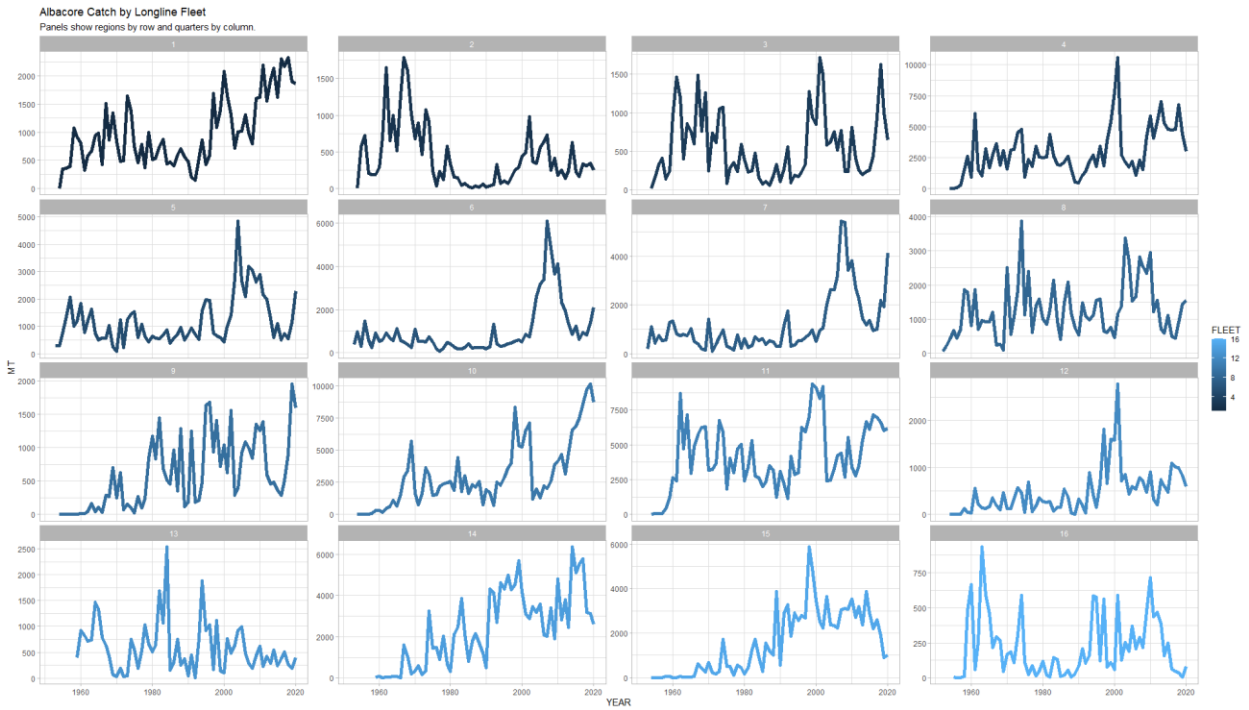

Figure 2. A comparison of Indian Ocean albacore longline catches. Rows indicate region (1-4) and columns indicate quarters (1-4).

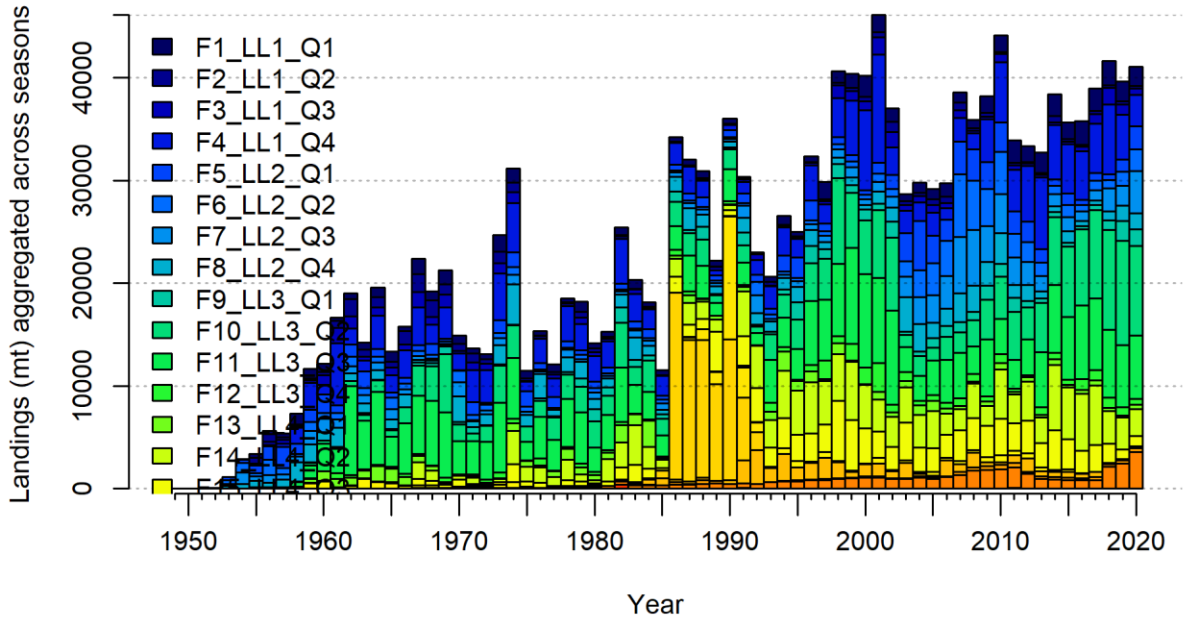

Figure 3. Total annual catch (1000s mt) of albacore tuna by fleet from 1950 to 2020. Fleet names indicate the fleet number, region and quarter, for example F1_LL1_Q1 is Fishing Fleet 1, in Region 1(LL1) and quarter1.

.

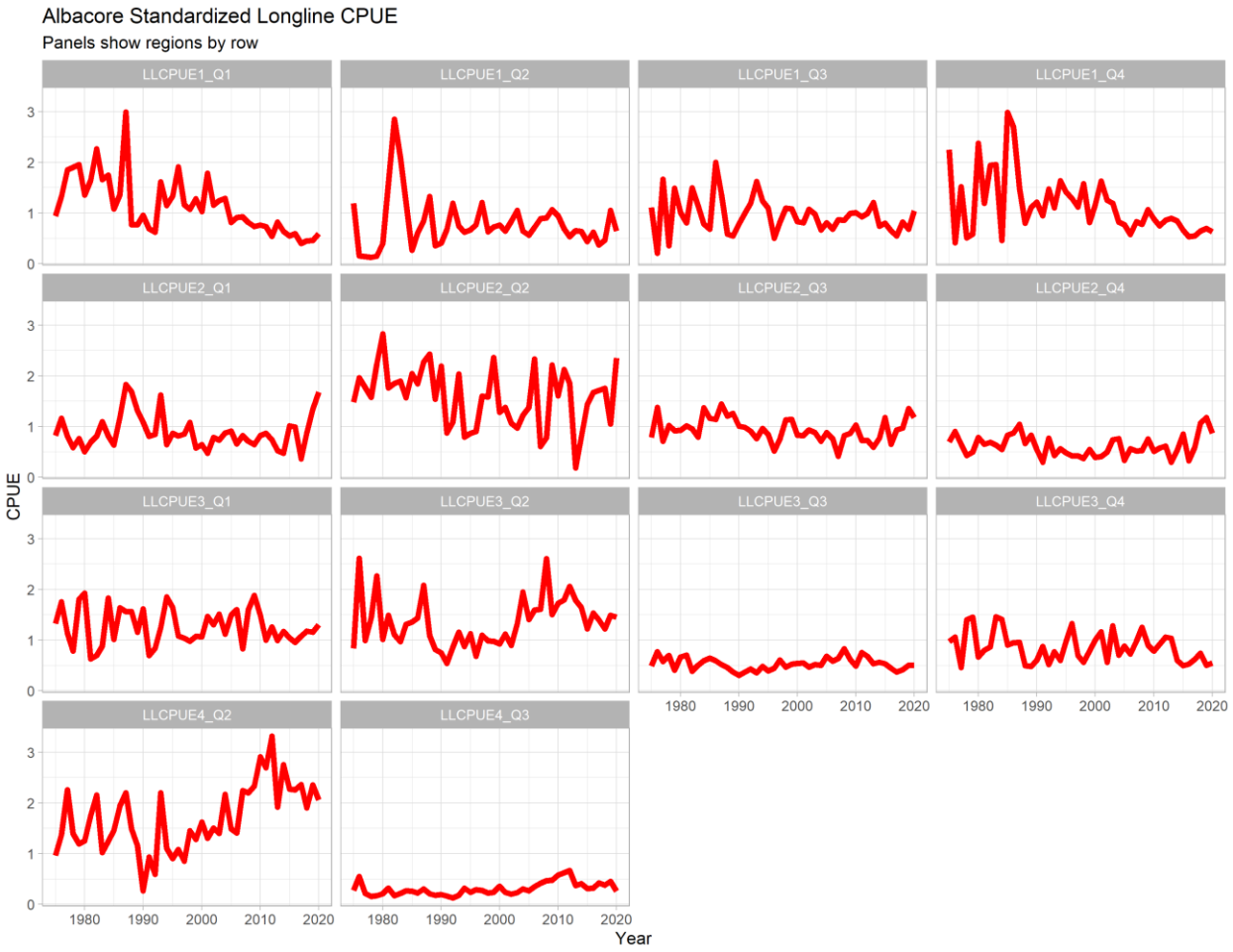

Figure 4. Standardized CPUE indices for the longline fisheries by region and quarter (LL 1-14) from 1975-2020. The figure shows the region in row, and individual quarterly CPUE series in each column. Note that the bottom row, representing region 4, only had CPUE series for quarters 2 and 3.

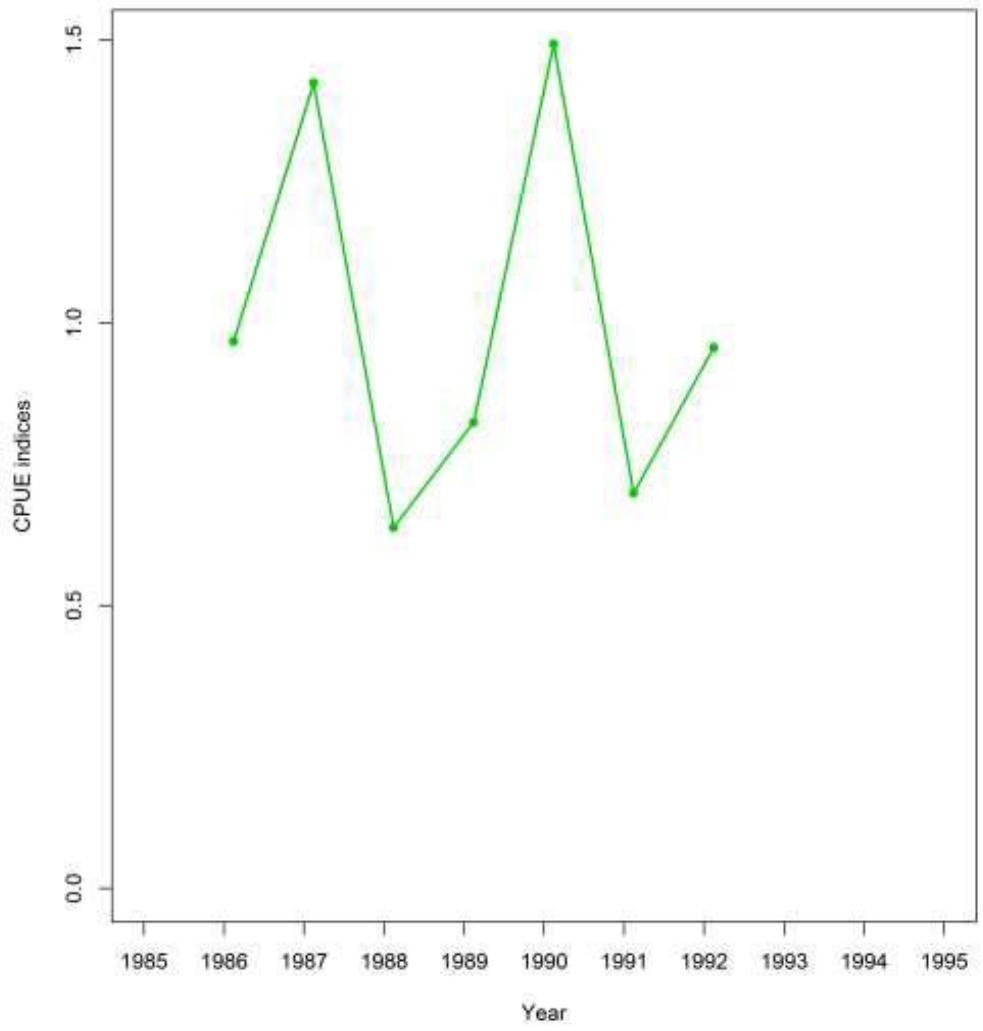

Figure 5 Annual drift net CPUE indices (source: Chang & Liu 1995).

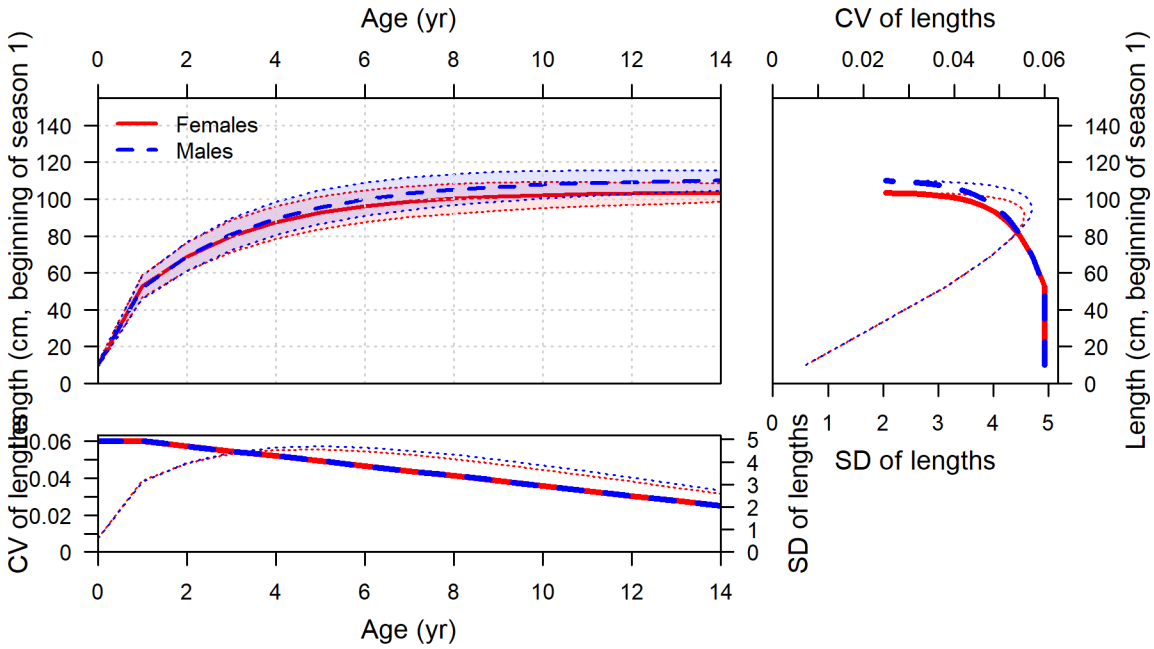

Figure 6. Growth function for female and male albacore.

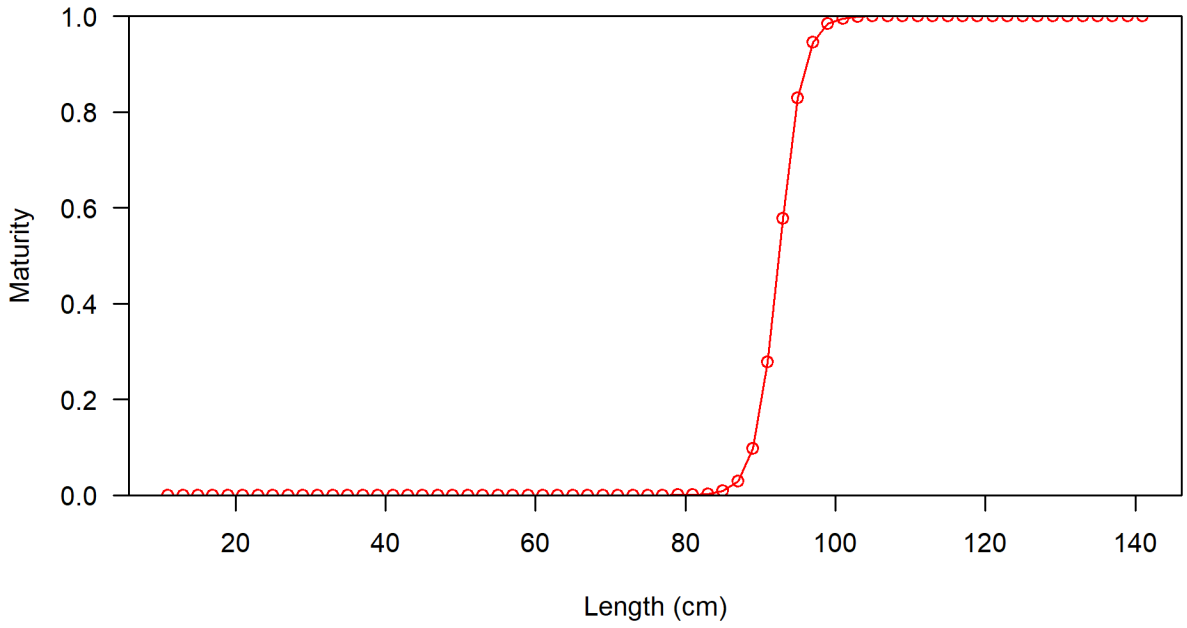

Figure 7 Maturity at length for female albacore in the Indian Ocean.

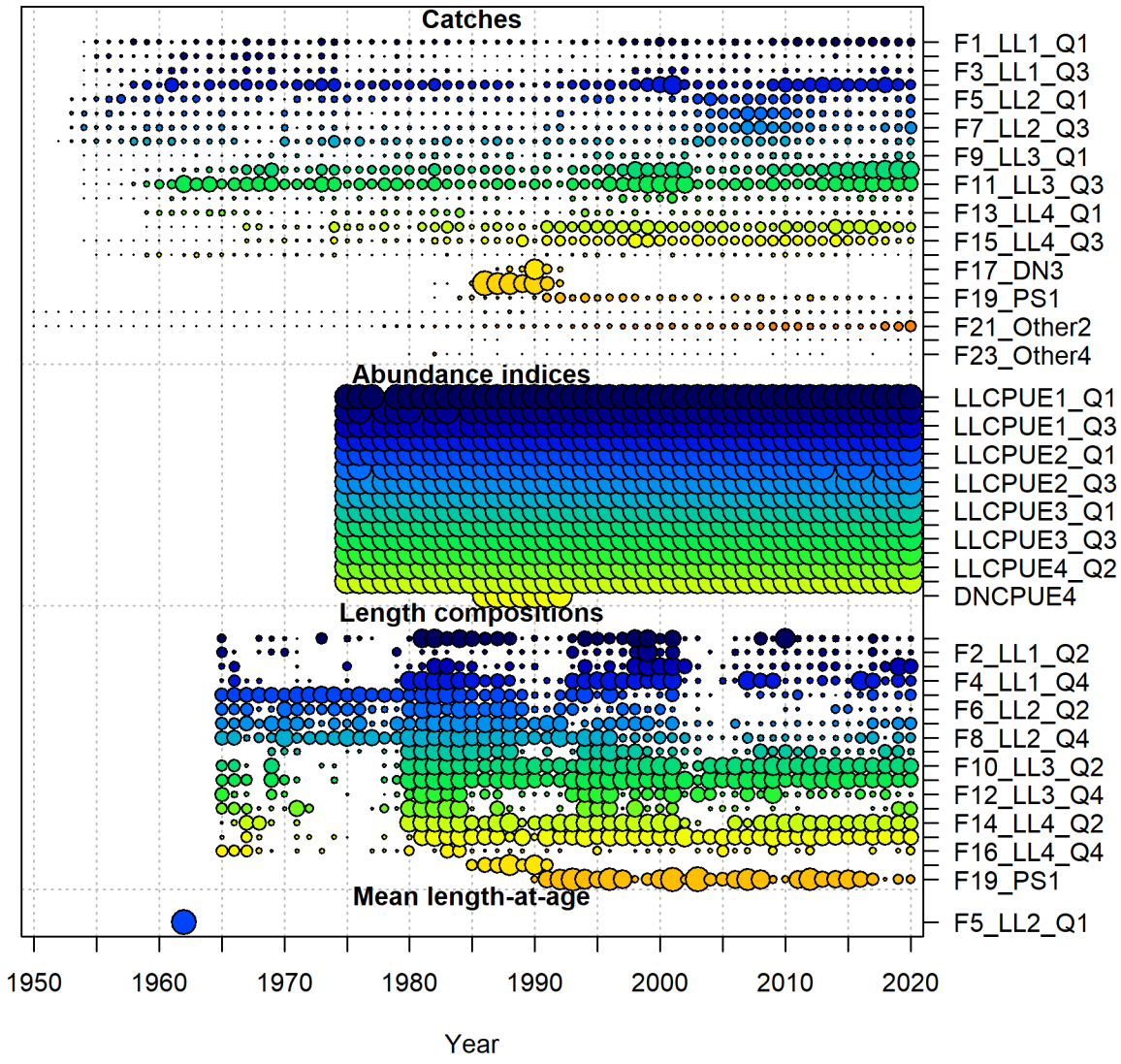

Figure 7: Temporal data coverage for the diagnostic case model for the assessment of albacore tuna.

