

PRELIMINARY ASSESSMENT OF INDIAN OCEAN YELLOWFIN TUNA 1950-2018 (STOCK SYNTHESIS, V3.30)

Agurtzane Urtizbera¹, Dan Fu², Gorka Merino³, Richard Methot⁴, Massimiliano Cardinale⁵,
Henning Winker⁶, John Walter⁷, Hilario Murua⁸

SUMMARY

This paper presents a preliminary reference model for the assessment of yellowfin tuna (*Thunnus albacares*) using the age and length structured integrated assessment model Stock Synthesis (SS) version 3.30.09. In this document we review the reference model that was used for the 2018 assessment as part of the 2019 workplan for yellowfin. The main features of the new model are a proposal for reducing or removing the influence of tagging data and for a reduced number of areas. The analyses that led to this proposal are explained throughout the document. In brief, the analyses and diagnostics of the model suggest that tagging data and environmental data do not contain enough information to estimate the movement between the 4 areas defined within the model: western-tropical, western-temperate, eastern-tropical and western tropical, and that these data make the model unstable. Therefore, we analyzed and compared three spatial configuration options: two area model defining East and West regions, three area model aggregating regions 3 and 4 of the 2018 model, and a four-area model comparable to the last year reference model but with a different version of (v3.30). The results suggest that the 2-area model is the most stable model and therefore, the 2-area model is proposed as the spatial structure of the reference model. From this, we update the model with the latest data available and analyse a number of sensitivity analyses. The catch and length frequency data were updated until 2018, new estimates of the joint index were introduced and two other additional indices were considered for inclusion as sensitivity runs; an acoustic index derived from echosounder receivers placed on FAD buoys prior to fishing and a purse seine free school index that improves upon the definition of effort in the purse seine fishery. However, the 2018 reference model was proven to be very sensitive to the new length frequency data. Similar conclusions arise from the reference model proposed in this paper. The model is very sensitive to the new length frequency data of longlines and the reference model proposed here uses length frequency data until 2014. This document is a draft and there are still analyses to complete before the 2019 WPTT. This document does not contain results on stock status nor reference points and these will be presented and discussed during the WPTT.

KEYWORDS: yellowfin tuna, stock assessment, stock synthesis

¹ AZTI, aurtizbera@azti.es;

² IOTC, Dan.fu@fao.org;

³ AZTI, gmerino@azti.es;

⁴ NOAA, richard.methot@noaa.gov

⁵ SLU, massimiliano.cardinale@slu.se;

⁶ DAFF, henning.winker@gmail.com

⁷ NOAA, john.f.walter@noaa.gov

⁸ ISSF, hmurua@issf-foundation.org

Contents

| | |
|---|----|
| Introduction | 3 |
| Methods | 4 |
| Stock Synthesis..... | 4 |
| Conversion to SS3.30..... | 4 |
| YFT 2018 model v3.30: Model spatial structure and movement..... | 6 |
| Updated 2-area model: | 12 |
| Data weighting, parameters, and likelihood..... | 29 |
| Model Diagnostics..... | 30 |
| Parameters Estimated..... | 30 |
| Benchmark and fishing mortality calculations | 33 |
| Uncertainty Quantification..... | 33 |
| Results | 36 |
| Building the reference case and update | 36 |

Introduction

This document presents a preliminary Stock Synthesis model for yellowfin tuna in the Indian Ocean including fishery data up to 2018. The assessment model is an age structured population model developed from the previous configuration (Fu et al., 2018). The procedure to develop the new model has been the following: First, the 2018 model has been migrated to the 3.30 version of Stock Synthesis. For that, a series of analyses have been made to verify that the new version produces similar results to the previous model. After that, we focused on elucidating the more appropriate regional structure of the model. Starting from the 4-area configuration of the previous version alternatives of 2 and 3-area configurations have been evaluated. Finally, a series of diagnostics have allowed evaluating the different pieces of information and to propose alternative treatments to catch, cpue, tagging and size frequency data. Finally, the proposed reference model configuration has been updated with the most recent series of data.

Starting from the new proposal for a reference model we also develop a series of sensitivity runs to help the WPTT characterize uncertainty within the fishery in a reference grid. The sensitivities include alternative biological parameters (steepness, mortality, growth) weighting and use of data sources (tagging, alternative indices of abundance) and other options. It is foreseen that statistical uncertainty will also be explored during the WPTT (REF a paper de Henning, IOTC–2019–WPTT21–51).