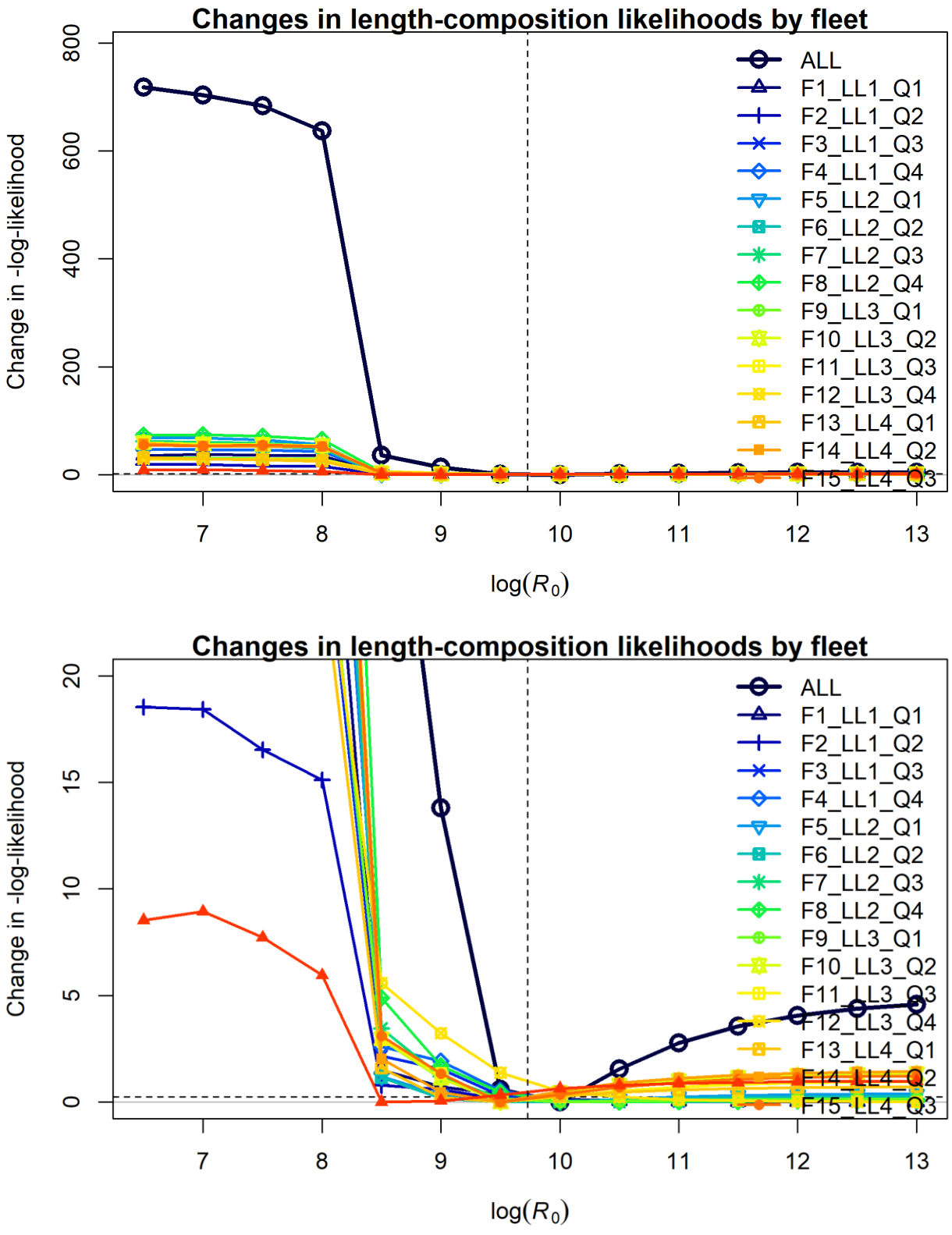

Figure 6: Likelihood profiles for length composition, the bottom panel is a close up version of the top..

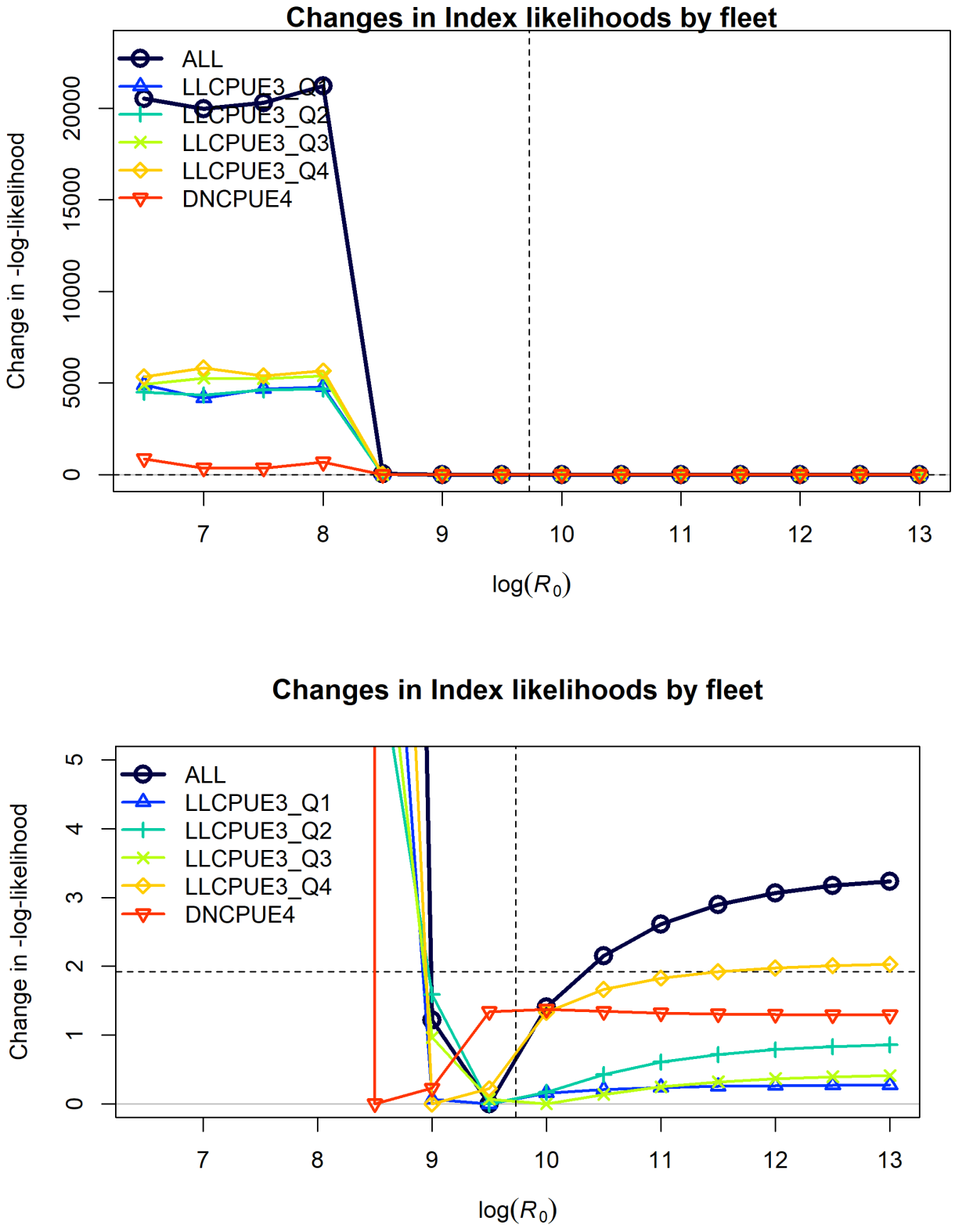

Figure 7: Likelihood profiles for the CPUE components for the reference run, The Bottom panel is a close up of the top.

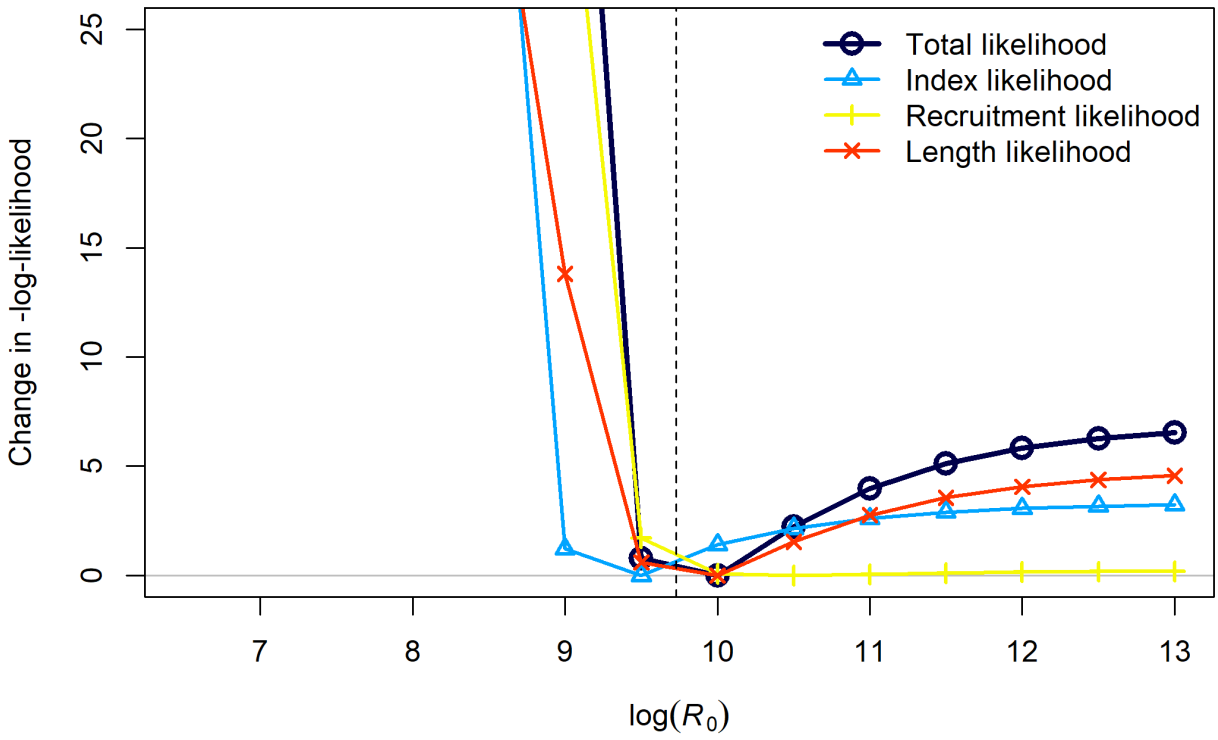

Figure 8 Likelihood profile for the total likelihood, based on the diagnostic run.

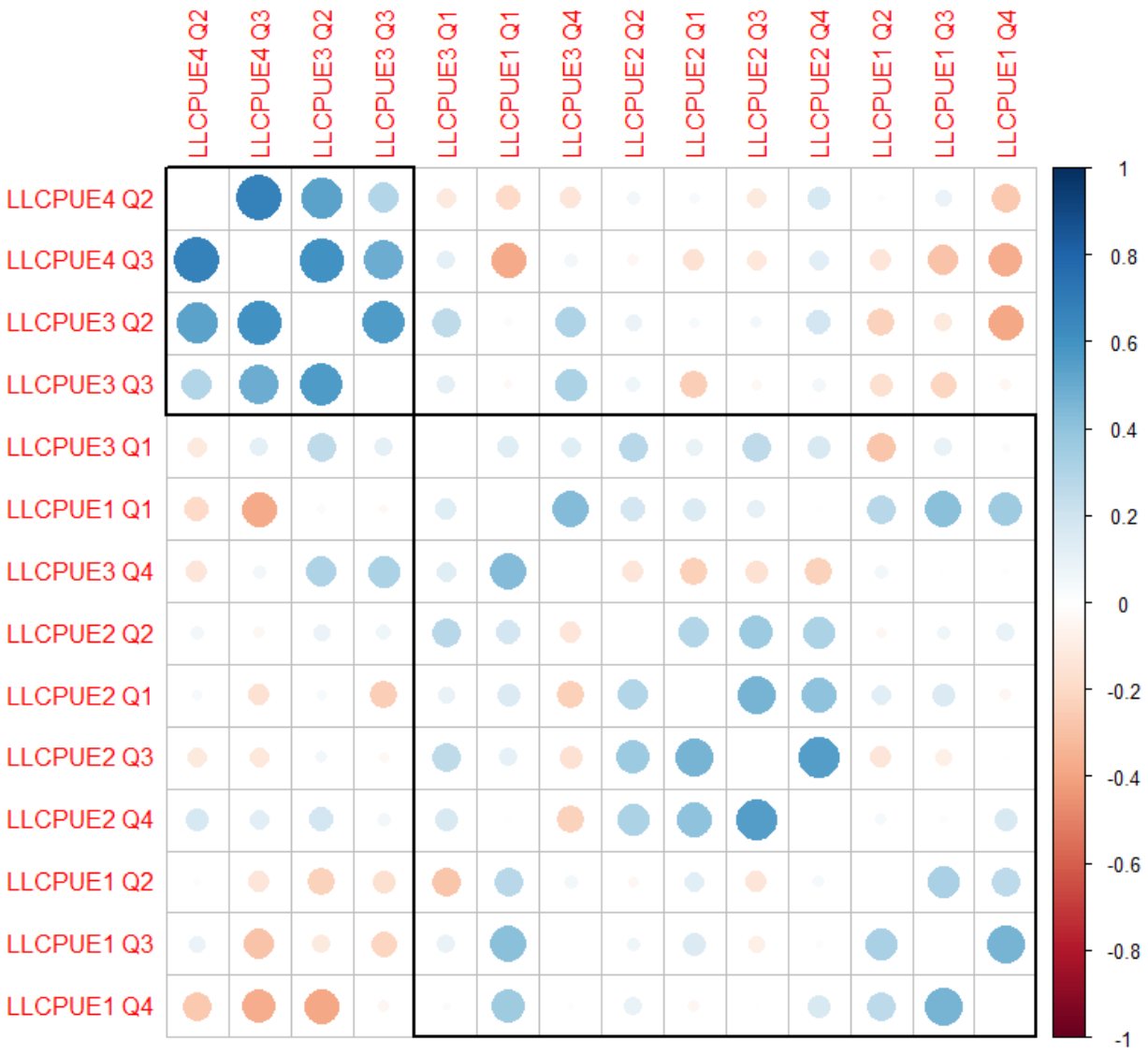

Figure 9. Correlation matrix for CPUE indices available for the Indian Ocean blue shark. Blue indicates positive and red negative correlations. The order of the indices and the rectangular boxes are chosen based on a hierarchical cluster analysis using a set of dissimilarities.

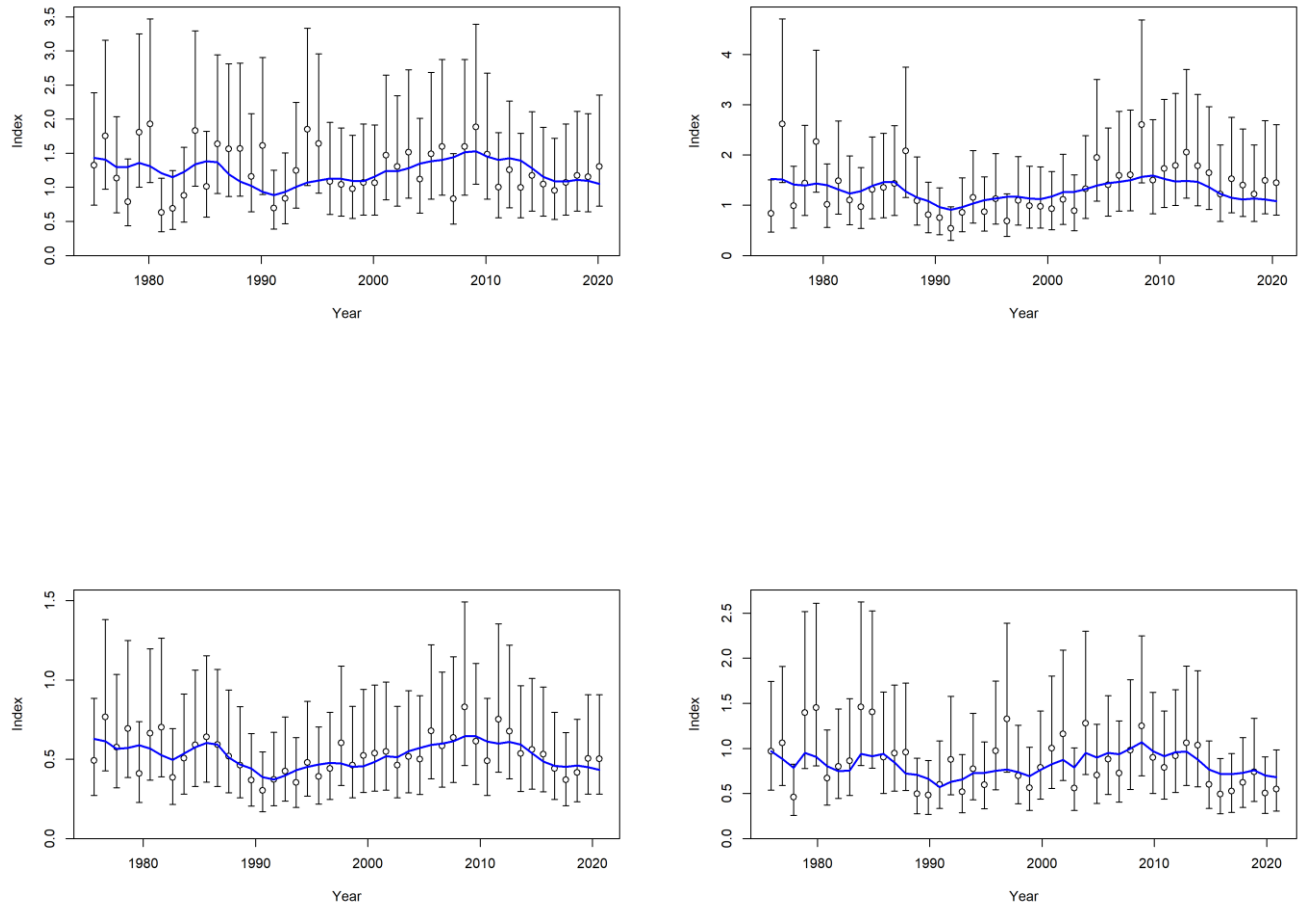

Figure 10: Diagnostic case fit to the CPUE series, from region 3. The top left panel is quarter 1, the top right is quarter 2, the bottom left is quarter 3, the bottom right is quarter 4.

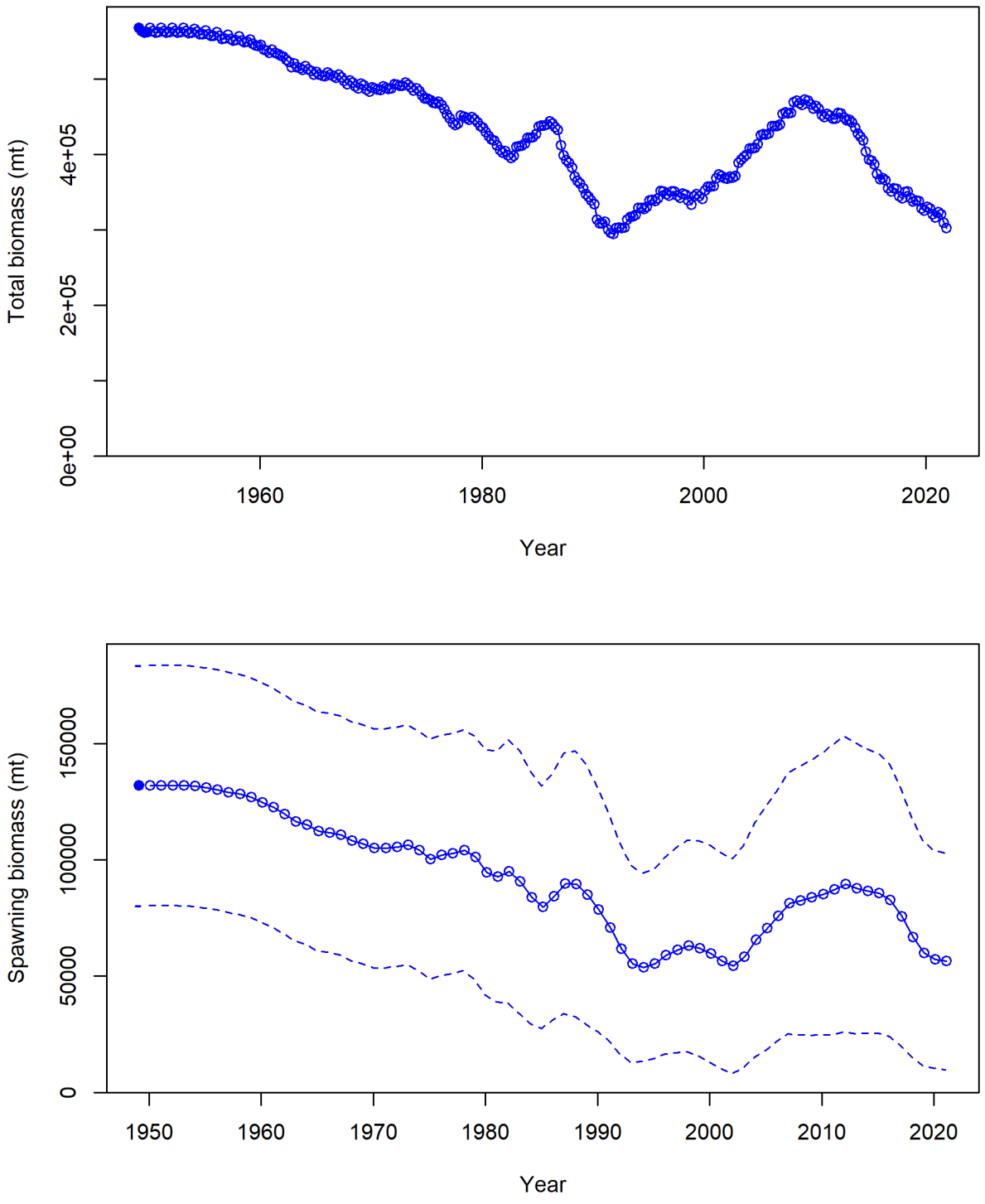

Figure 11: Total biomass (top) and spawning biomass for the diagnostic case parameterization model. The filled dot represents the pre-model estimate of unfished biomass.

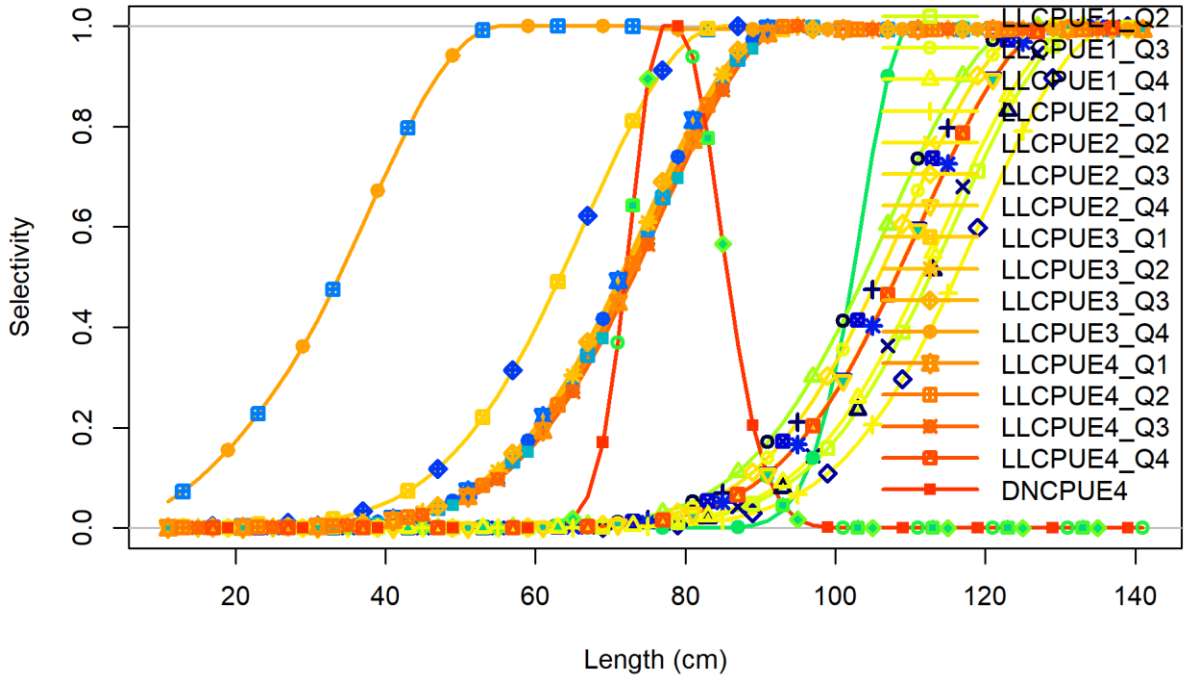

Figure 12: Selectivity curves estimated from the diagnostic case model for the assessment of albacore tuna in the Indian Ocean.

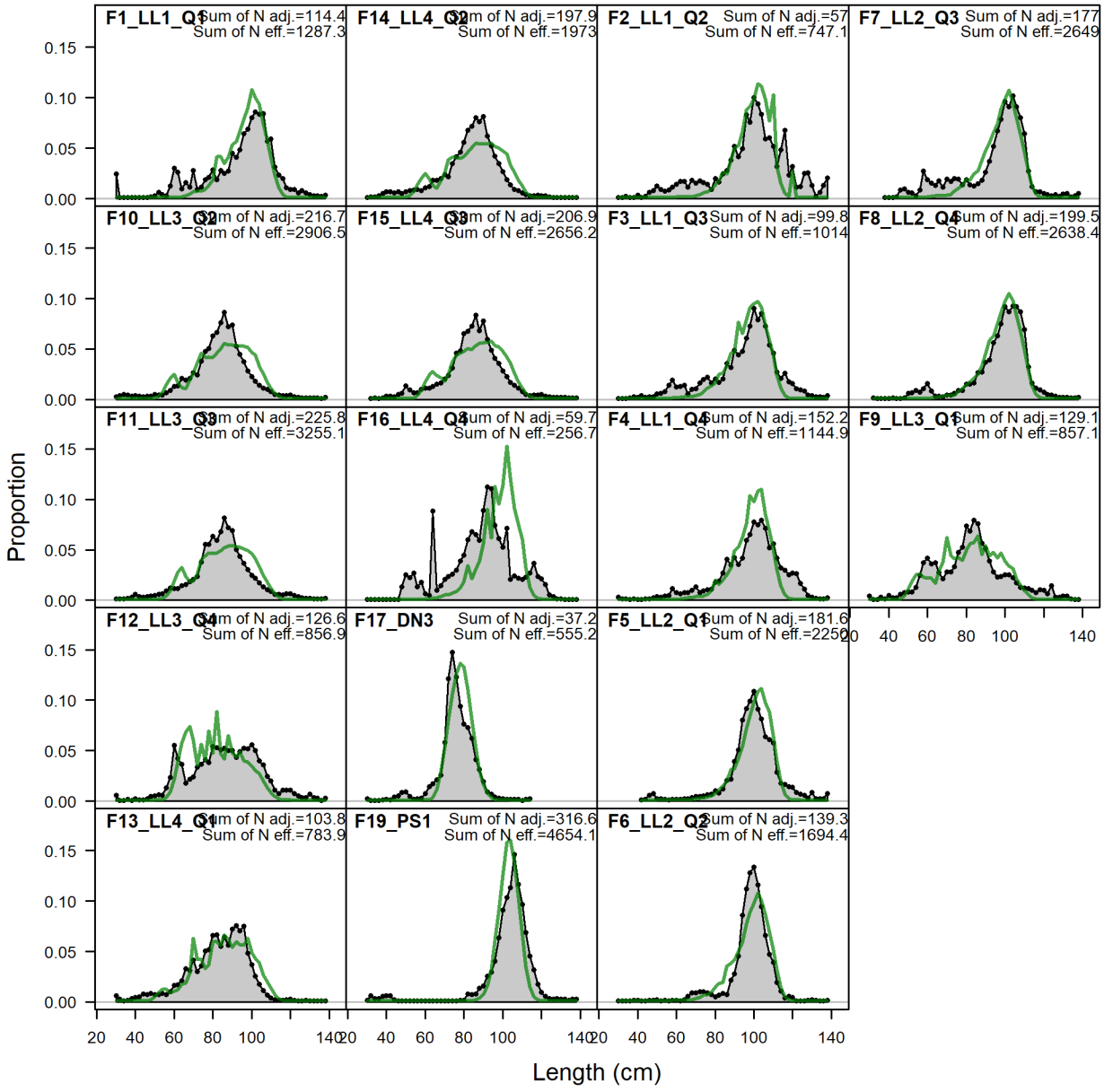

Figure 13 Fit to the length frequency data for the diagnostic case model for the assessment of albacore tuna in the Indian Ocean.

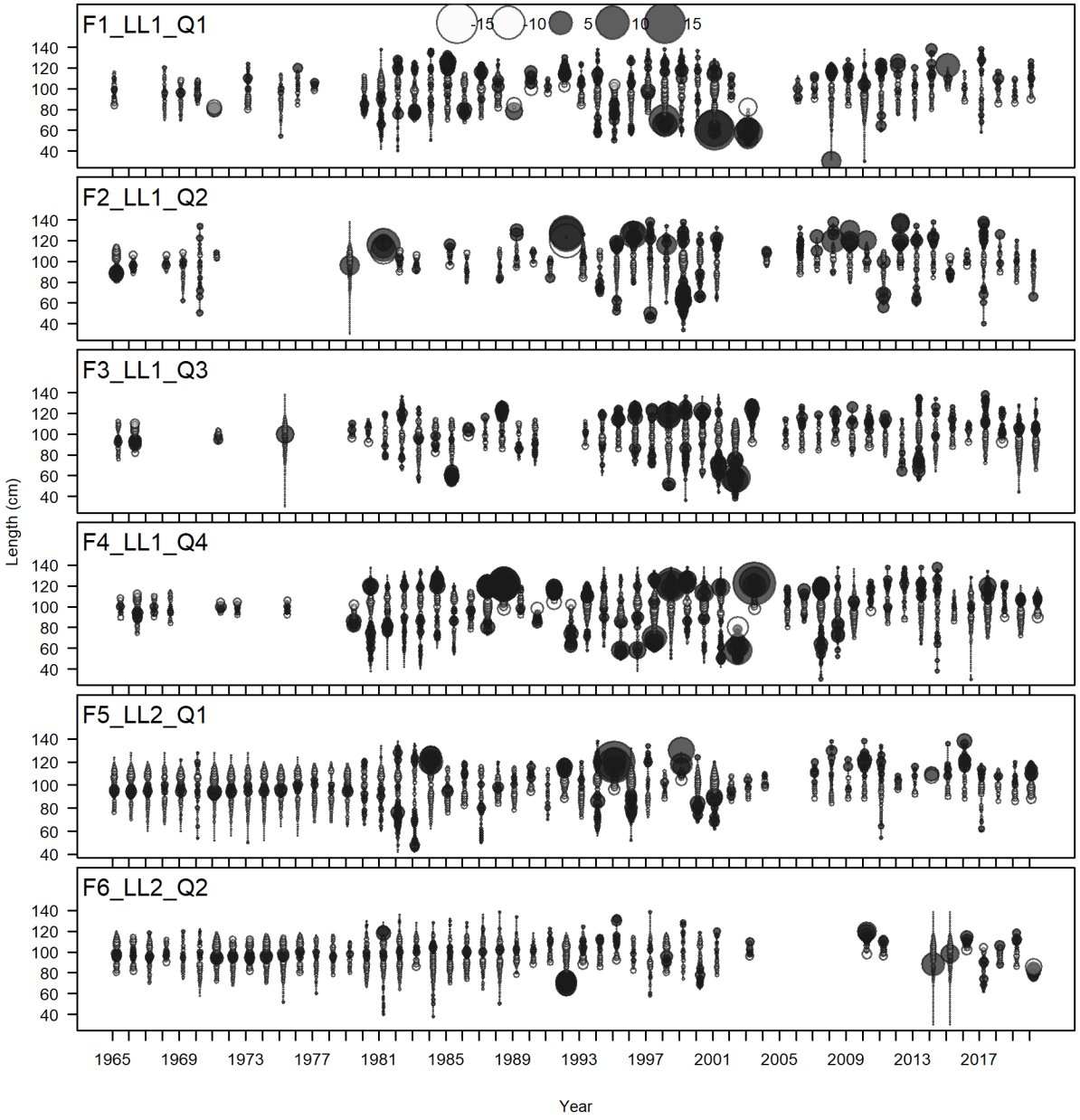

Figure 14. Residuals from the fit to the length frequency data for the diagnostic case model for the assessment of albacore in the Indian Ocean, Fleets 1-6. Closed bubbles are positive residuals and open bubbles are negative residuals, bubble sizes are scaled to maximum within each panel. Thus, comparisons across panels should focus on patterns, not bubble sizes.

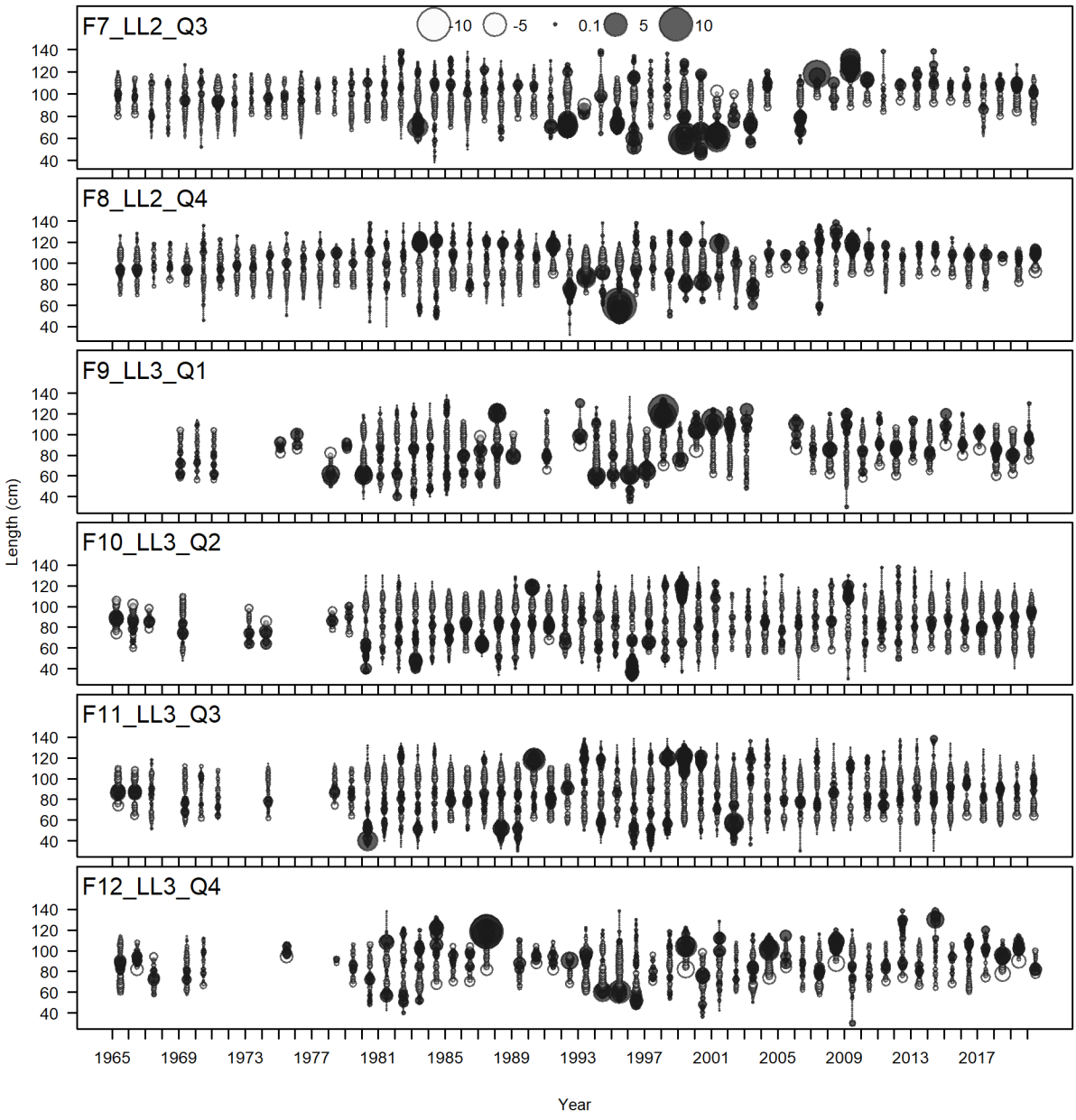

Figure 15. Residuals from the fit to the length frequency data for the diagnostic case model for the assessment of albacore in the Indian Ocean, Fleets 7-12. Closed bubbles are positive residuals and open bubbles are negative residuals, bubble sizes are scaled to maximum within each panel. Thus, comparisons across panels should focus on patterns, not bubble sizes.

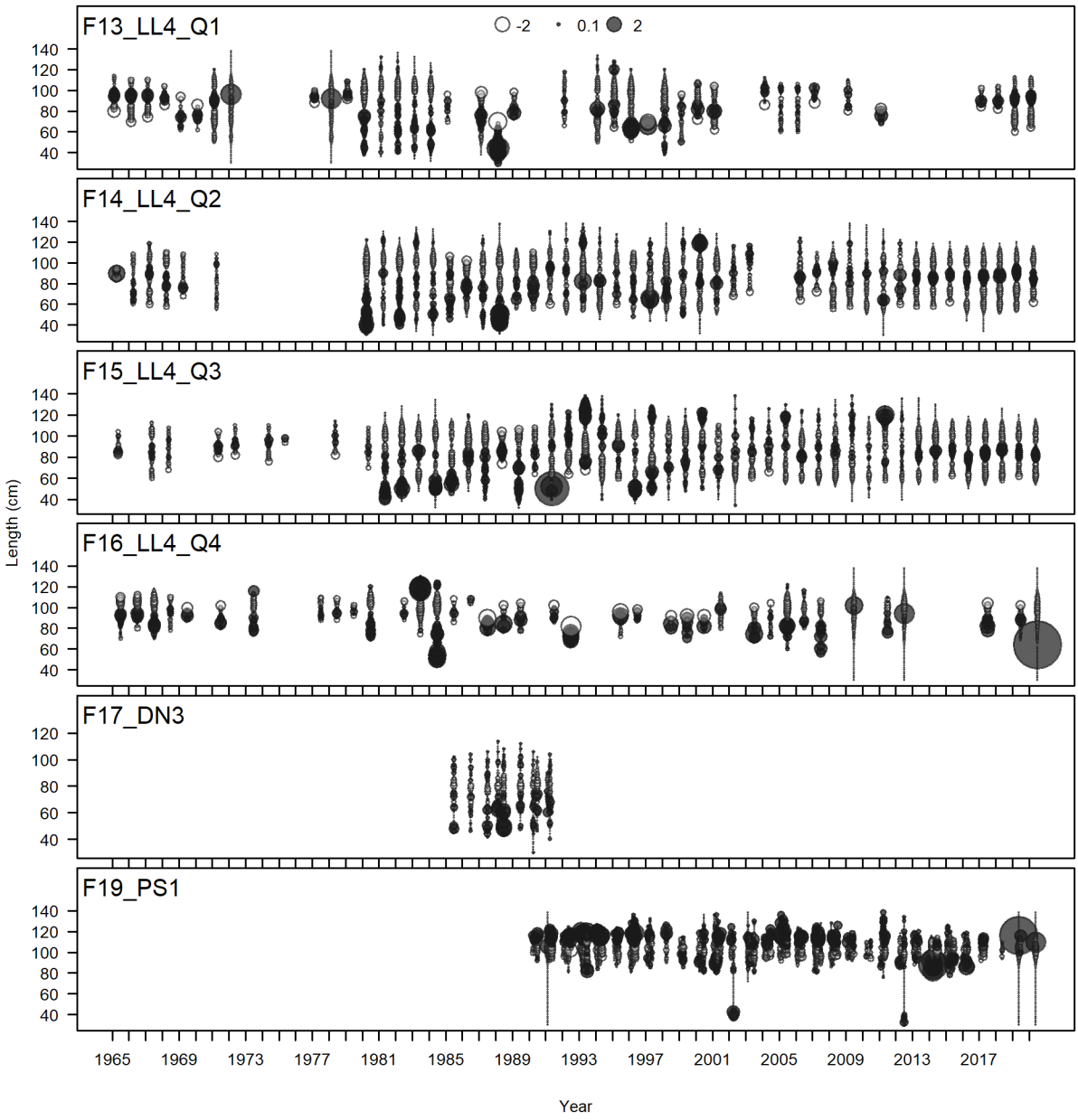

Figure 16 Residuals from the fit to the length frequency data for the diagnostic case model for the assessment of albacore in the Indian Ocean, Fleets 13-19. Closed bubbles are positive residuals and open bubbles are negative residuals, bubble sizes are scaled to maximum within each panel. Thus, comparisons across panels should focus on patterns, not bubble sizes.