As said, the objective of this document is to propose a reference case for the Indian Ocean yellowfin and help the WPTT deciding a grid of models to provide scientific advice on the status of this stock and to recommend catch limits in a probabilistic manner.

At the moment of the submission of this document there are still analyses to make prior to the WPTT. The results of the model developed here will be presented to the WPTT and have not been included in this document.

Methods

Stock Synthesis

Stock Synthesis (SS3) is an integrated statistical catch-at-age model that is widely used for many stock assessments across tuna RFMOs (Methot and Wetzel 2013). SS3 takes relatively unprocessed input data and incorporates many of the important processes (mortality, selectivity, growth, etc.) that operate in conjunction to produce fits to observed catch, size and age composition and CPUE indices. Because many of these inputs are correlated, the concept behind SS3 is that they should be modeled together, which helps to ensure that uncertainties in the input data are adequately represented in the assessment. SS3 is comprised of three subcomponents: 1) a population subcomponent that recreates the numbers/biomass at age using estimates of natural mortality, growth, fecundity, etc; 2) an observational sub-component that consists of observed (measured) quantities such as CPUE or proportion at length/age; and 3) a statistical sub-component that uses likelihoods to quantify the fit of the observations to the recreated population. Basic equations and technical specifications underlying Stock Synthesis can be found in Methot (2000). In these models, we use SS version 3.30.09. SS Version 3.30 has many updated features from previous versions, notably it allows for greater precision in modeling temporal dynamics, in specifying future recruitment and more streamlined modeling of time-varying processes.

Conversion to SS3.30

As a first step towards building a reference model for the 2019 stock assessment, the 2018 model was converted from SS 3.24 to SS 3.30. The comparison of both versions is made with the reference case (*io_h80_q1_tm30_dw1*) defined in the 2018 assessment with the spatial structure and movement defined in Figure 1 (steepness=0.8, single catchability for longline, tag-release mortality of 28,5% and tag data not downweighed). During the conversion the movement parameters were hitting bounds so the boundaries were increased from -12,12 and to -15,15 and catchability (q) was changed to a parameter, so the 4 seasons contribute to the scaling of q . There are some differences in the likelihood due to these changes (Table 1) and also in the spawning biomass and R_0 (Figure 2). But both models estimate different movement rates (Figure 3), so in order to make them comparable, we analyzed the v3.30 model but with fixed movement rates; the same movement rates as the estimated by v3.24 model (Figure 2, Table 1). The estimates of models with different SS version but the same movement rates are very similar, so most of the difference between the models with different versions is a consequence of local minimum estimates of movement rates. Then the differences between both models were not due to the conversion and thus, we followed the analysis with the v3.30 version.