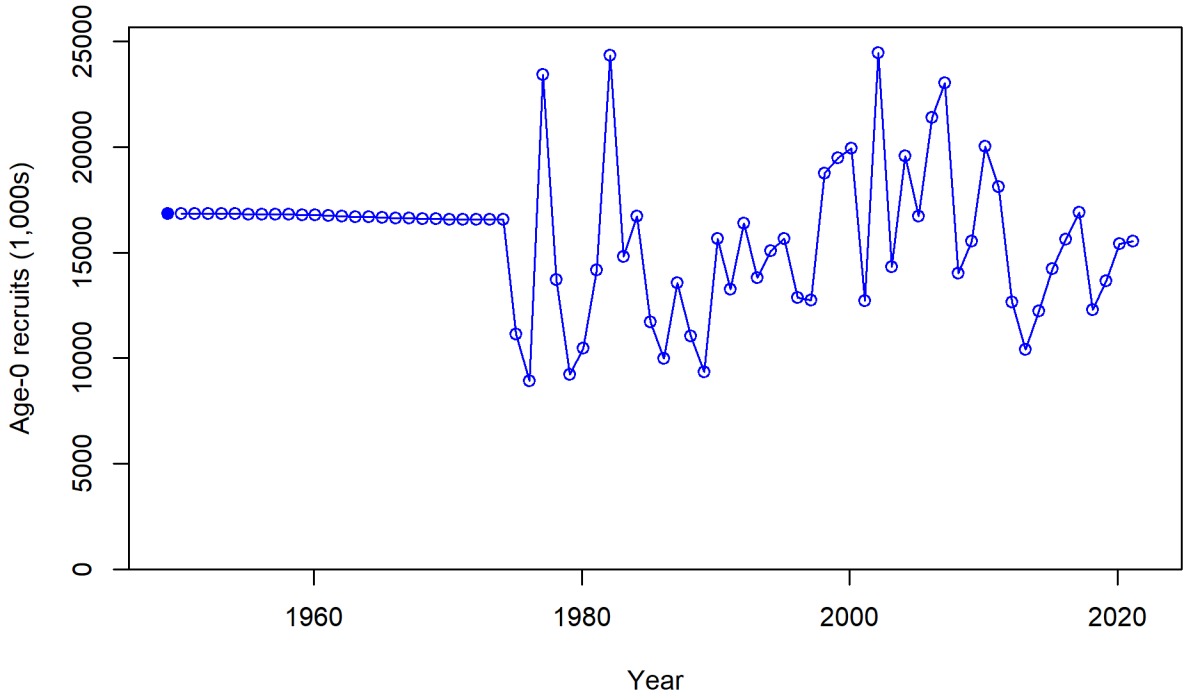

Figure 17. Estimated recruitment including the estimate of virgin recruitment (filled circle at the start of the time series) for the diagnostic case model for the assessment of albacore tuna in the Indian Ocean.

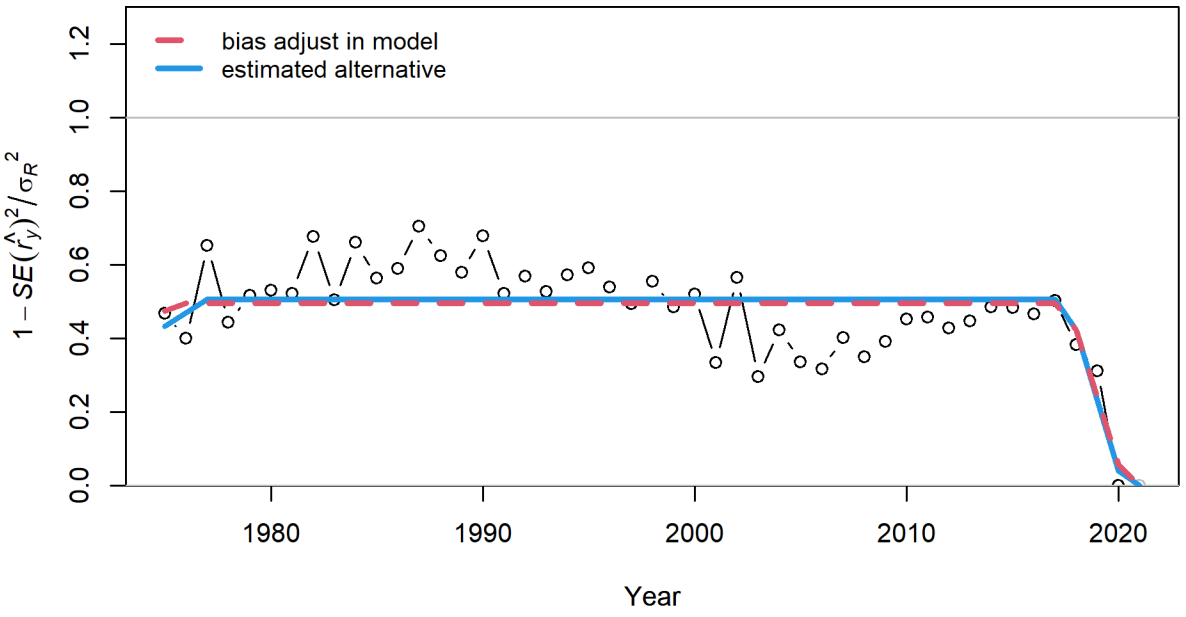

Figure 18 .Estimated bias adjustment in the model.

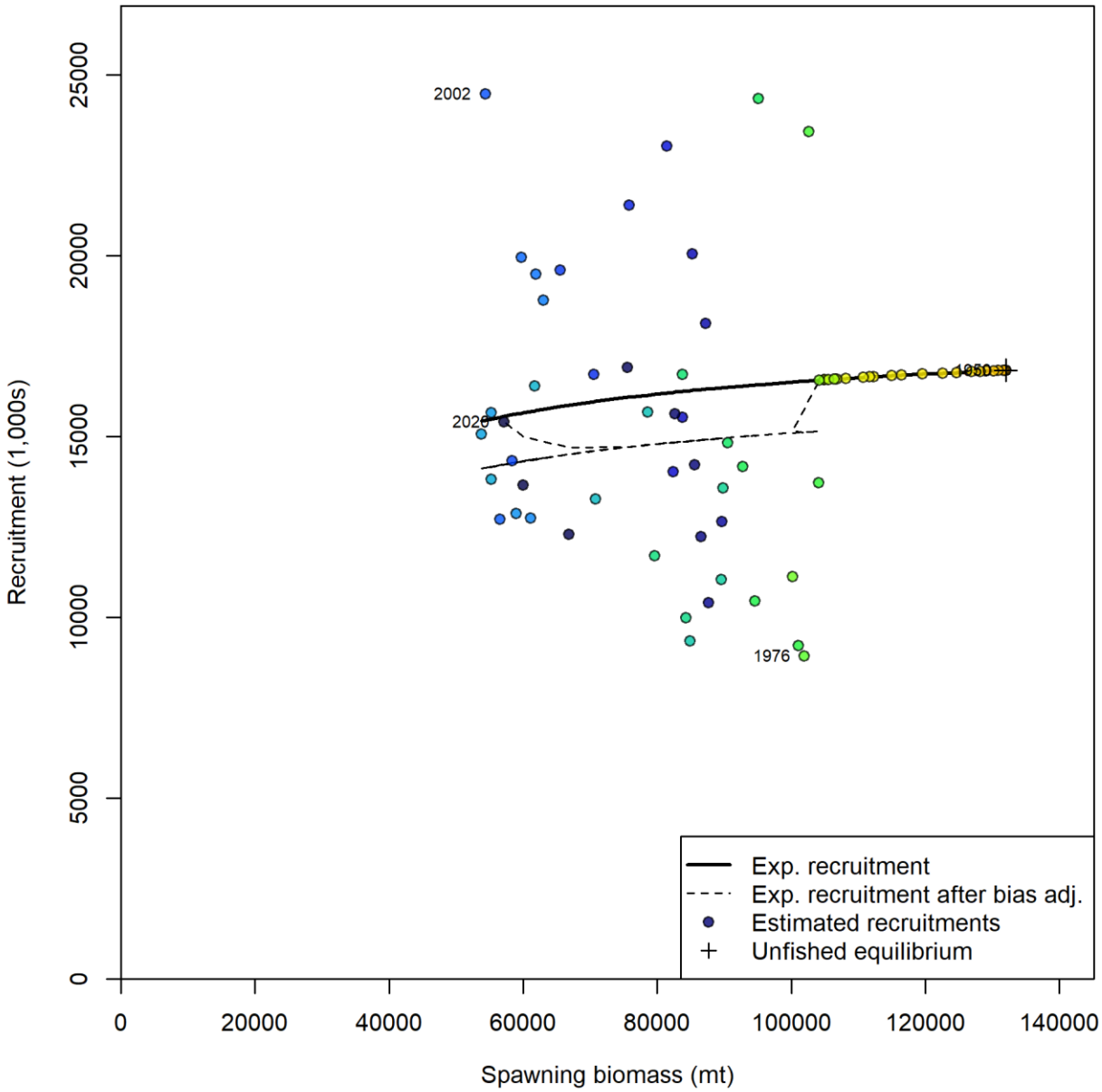

Figure 19. Stock recruitment curve used in the assessment and time series of estimates of recruitment deviations (colored points).

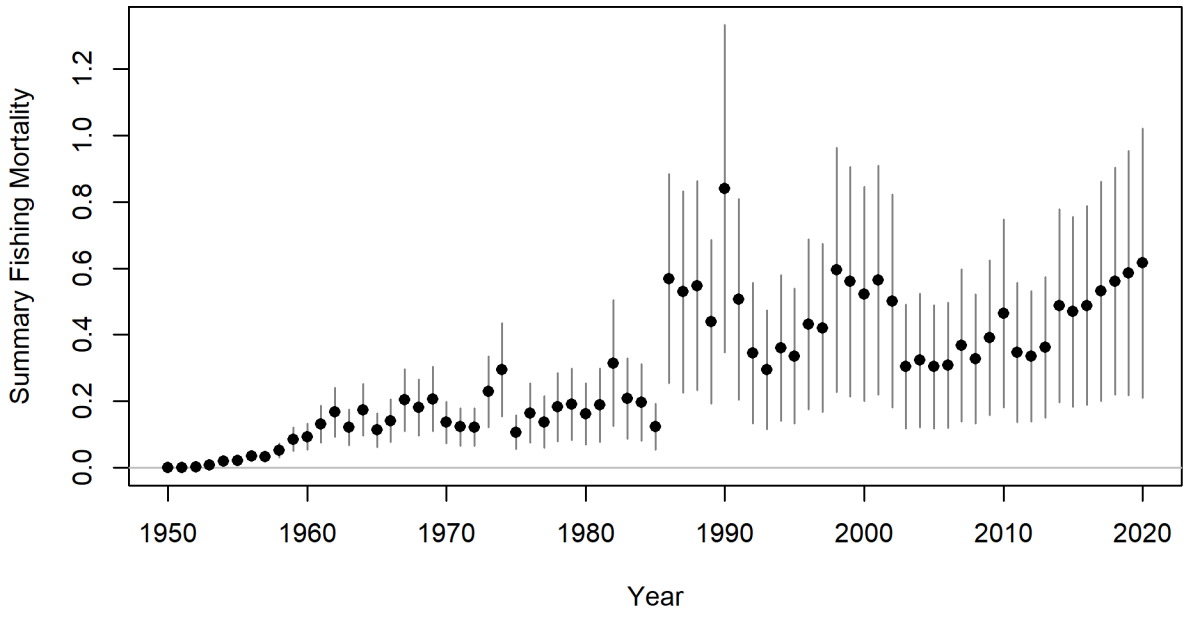

Figure 20. Estimated total fishing mortality/FMSY.

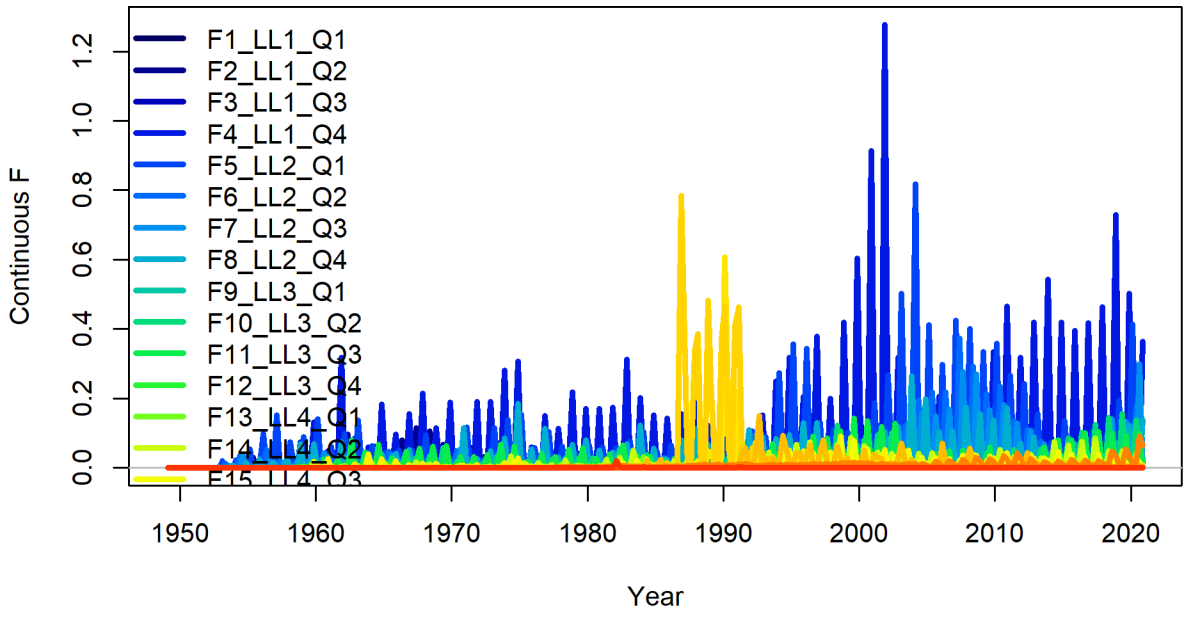

Figure 21. Estimated fleet specific fishing mortality by year for the base case model configuration.

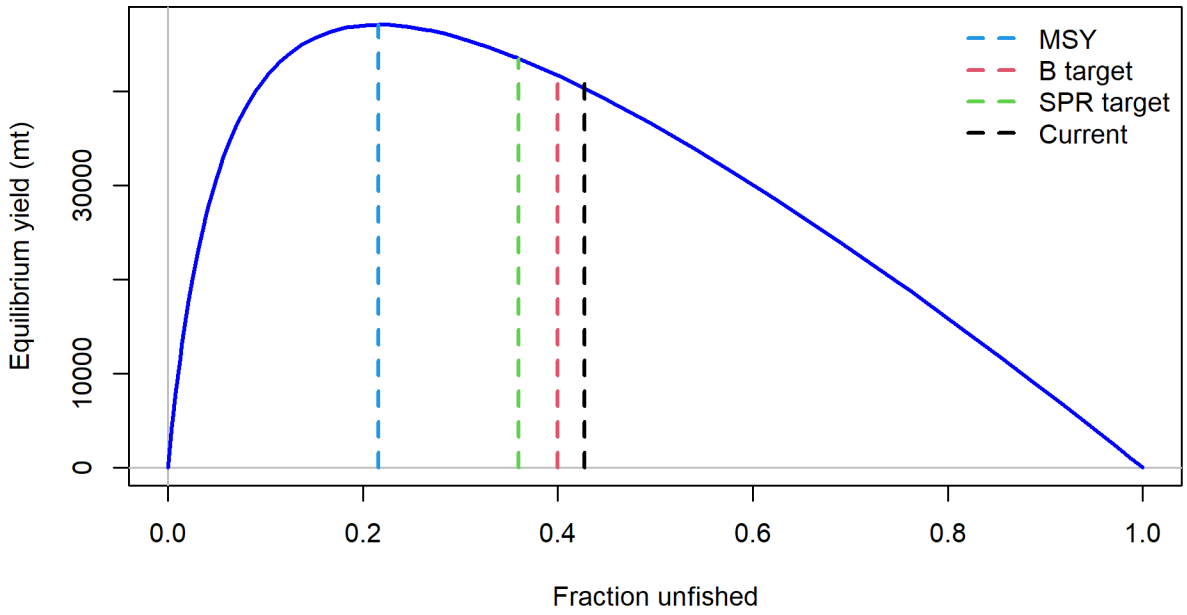

Figure 22. Equilibrium yield curve for the diagnostic case model for the assessment of albacore tuna in the Indian Ocean.

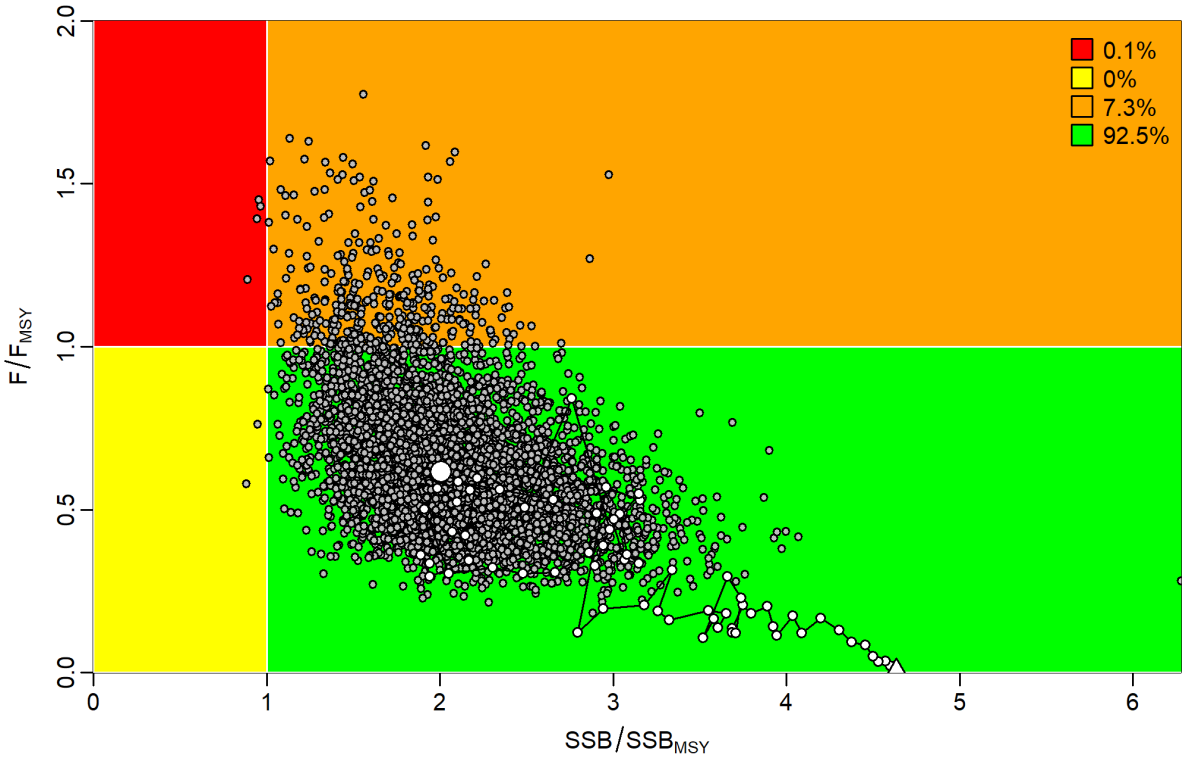

Figure 23. Kobe plot of the annual stock status

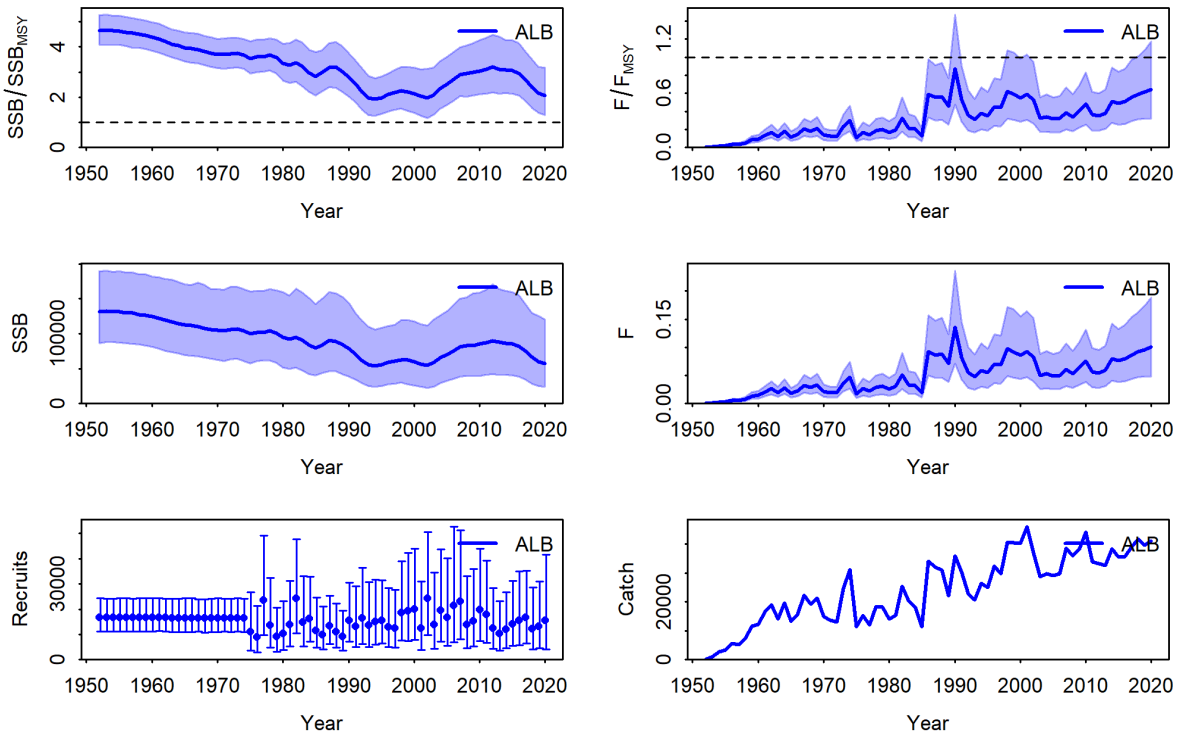

Figure 24. Estimated timeseries based on the MVLN approximation

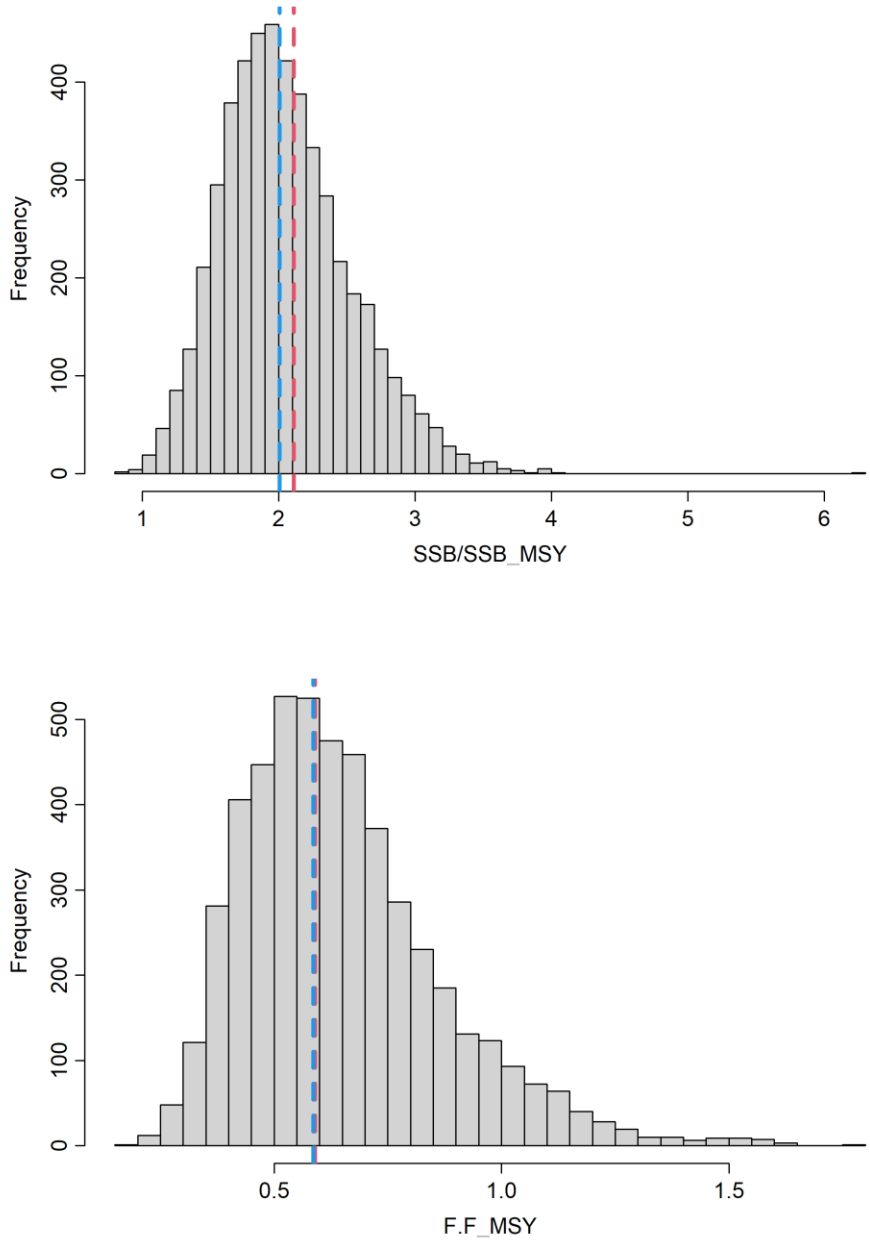

Figure 25. Estimated spawning biomass in 2020 relative to MSY (SSB2020/SSBMSY, top panel) and estimated total fishing mortality in 2019 relative to MSY (F2020/FMSY, bottom panel) for the diagnostic case model configuration, dashed lines indicate the 50th quantile.

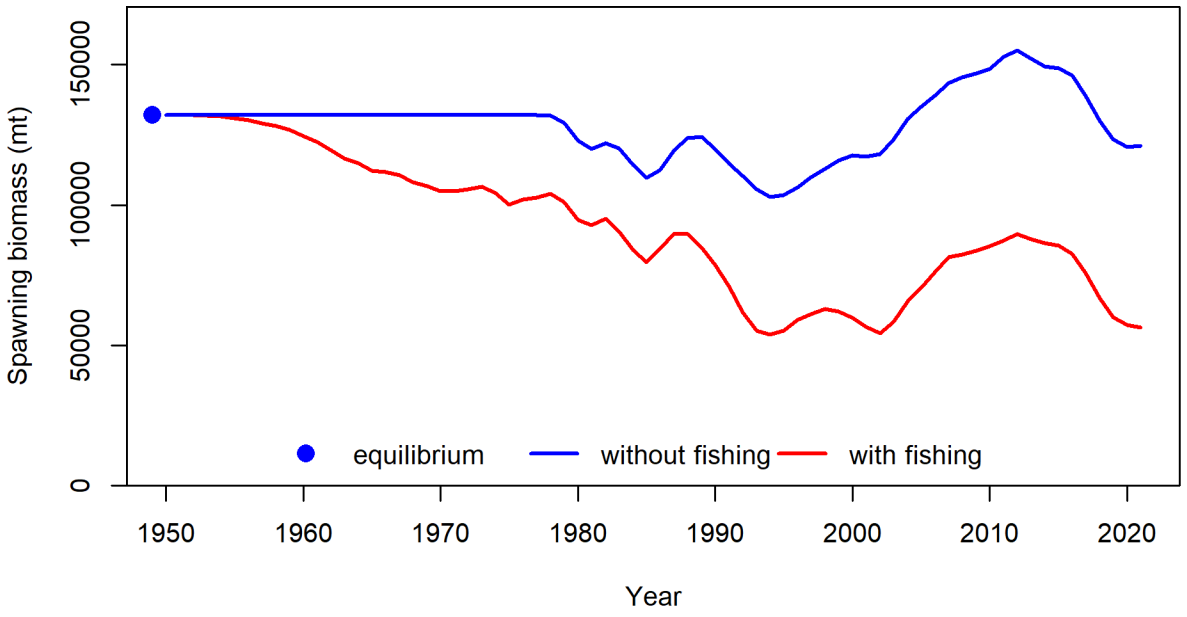

Figure 26. Dynamic B0 plot showing the spawning biomass under conditions no fishing and assuming the input catch series.

ANNEX 1

Model Diagnostics

Stock assessment of albacore tuna (*Thunnus alalunga*) in the Indian Ocean using SS3.

10 Annex 1. Diagnostic Tests for the Diagnostic Case SS3 Model

Diagnostic tests are important in determining the robustness of estimates for management advice in integrated stock assessment models. The diagnostics tests included in are based on diagnostics prepared for the previous assessment as well as recently developed methods (Carvallo et al. 2021). Here we present the model diagnostics for the diagnostic model run presented in the main assessment analysis.

10.1 Goodness of fit

Residual and Hierarchal cluster analysis

Data misfit often stems from inappropriate model structure, particularly with respect to the information contained in the CPUE data. By including divergent CPUE trends model mist fit and The CPUE time series are plotted in Figure A1, along with a lowess smoother fitted to CPUE each year using a general additive model (GAM) to compare trends for the submitted CPUEs.

The overall trend for fit to the indices is an initial decrease, a more dramatic decrease beginning in the early 1990s, following the Japanese CPUE, with a decrease through the 2000's and a nearly stable or slightly declining trend during the 2010-2019 timeframe.

Residuals from the smoother fits to CPUE are compared in Figure A2 to look at deviations from the overall trends. This allows conflicts between indices (e.g. highlighted by patterns in the residuals) to be identified. For example, in both the EU Portugal and the EU France (Reunion) time-series, the early part is mostly positive and the latter part is mostly a series negative residuals indicating that these time-series do not follow the overall trend, and provide evidence of a more rapidly decreasing trend in the stock trajectory in recent years than the overall trend. In contrast, The Japanese and South African series provide evidence of a more gradually increasing trend in the stock trajectory in recent years than the overall trend.

Correlations between indices are evaluated in Figure A3. The lower triangle shows the pairwise scatter plots between indices with a regression line, the upper triangle provides the correlation coefficients, and the diagonal provides the range of observations. A single influential point may cause a strong spurious correlation, so it is important to look at the plots as well as the correlation coefficients. Also, a strong correlation could be found by chance if two series only overlap for a few years.

A hierarchical cluster analysis evaluated for the indices using a set of dissimilarities is provided in Figure A4. If indices represent the same stock components, then it is reasonable to expect them to be correlated. If indices are not correlated or are negatively correlated, i.e. they show conflicting trends, then this may result in poor fits to the data and bias in the parameter estimates obtained within a stock assessment model. Therefore, the correlations can be used to select groups of indices that represent a common hypothesis about the evolution of the stock (ICCAT 2017).

The hierarchical cluster analysis (HCA) identified two groupings of time-series. The first group includes Portugal and Reunion. This group is characterized by time-series which are highly correlated with each other and which have some highly negative correlations with some time-series not included in the group. The second group includes the other indices. This group is characterized by time-series which are less highly correlated with each other or slightly negatively correlated with each other. Notably the HCA identified that the Portuguese CPUE was positively correlated with all the other CPUE series except the Taiwanese CPUE. The Taiwanese CPUE was negatively correlated with all of the other CPUE series, though only minimally with the CPUE series from Reunion and Japan. The South African CPUE series was positively correlated with only the Portuguese and Japanese CPUE series.

Cross-correlations for the CPUE series are plotted in Figure A5 (i.e., the correlations between series when they are lagged by -10 to 10 years). The diagonals show the autocorrelations of an index lagged against itself. The lag refers to how far the series are offset, and its sign determines which series is shifted. Note that as the lag increases, the number of possible matches decreases because the series overlap at the ends and do not overlap. The value of the lag with the highest correlation coefficient represents the best fit between the two series.

You can plot the correlation coefficients versus lag to look for periodicities in the original time series. If the data is periodic, there will be an oscillation in the correlation coefficients with lag. They will be positive and have large values when the two series are in phase, and negative with large values when the two series are out of phase (peaks aligned with troughs).

Runs test and joint residual plots.
The goodness of fit of the model can be used as an indication of whether there is presence of significant model misspecification. Models that do not fit the data should be considered suspect, and further investigated. Here we use residual plots to investigate the trends and patterns in the data over time. Temporal correlation (autocorrelation) can drive bias and drift in the model estimates over time. A runs test (Wald and Wolfowitz, 1940) can test for randomness in a data sequence, such as model residuals (Carvalho et al., 2021). Residuals can also be investigated along side the root mean square error (RMSE, Carvalho et al., 2017), and a joint residual plot (Winker et al., 2018), which can highlight the systematically auto-correlated residual patterns.

10.2 Model consistency

R0 Profile
Use of a likelihood component profile on the a global scaling parameter (or other parameter) has been identified as a key model diagnostic to identify the influence of information sources on model estimates (Carvalho et al. 2017, Ichinokawa et al., 2014; Lee et al., 2014; Wang et al., 2014). Here the equilibrium recruitment parameter, R0, is used because it represents an ideal global scaling parameter given that unfished (virgin) recruitment is proportional to unfished biomass (Carvalho et al 2021, Lee et al., 2014; Maunder and Piner, 2015; Wang et al., 2014).

A relatively large change in negative log-likelihood units along the profile suggests a relatively informative data source for that particular model. Close association in the location of the minimum negative log-likelihood along the profile between data sources suggest that model consistency, and lack of conflict in the data. Figures A9-A11 show the profile likelihoods of R0 for the overall, CPUE and length components of the model. Figure A12 shows the fit to the CPUE series for the range of LN(R0) assumed. The likelihood profiles show that overall the *LN(R0)* parameter is well estimated, led by the length likelihood then the index likelihood and then the recruitment. Interestingly the lower edge of the likelihood is better defined (steeper) for all three components than the upper (Figure AA9). The fleet indices are generally in agreement (Figure AA10), however the length data shows different minimums (Figure A11), which is consistent with the fleets that encounter different components of the stock. The fits to the CPUE series at different values of *LN(R0)* show that values of approximately 7.5 fit the series (Figure A12).

Age-Structured Production Model (ASPM)

This diagnostic can help evaluate whether the catch and CPUE data give evidence for a production function within the model (Carvalho 2017). Overall the ASPM evaluates whether the effect of surplus production and observed catches alone could explain trends in the CPUE, in contrast to a more complex model (i.e. SS3) that incorporates annual recruitment deviations to improve the fit (Carvalho et al 2021). Maunder and Piner (2017) note that if the ASPM fits well to the indices of abundance with contrast the production function is likely to drive the stock dynamics and the indices will provide information about absolute abundance (Minte-Vera et al., 2017). Figure A13 shows that the biomass trajectories for both models (ASPM and the diagnostic) follow the same trend and that the estimates of *LN(R0)* are comparable. The fits to the indices are shown in figure Figure A14 and indicate an overall good fit, indicating that the information content in the data is sufficient.

Retrospective analysis

Retrospective analysis is common in fisheries stock assessment to check the consistency of model estimates (Brooks and Legault, 2016; Carvalho et al., 2017; Hurtado-Ferro et al., 2015; Miller and Legault, 2017). A retrospective analysis is carried out by sequentially deleting a number of years of day (i.e. from 0 to 7) and re-running the model. Comparisons of model estimates from the full time-series and the truncated time-series can illuminate the bias and accuracy of the modelled quantities. Statistical analysis in the form of calculating the retrospective bias, rho (ρM, Mohn (1999), is common, with values between -0.15 and 0.2 being considered indicative of no bias. Figure A15 shows the analysis of spawning stock biomass (SSB) and fishing mortality estimates for Indian Ocean albacore tunaalong with the Mohn's rho which indicates no retrospective bias. Forecasting the next year based on the retrospective analysis shows similar analysis (Figure A16).

10.3 Prediction Skill

Kell et al. (2016) proposed the hindcasting cross-validation technique (HCXval) where observations are compared to their predicted future values. The key concept behind the HCXval approach is 'prediction skill', which is defined as any measure of the accuracy of a forecasted value to the actual observed that is not known by the model (Kell et al., 2021). The difference, which is referred to as the 'prediction residual' (Michaelsen, 1987) can be evaluated by the mean absolute scaled error (MASE; Hyndman and Koehler, 2006). Carvalho et al note that a MASE score > 1 indicates that the average model forecasts are worse than a random walk. Conversely, a MASE score of 0.5 indicates that the model forecasts twice as accurately as a naïve baseline prediction; i.e. the model has prediction skill. For the CPUE series that constitute the diagnostic case the MASE values are 1.36, 0.93 and 1.09 for the Japanese, Portugal and Reunion series (respectively, Figure A17), this indicates a mix of poor, good and decent prediction skill.

11 Figures

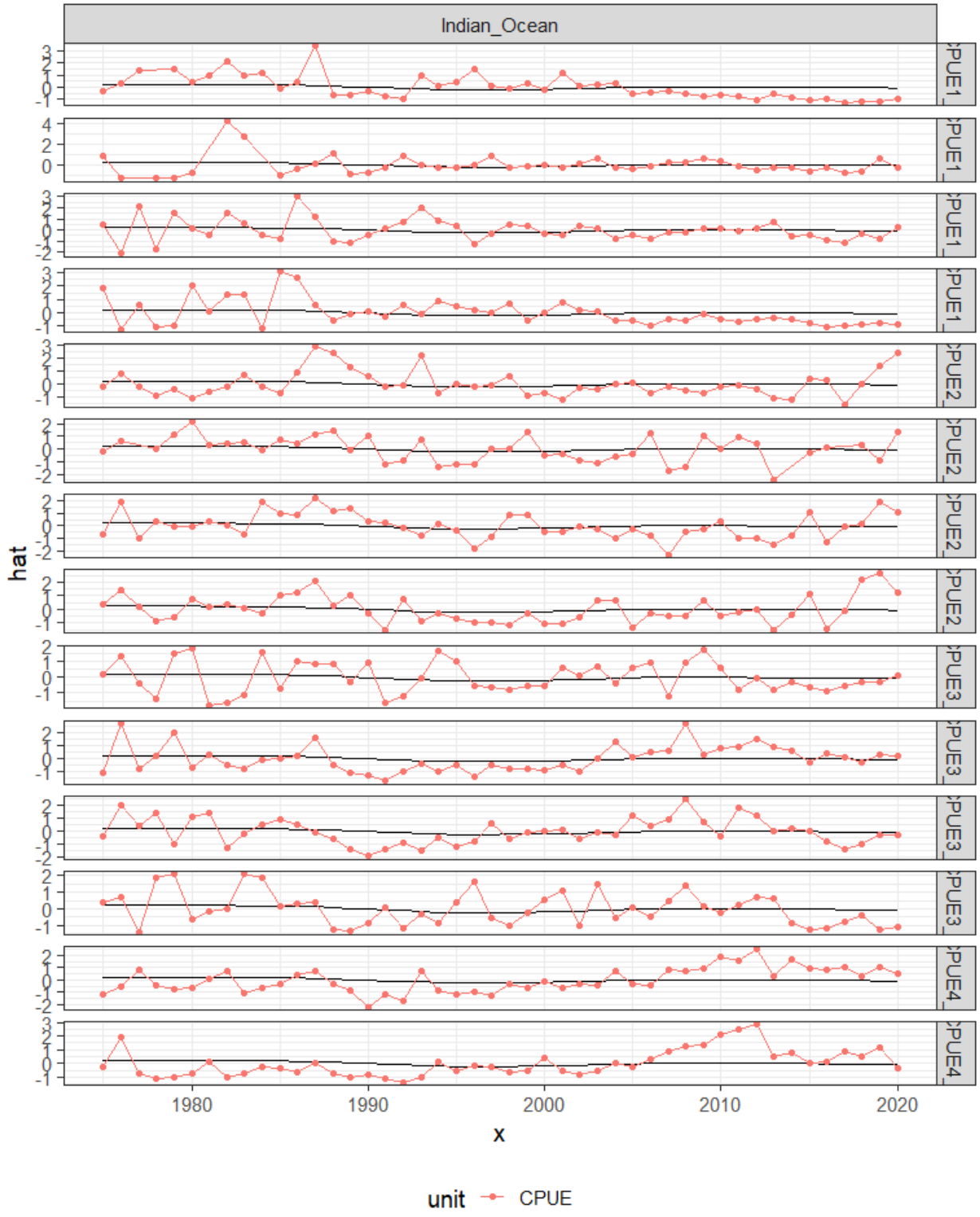

Figure A1. Indian Ocean time series of LL CPUE indices; Points are the standardized values, continuous black lines are a lowess smoother showing the average trend by area (i.e. fitted to year for each area with series as a factor). X-axis is time, Y-axis are the scaled indices.