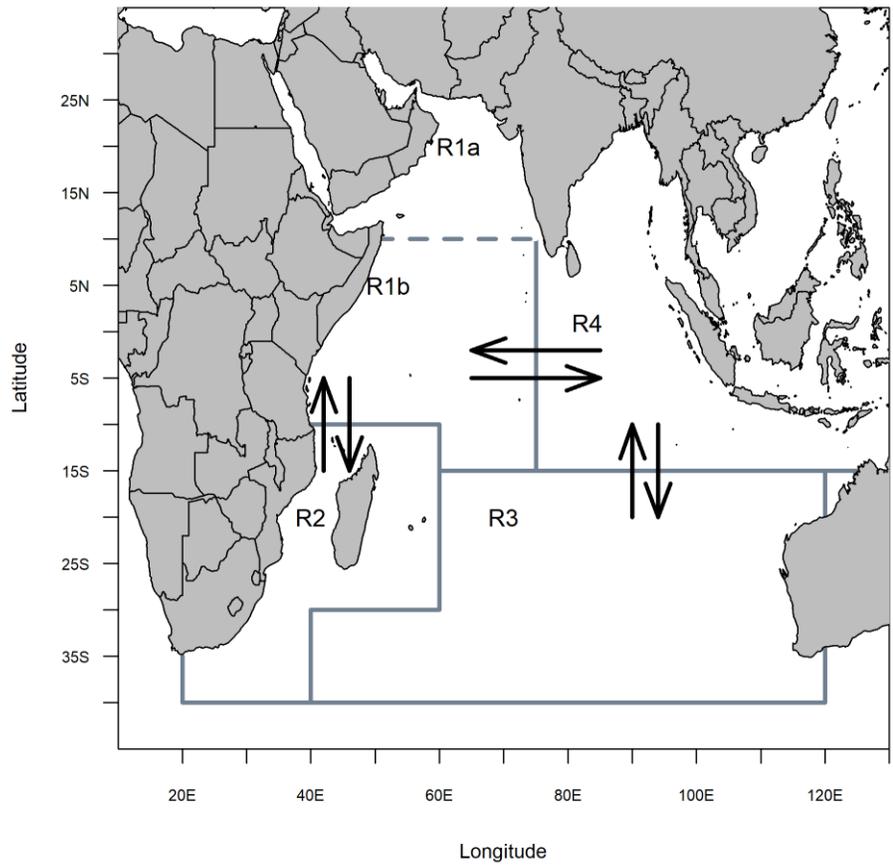


Figure 1. Spatial stratification of the Indian Ocean for the 4-area assessment model. The black arrows represent the configuration of the movement parameterization of the base assessment model.

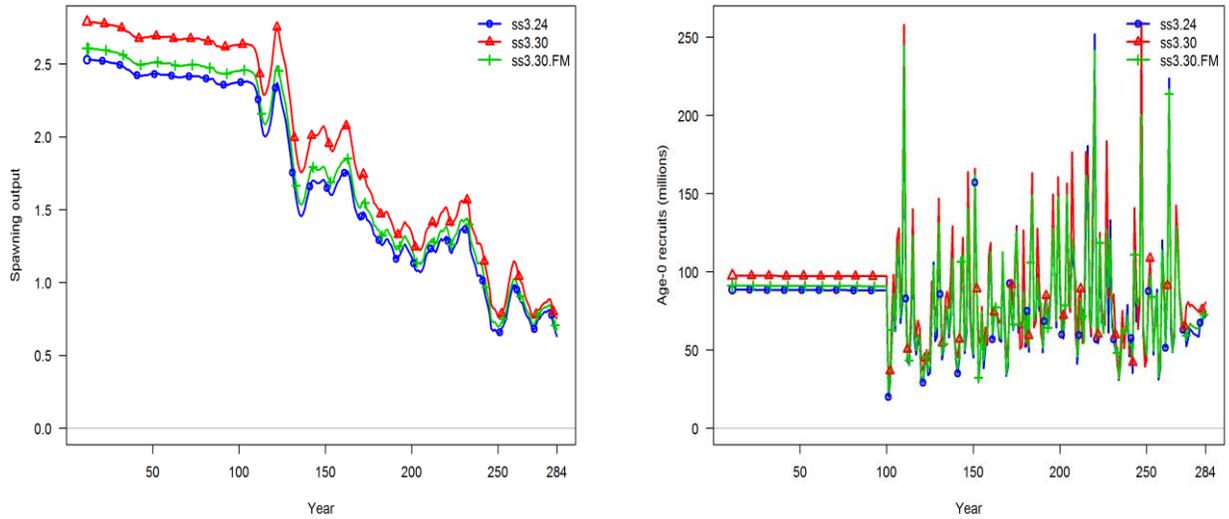


Figure 2. Comparison of SSB for run 5 model in 2016 in SS 3.24, similar model but with v3.30 and the v3.30 model but with fixed movement rates (the same as in v3.24).

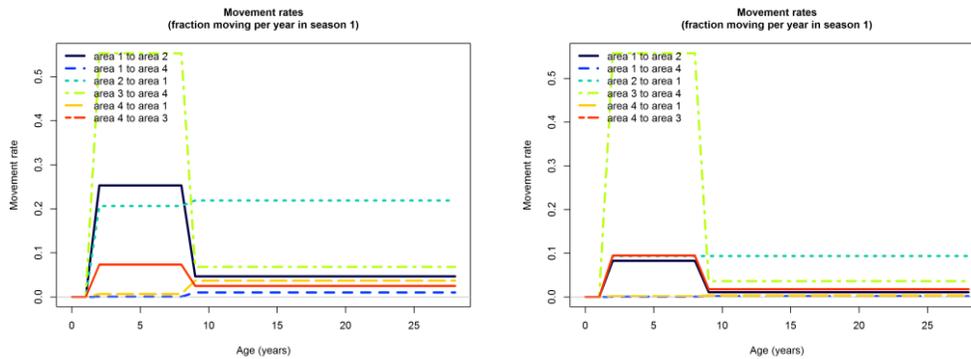


Figure 3. Estimated movement rates with v3.24 and v3.30 model.