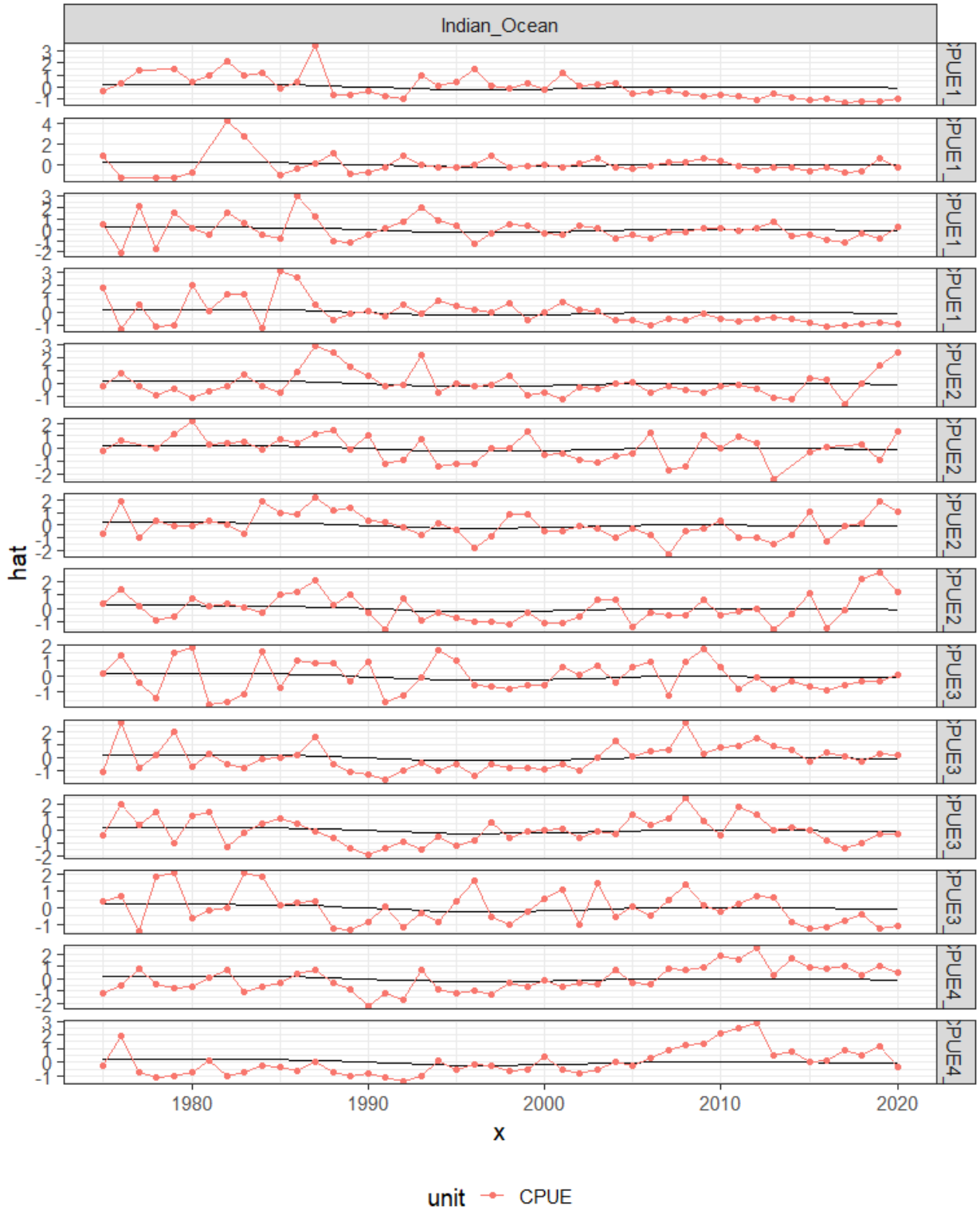

Figure A2. Time series of residuals from the smooth fit to CPUE indices. X-axis is time, Y-axis are the scaled indices.

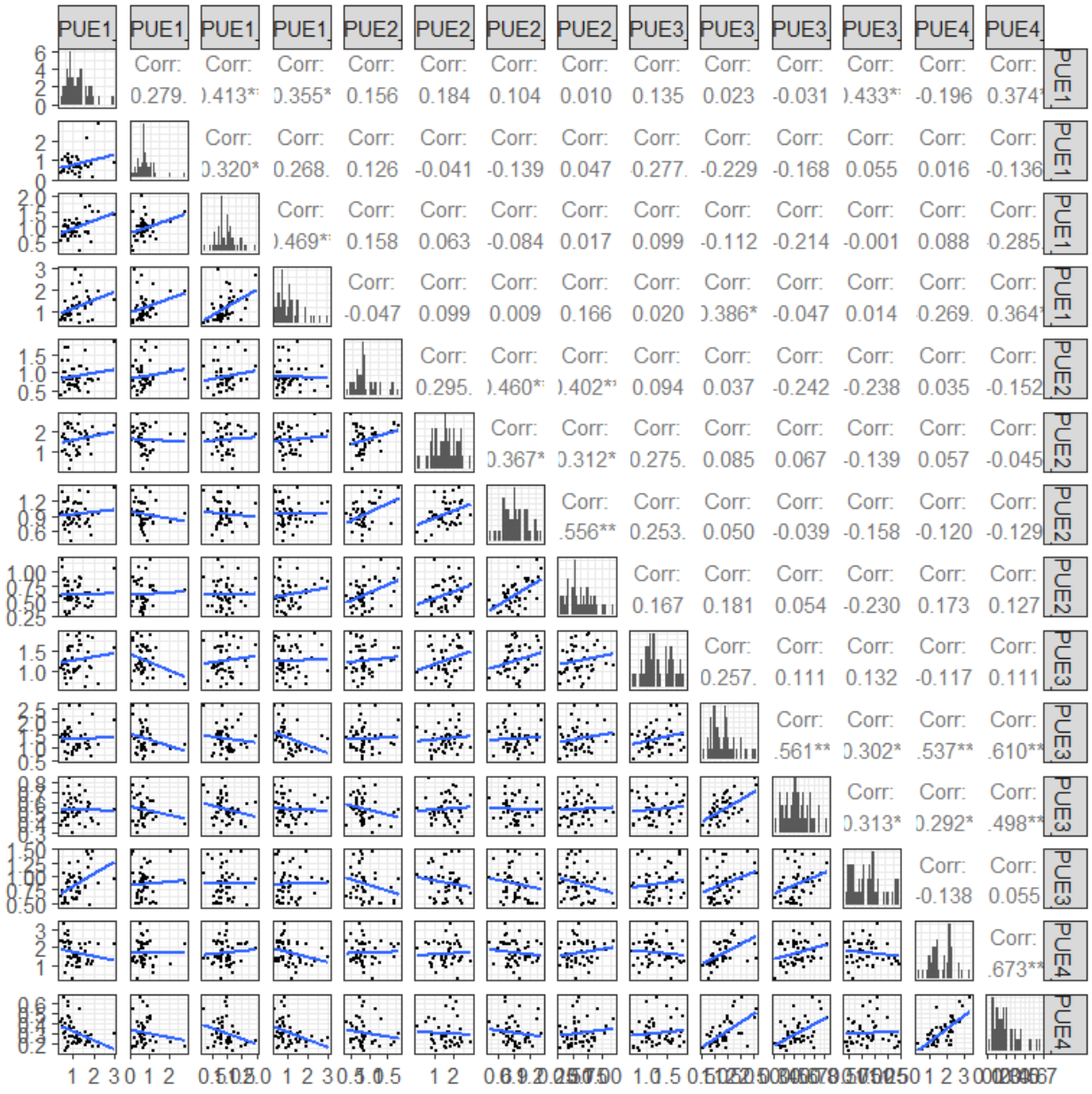

Figure A3. Pairwise scatter plots for CPUE indices. X- and Y-axis are scaled indices.

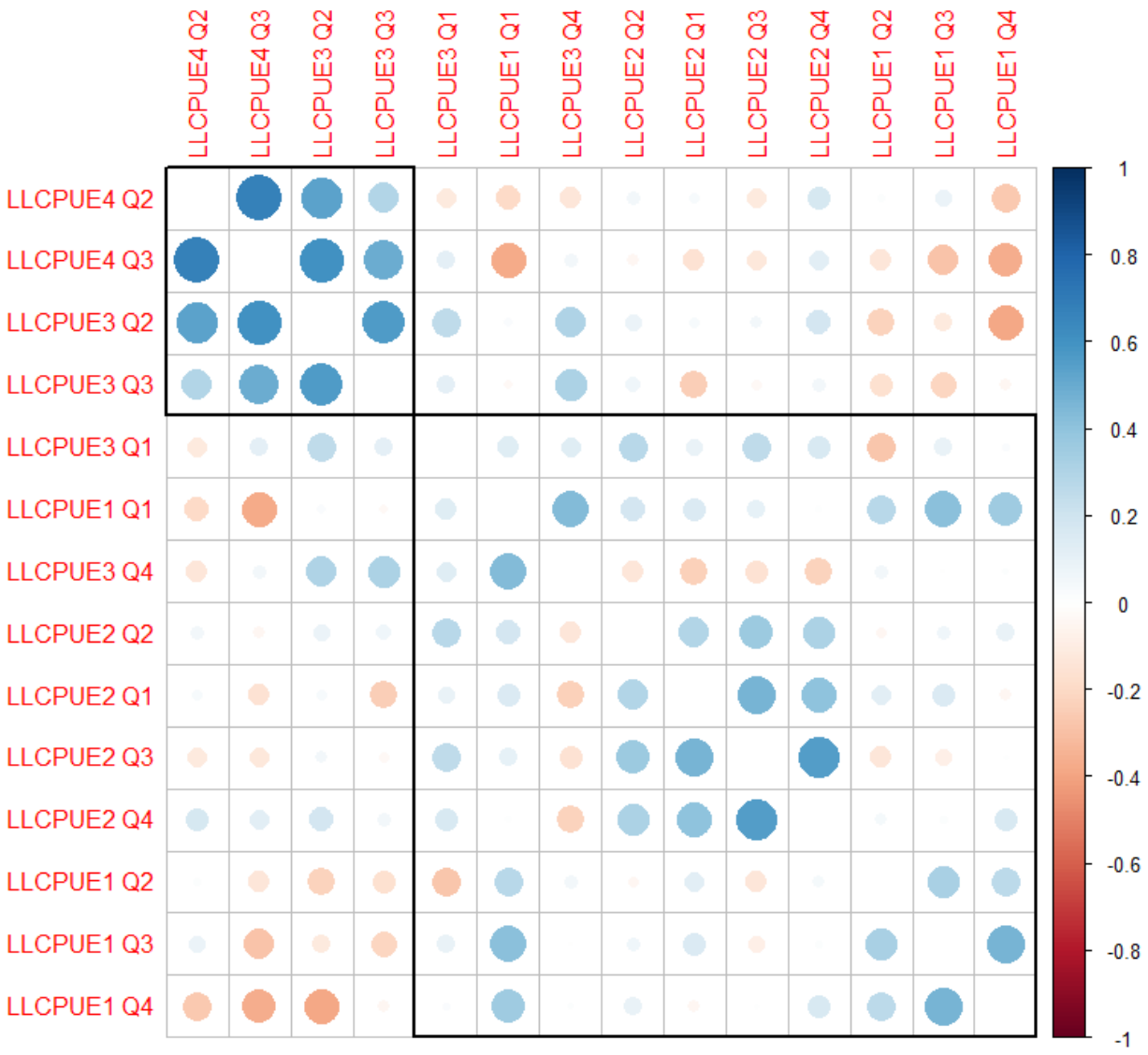

Figure A4. Correlation matrix for CPUE indices; blue indicates positive and red negative correlations, the order of the indices and the rectangular boxes are chosen based on a hierarchical cluster analysis using a set of dissimilarities.

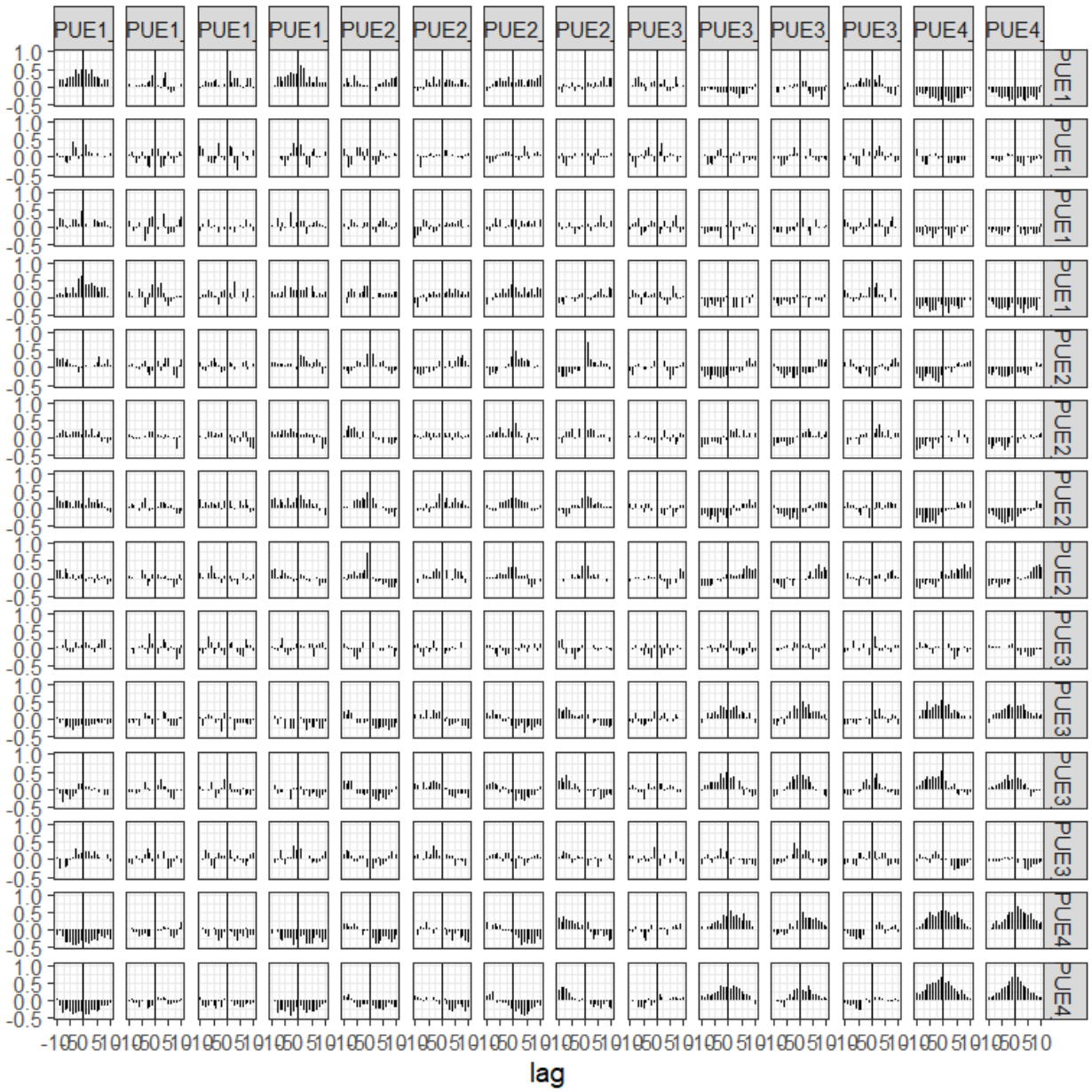

Figure A5 Cross-correlations between CPUE indices to identify lagged correlations (e.g., due to year-class effects). X-axis is lag number, and y-axis is cross-correlation.

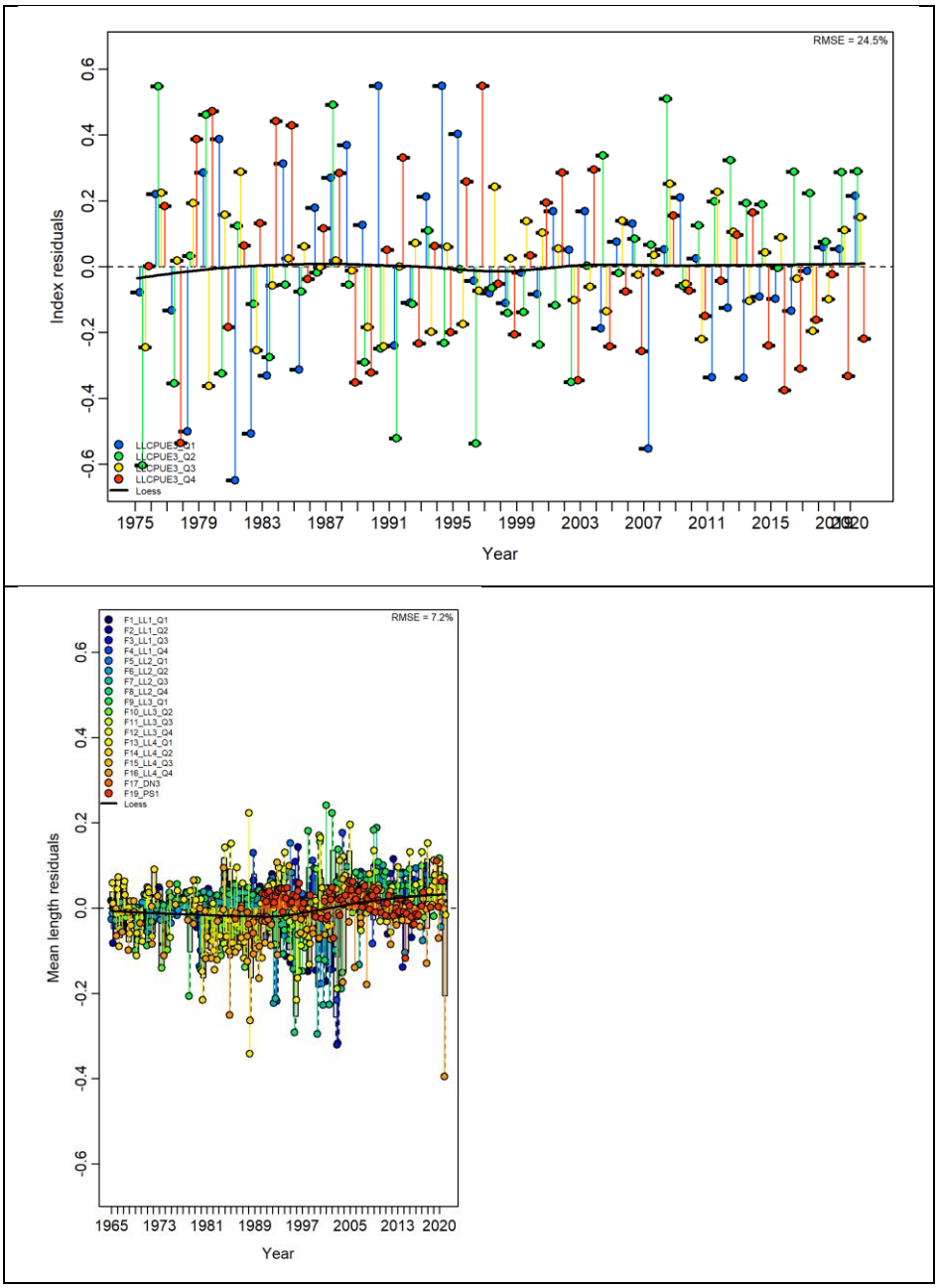

Figure A6. Joint residual plots for CPUE fits (top),. annual mean length estimates for multiple fishing fleets (lower). Vertical lines with points show the residuals (in colors by index), and solid black lines show loess smoother through all residuals. Boxplots indicate the median and quantiles in cases where residuals from the multiple indices are available for any given year. Root-mean squared errors (RMSE) are included in the upper right-hand corner of each plot.

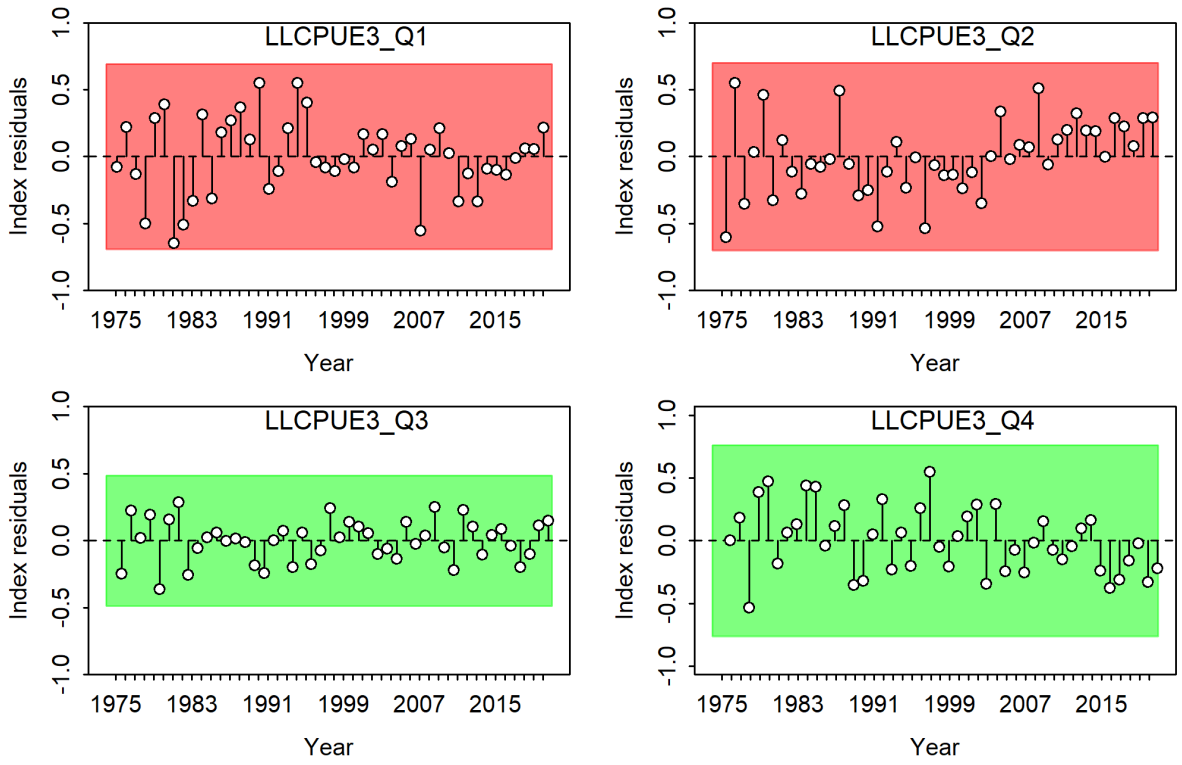

Figure A7. Runs tests results illustrated for three catch-per-unit-effort (CPUE) fits, panels indicate separate quarters. Green shading indicates no evidence ($p \geq 0.05$) and red shading evidence ($p < 0.05$) to reject the hypothesis of a randomly distributed time-series of residuals, respectively. The shaded (green/red) area spans three residual standard deviations to either side from zero, and the red points outside of the shading violate the 'three-sigma limit' for that series.

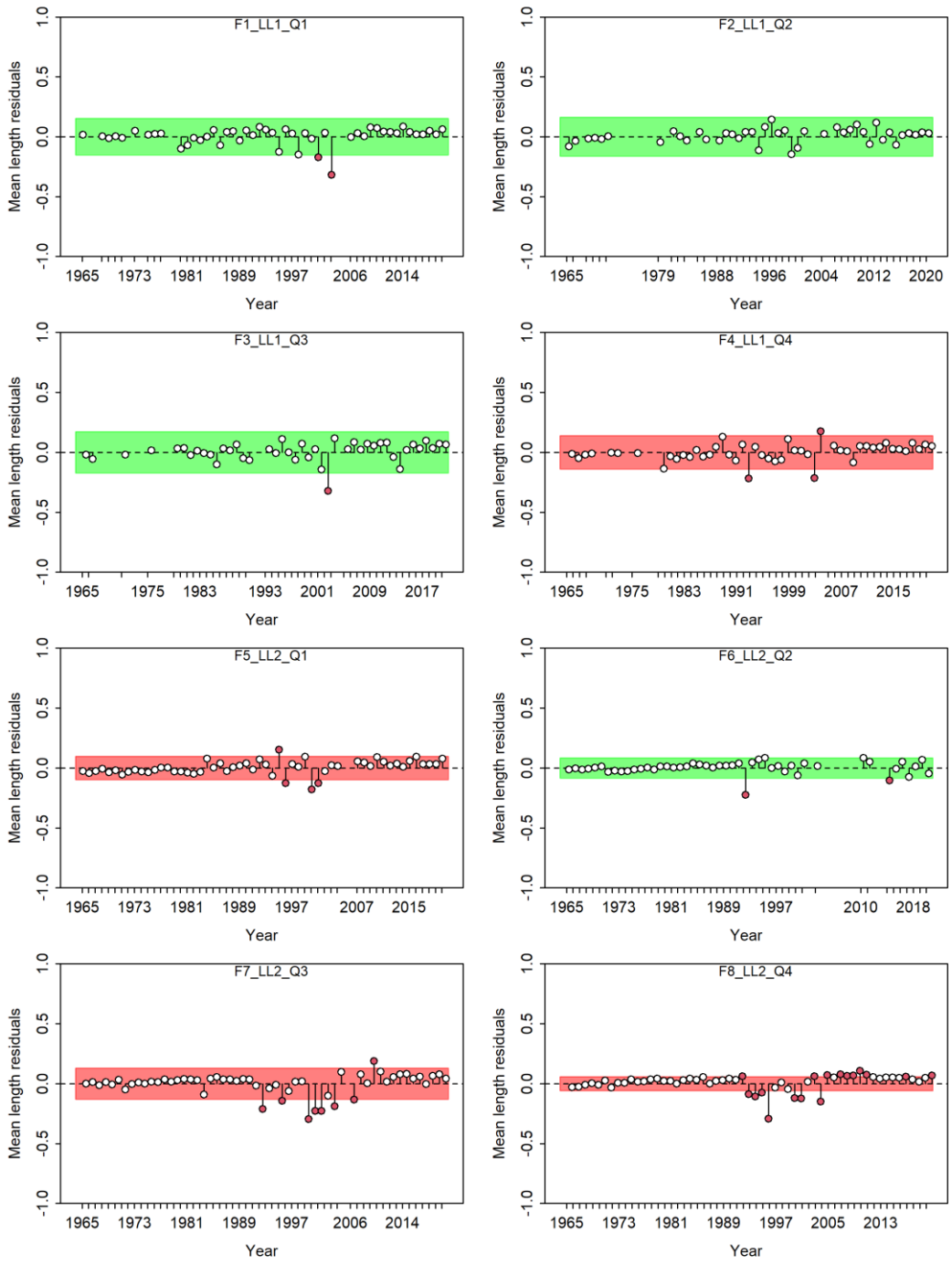

Figure A8. Runs tests results illustrated for the length composition fits (with names and fleet numbers on top) from the Indian Ocean SS3 albacore tuna(ALB). Green shading indicates no evidence ($p \geq 0.05$) and red shading evidence ($p < 0.05$) to reject the hypothesis of a randomly distributed time-series of residuals, respectively. The shaded (green/red) area spans three residual standard deviations to either side from zero, and the red points outside of the shading violate the 'three-sigma limit' for that series.

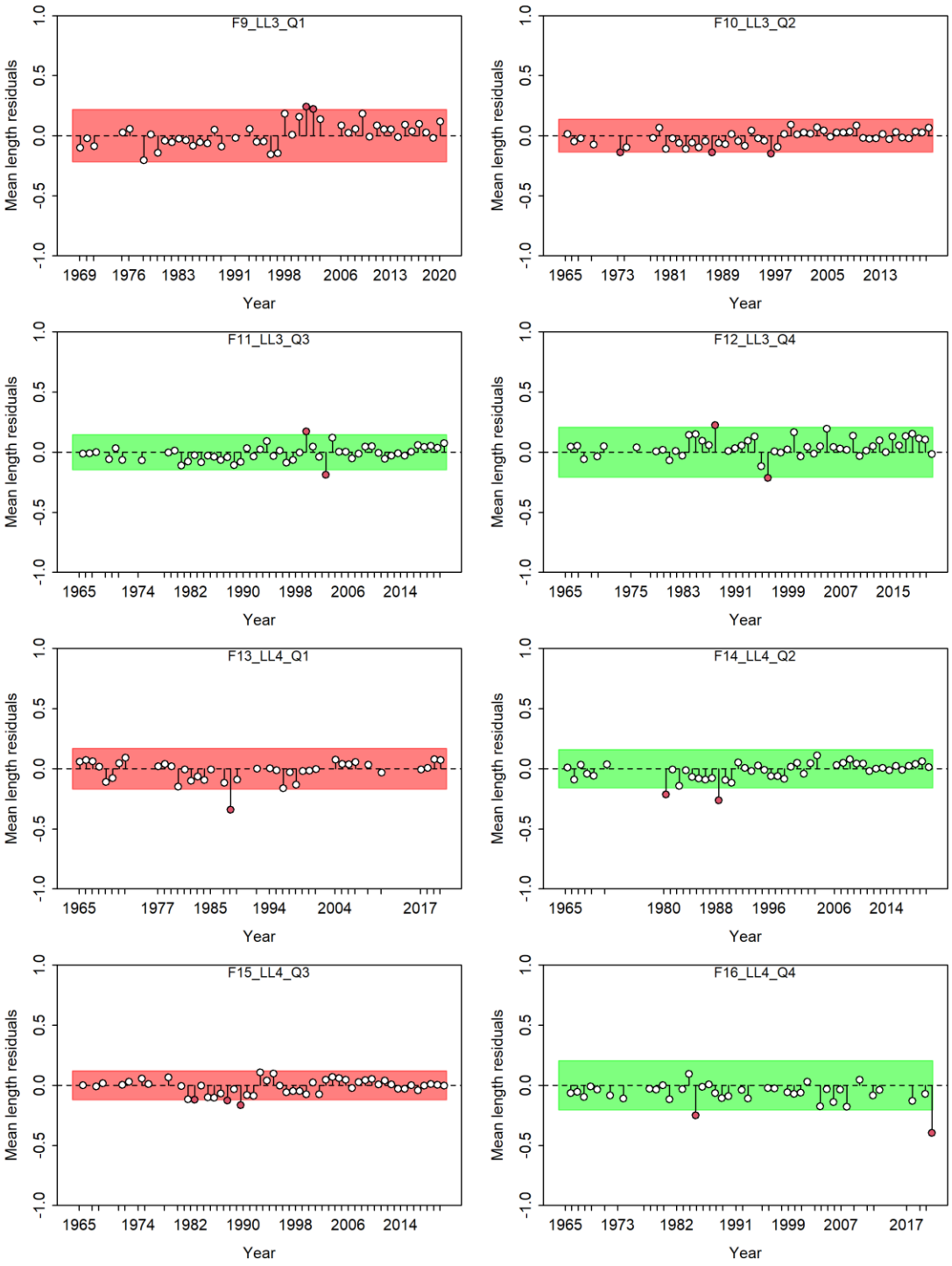

Figure A8. Runs tests results illustrated for the length composition fits (with names and fleet numbers on top) from the Indian Ocean SS3 albacore tuna(ALB). Green shading indicates no evidence ($p \geq 0.05$) and red shading evidence ($p < 0.05$) to reject the hypothesis of a randomly distributed time-series of residuals, respectively. The shaded (green/red) area spans three

residual standard deviations to either side from zero, and the red points outside of the shading violate the 'three-sigma limit' for that series.

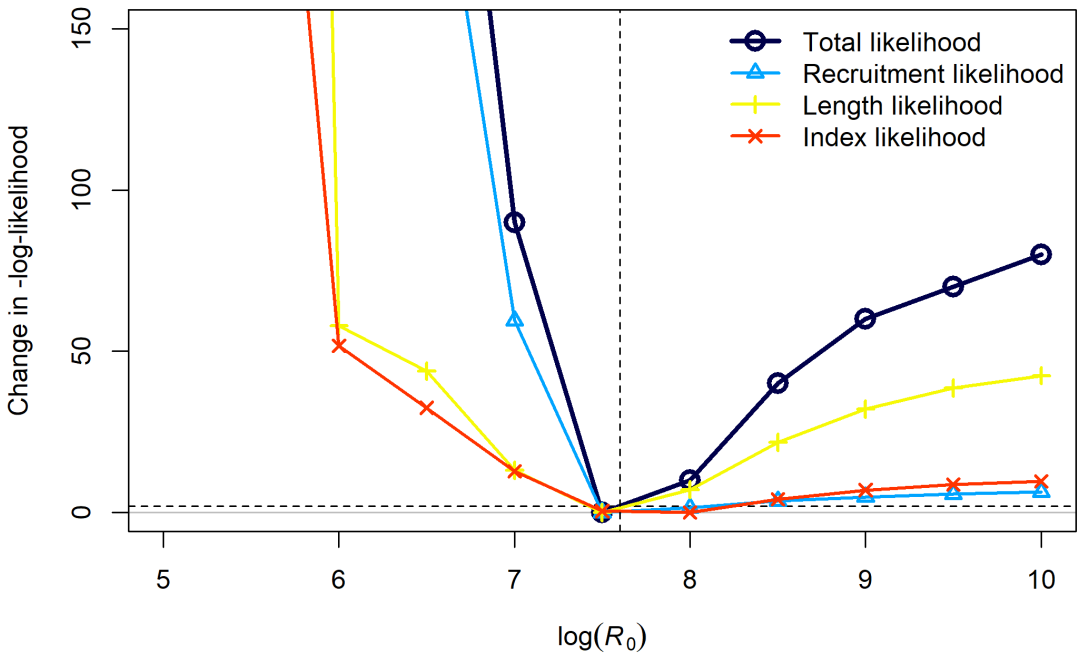

Figure A9. Total likelihood and component profiles (recruitment, length and index (CPUE) components).

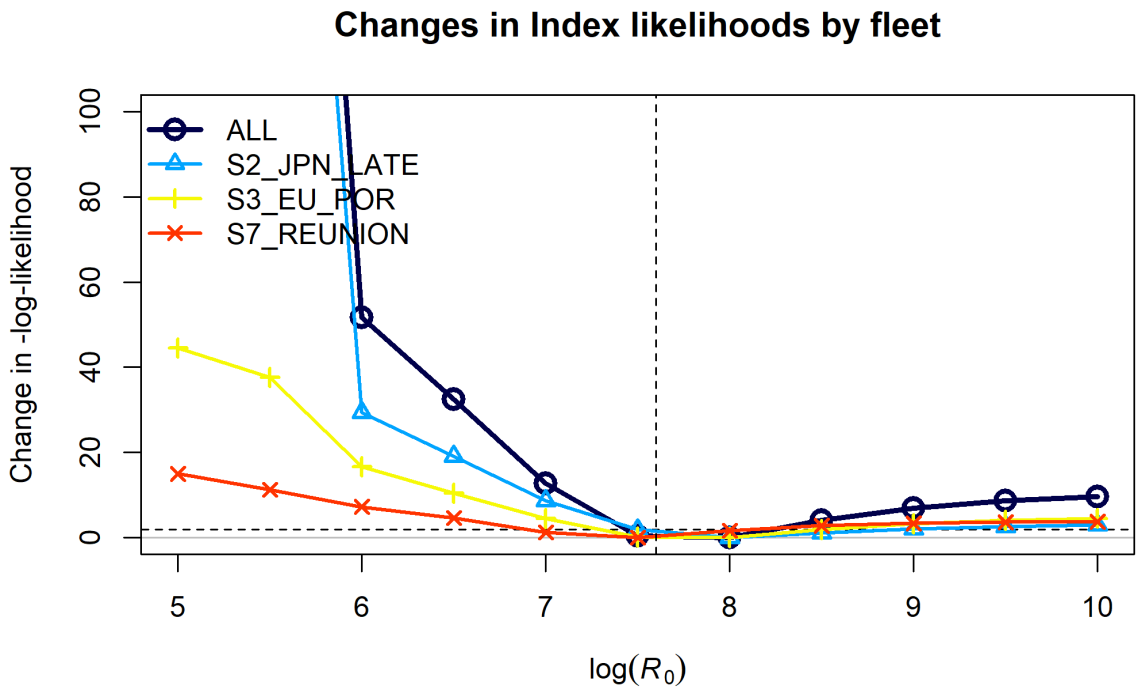

Figure A10. CPUE likelihoods for the diagnostic model.

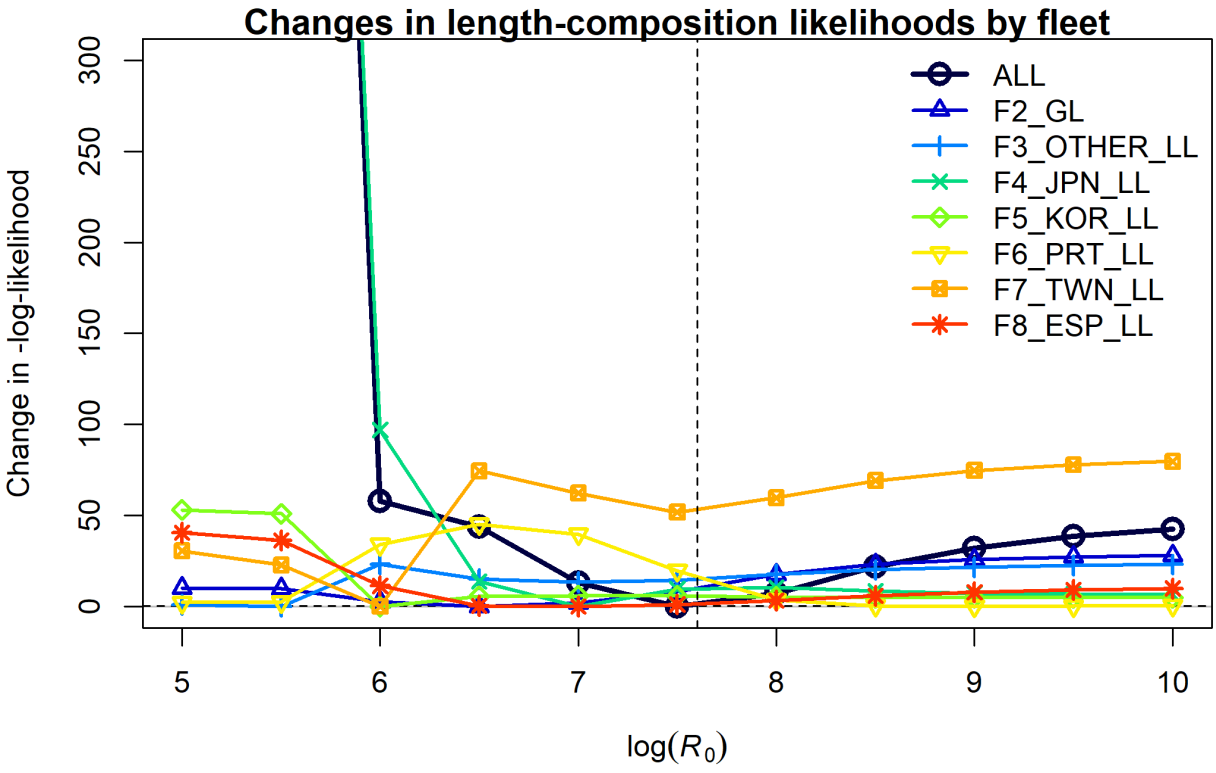

Figure A11.R0 profiles likelihoods for the fit to the length data.

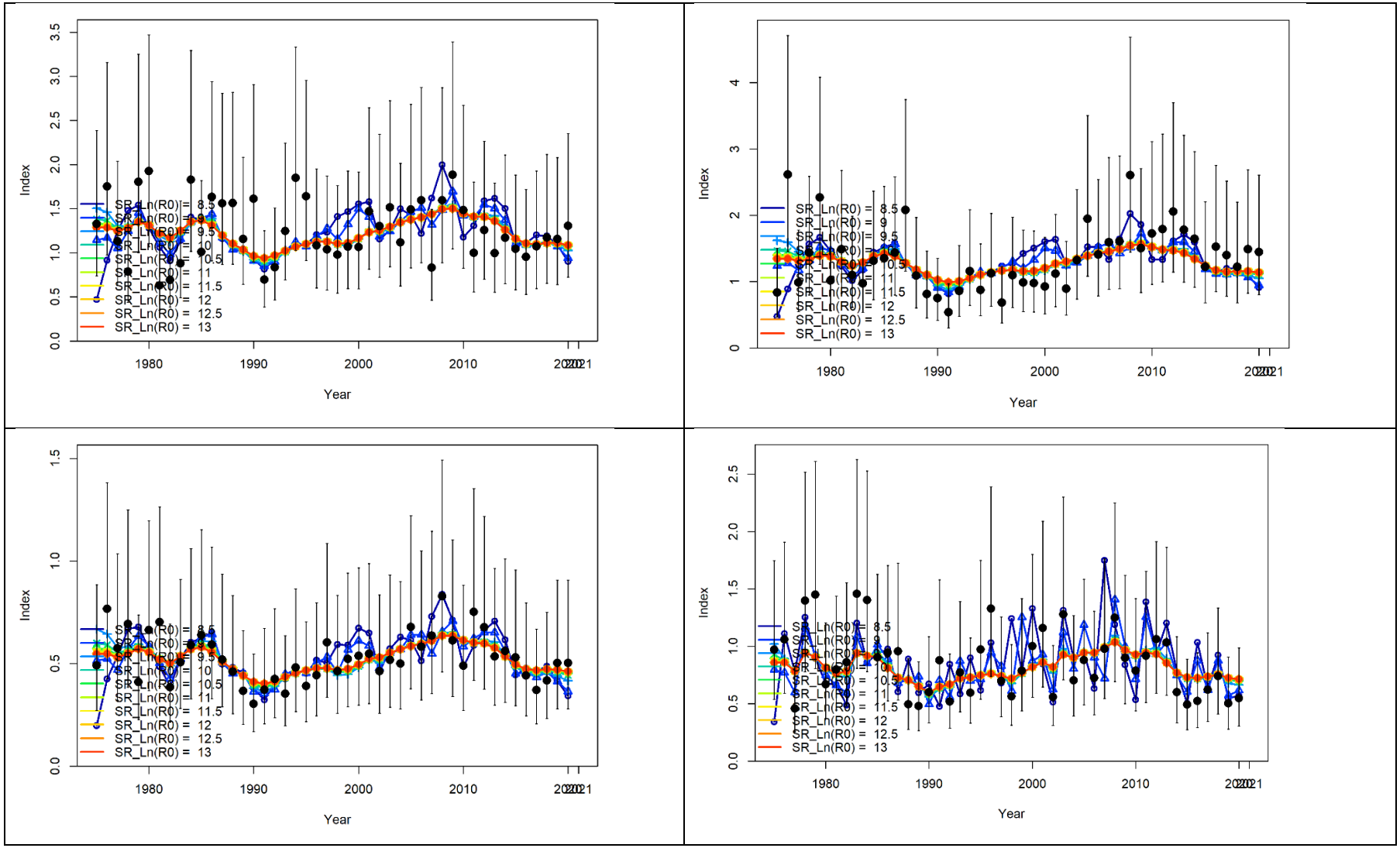

Figure A12. Fits to the component CPUE series (black dots and lines) for various profile likelihood values colored lines. Only the CPUE series that were fit in the diagnostic case are presented. The top panel is the fit to the Japanese CPUE, middle panel is the fit to the Portuguese CPUE and the bottom panel is the fit to the Reunion CPUE series

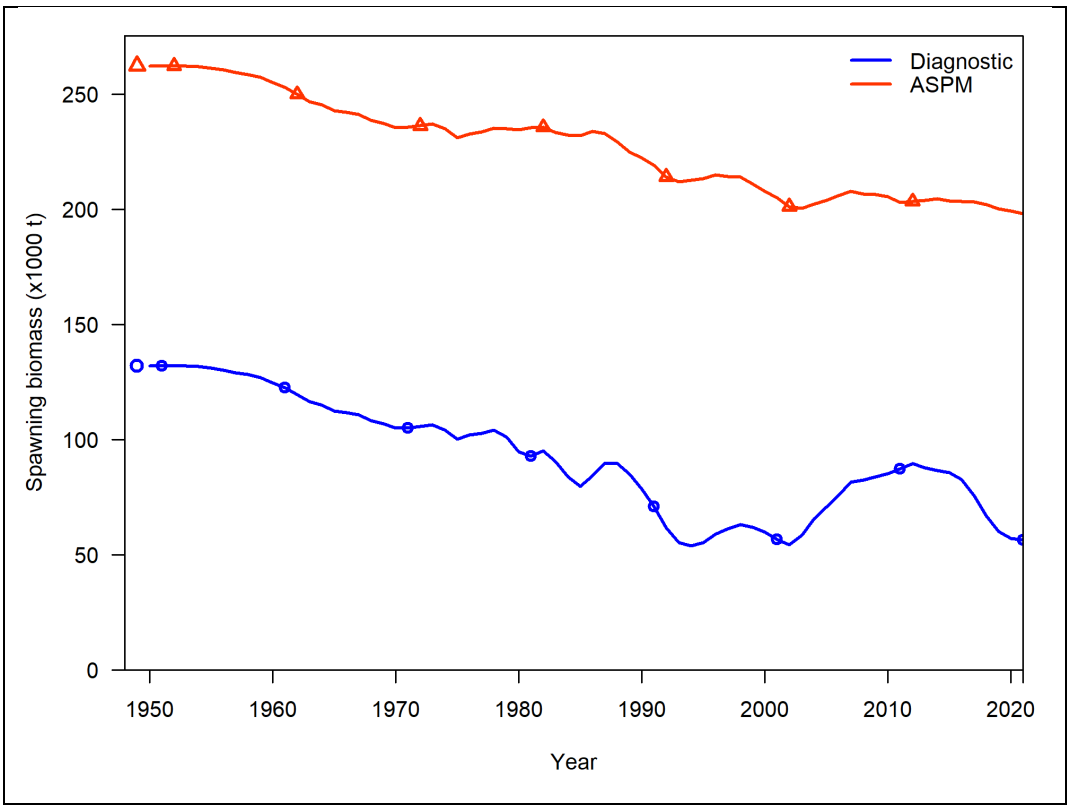

Figure A13. Comparison of spawning biomass trajectories for the ASPM and the diagnostic case of the assessment model carried out in stock synthesis.

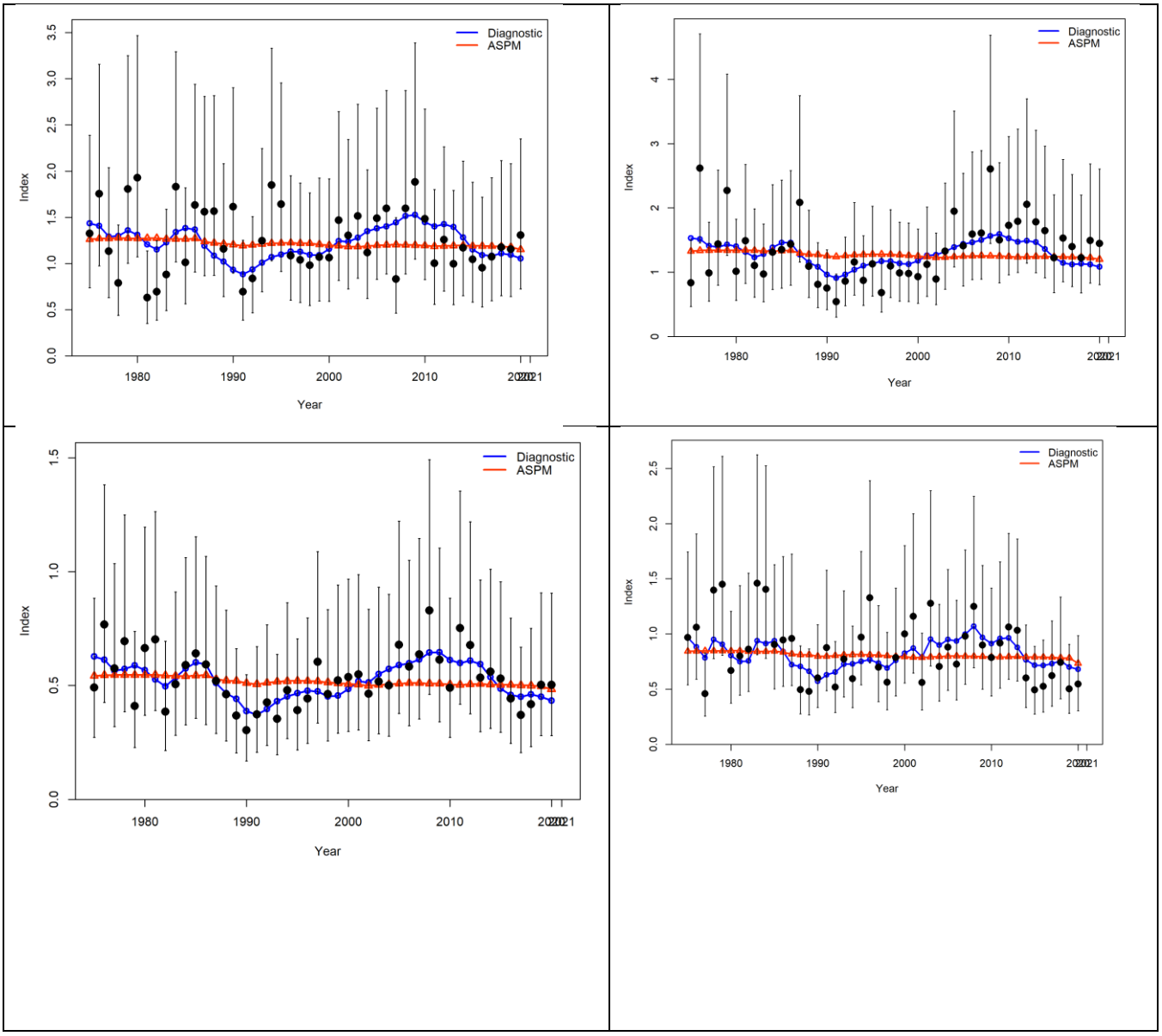

Figure A14. Fits to the indices for the ASPM and Diagnostic case. The panels indicate the fits to the CPUE from region 3 (SW region) by quarter, top left and right are quarters 1 &3, bottom left and right are quarters 3 and 4.

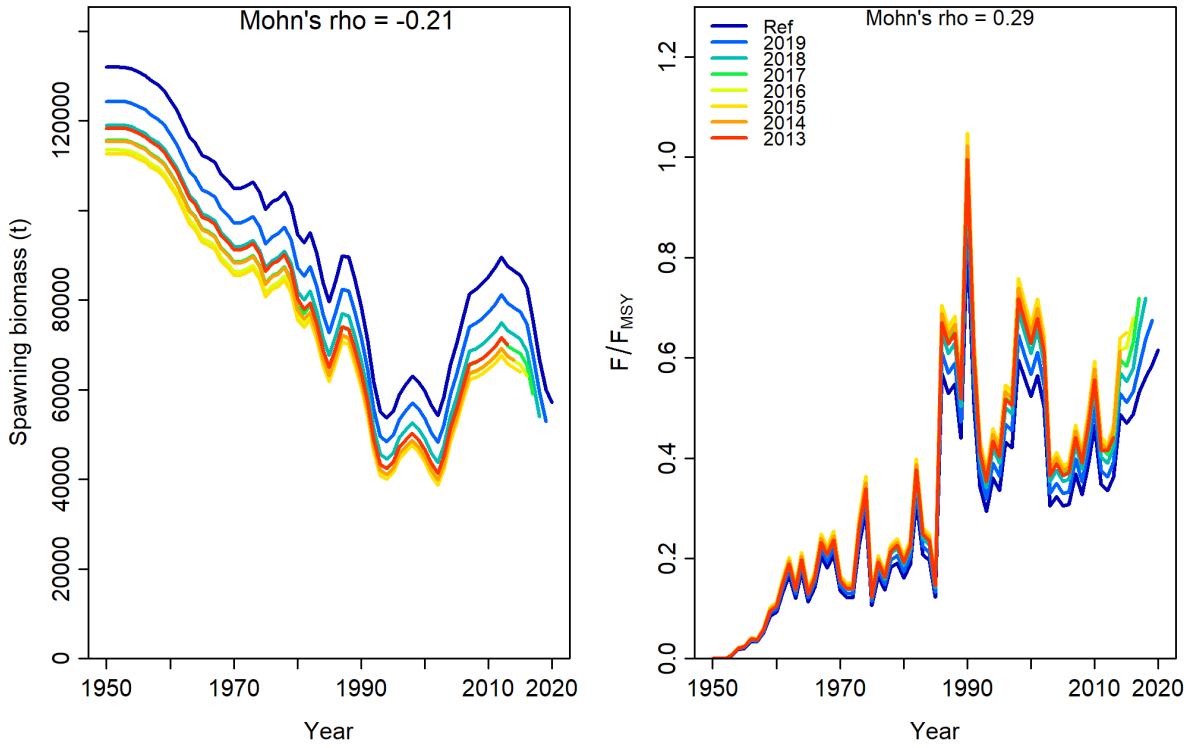

Figure A15. Retrospective analysis of spawning stock biomass (SSB) and fishing mortality estimates for Indian Ocean Albacore tuna conducted by re-fitting the reference model (Ref) after seven years, one year at a time sequentially. Mohn's rho statistic are denoted on top of the panels.

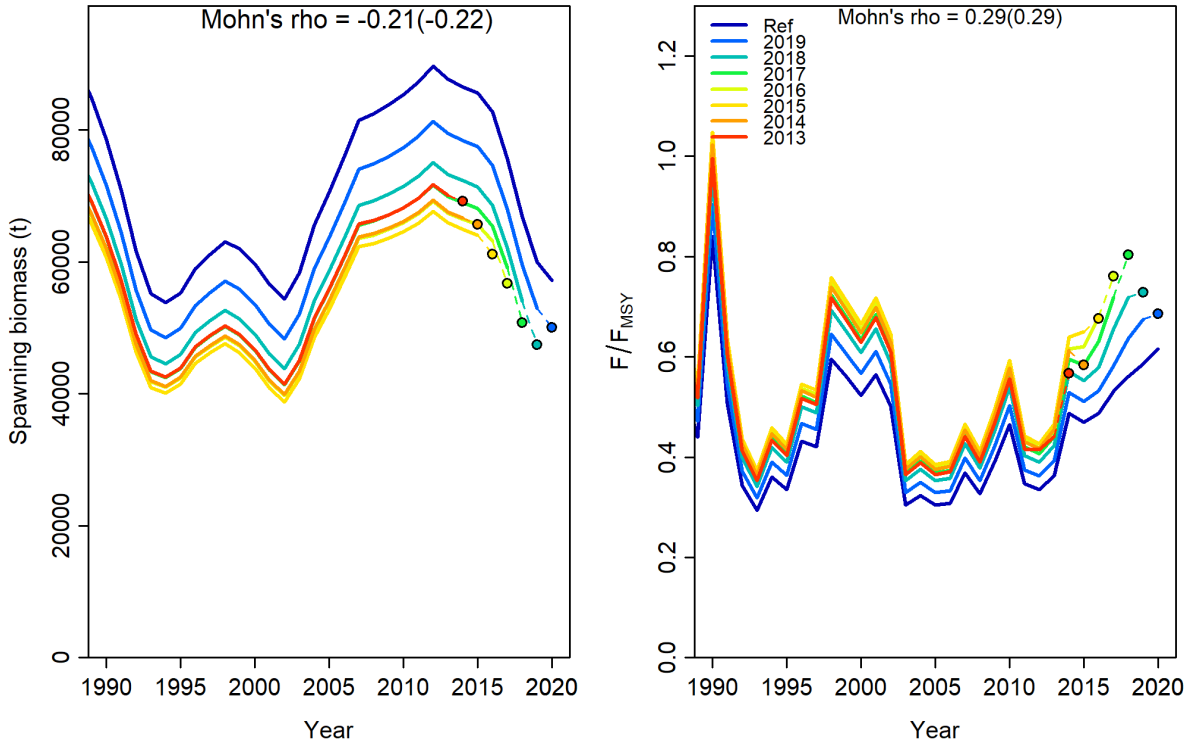

Figure A16. Retrospective results shown for the most recent years only. Mohn's rho statistic and the corresponding 'hindcast rho' values (in brackets) are printed at the top of the panels. One-year-ahead projections denoted by color-coded dashed lines with terminal points are shown for each model.

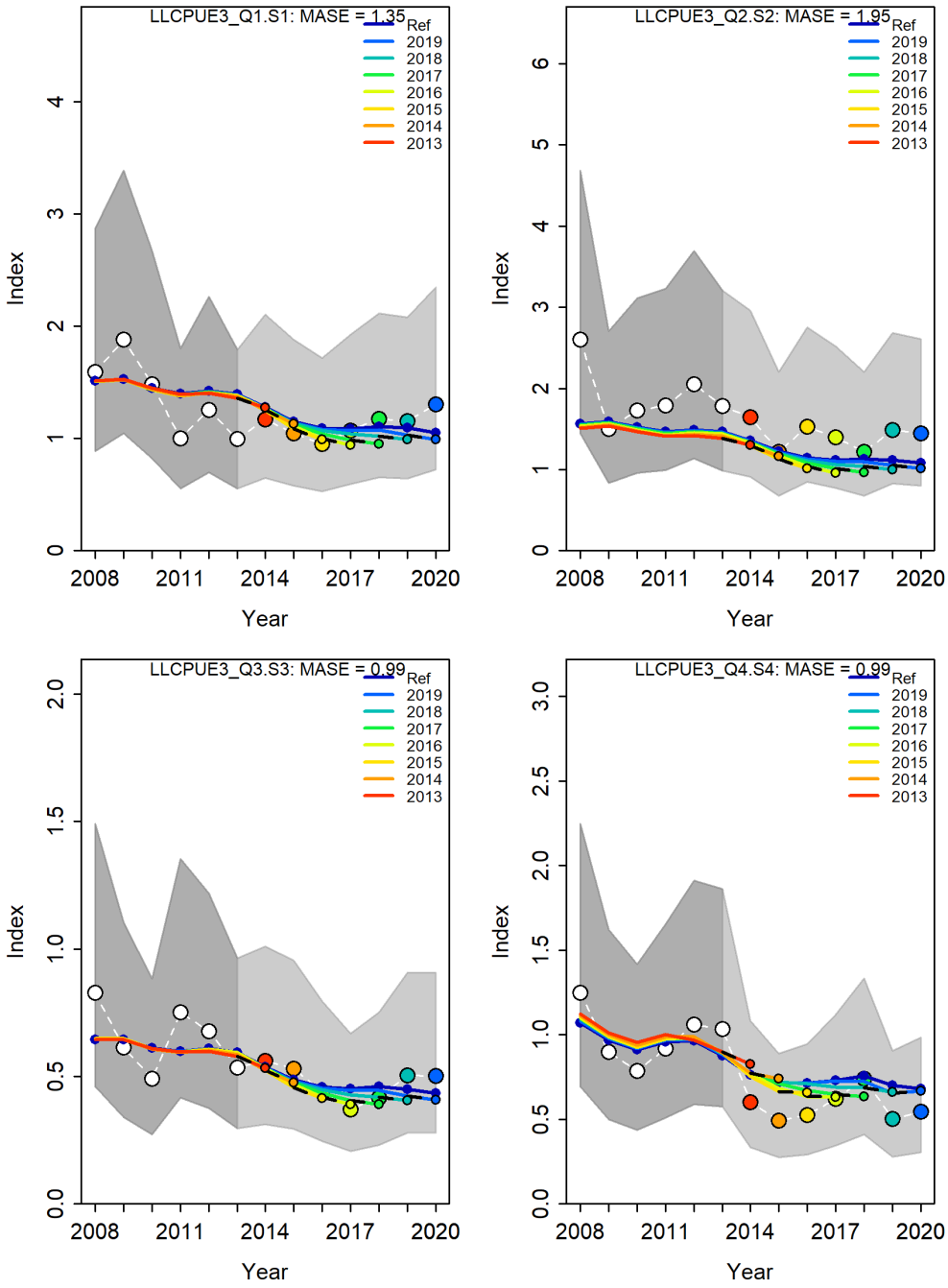

Figure A17. Hindcasting cross-validation (HCxval) results from CPUE fits, showing observed (large points connected with dashed line), fitted (solid lines) and one-year ahead forecast values (small terminal points). HCxval was performed using one reference model (Ref) and seven hindcast model runs (solid lines) relative to the expected CPUE. The observations used for cross-validation are highlighted as color-coded solid circles with associated 95 % confidence intervals. The model reference year refers to the endpoints of each one-year-ahead forecast

and the corresponding observation (i.e., year of retrospective + 1). The mean absolute scaled error (MASE) score associated with each CPUE.

12 References

Brooks, E.N., Legault, C.M., 2016. Retrospective forecasting — evaluating performance of stock projections for New England groundfish stocks. Can. J. Fish. Aquat. Sci. 73, 935–950.

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A., Whitten, A.R., Punt, A.E., 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. Ices J. Mar. Sci. 72, 99–110. https://doi.org/10.1093/icesjms/fsu198.

International Commission for the Conservation of Atlantic Tunas (ICCAT). 2017. Report of the 2017 ICCAT Shortfin Mako Data Preparatory Meeting (Madrid, Spain 28-31 March, 2017).

Kell, L. T., I. Mosqueira, P. Grosjean, J-M. Fromentin, D. Garcia, R. Hillary, E. Jardim, S. Mardle, M. A. Pastoors, J. J. Poos, F. Scott, R. D. Scott. 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES J Mar Sci, 64 (4): 640-646. doi: 10.1093/icesjms/fsm012

Lee, H.-H., Piner, K.R., Methot, R.D., Maunder, M.N., 2014. Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: An example using blue marlin in the Pacific Ocean. Fish. Res. 158, 138–146. https://doi.org/10.1016/j.fishres.2013.12.017.

Maunder, M.N., Piner, K.R., 2015. Contemporary fisheries stock assessment: many issues still remain. ICES J. Mar. Sci. 72, 7–18. https://doi.org/10.1093/icesjms/fsu015.

Michaelsen, J., 1987. Cross-validation in statistical climate forecast models. J. Clim. Appl. Meteorol. 26, 1589–1600. https://doi.org/10.1175/1520-0450(1987)026<1589:CVISCF>2.0.CO;2.

Miller, T.J., Legault, C.M., 2017. Statistical behavior of retrospective patterns and their effects on estimation of stock and harvest status. Fish. Res. 186, 109–120. https://doi.org/10.1016/j.fishres.2016.08.002.

Minte-Vera, C.V., Maunder, M.N., Aires-da-Silva, A.M., Satoh, K., Uosaki, K., 2017. Get the biology right, or use size-composition data at your own risk. Fish. Res. 192, 114–125. https://doi.org/10.1016/j.fishres.2017.01.014.

Mohn, R., 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES J. Mar. Sci. 56, 473–488. https://doi.org/10.1006/jmsc.1999.0481

Wald, A., Wolfowitz, J., 1940. On a test whether two samples are from the same population. Ann. Math. Stat. 11, 147–162. http://www.jstor.org/stable/2235872.

Winker, H., Carvalho, F., Kerwath, S., 2020. Age-structured biomass dynamics of North Atlantic shortfin mako with implications for the interpretation of surplus production models. Col. Vol. Sci. Pap. ICCAT 76, 316–336. https://www.iccat.int/Documents/CVSP/CV076_2019/colvol76.html.