Table 1. Likelihoods of the YFT 2018 reference model v3.24, similar model with version v3.30, and the v3.30 but fixing the movement rate with the estimated movement rate with v3.24.

| | v3.24 | v3.30 | v3.30_FM |
|----------------------|----------|-----------|----------|
| TOTAL | 9360.71 | 9216.73 | 9310.93 |
| Catch | 7.61E-05 | 0.0365694 | 5.71E-05 |
| Equil_catch | 0 | 0 | 0 |
| Survey | -303.702 | -337.498 | -311.751 |
| Length_comp | 3869.15 | 3833.16 | 3858.32 |
| Tag_comp | 4053.92 | 4021.68 | 4047.85 |
| Tag_negbin | 1700.31 | 1679.47 | 1688.92 |
| Recruitment | -45.6385 | -50.8727 | -50.2307 |
| Forecast_Recruitment | 1.93E-05 | 1.93E-05 | 7.42E-06 |
| Parm_priors | 59.6685 | 56.6686 | 58.5807 |

| | | | |
|-----------------|------------|------------|------------|
| Parm_softbounds | 0.00585596 | 0.00584073 | 0.00617977 |
| Parm_devs | 26.1086 | 12.9958 | 18.2173 |
| F_Ballpark | 0.893315 | 1.08146 | 1.0198 |
| Crash_Pen | 0 | 0 | 0 |

YFT 2018 model v3.30: Model spatial structure and movement

The YFT 2018 model is constructed as a seasonal model with 4 seasons and a timeframe from 1950 – 2017. The model was very sensitive to the new length composition data 2015-2017, therefore, only data until 2014 was used during the assessment. The initial model estimates the rates of movement between regions from the available tagging data from 2005 to 2007 (Figure 1), assuming 3 quarters as the mixing latency period, so only 38% of tag-recapture data is used (n=3480.1). All recoveries were in areas 1 and 2 and very small number in area 4 (Table 2). The movement rates in the model are bi-directional between areas 1 and 2, 1 and 4, and 3 and 4 (Figure 1). The model does not estimate almost any movement rates between area 1 and 4, but it does between area 1 and 2, and quite high movement rates between area 3 and 4 (Table 3). However, in addition of the tagging data, environmental effects are also included in the estimates of the movement mainly to create a seasonal pattern in the movement (Table 3). We performed a jitter analysis in order to analyze the sensitivity of the model to initial values and the results show that the model gives quite different results in terms of likelihoods, spawning biomass, fishery mortality or recruitment (Figure 4) depending on the initial values, and therefore it does not find stable results. We already observed that the results of the model are very sensitive to the movement rates (Figure 2) so the sensitivity of the model to the initial values it could be due to the lack of information to estimate movement rates between the areas. Therefore, we analyzed the possibility of different spatial structures in the model in order to improve the convergency and stability of the model.

Table 2. Percentage and numbers of tagging release and recover assuming 3 quarters of mixing latency period.

| | areaRec_1 | areaRec_2 | areaRec_4 |
|------------|------------------|--------------------|----------------|
| areaReal_1 | 0.93 (n=3709) | 0.069 (n=269.2) | 0.001 (n=7) |
| areaReal_2 | 0.89 (n=46.6) | 0.11 (n=6) | 0 |

Table 3. Movement probabilities estimated with v3.30 model.

| Source_area | Dest_area | Ages 2-8 | 9-28 | ENV link |
|-------------|-----------|----------|------|----------|
| | 1 | 0.92 | 0.99 | |

| | | | | |
|---|---|------|------|-----|
| 1 | 2 | 0.08 | 0.01 | 7,1 |
| 1 | 4 | 0.00 | 0.00 | 5,4 |
| 2 | 1 | 0.10 | 0.10 | 7,1 |
| 2 | 2 | 0.90 | 0.90 | |
| 3 | 3 | 0.38 | 0.96 | |
| 3 | 4 | 0.62 | 0.04 | 3,3 |
| 4 | 1 | 0.00 | 0.00 | 5,4 |
| 4 | 3 | 0.10 | 0.02 | 3,3 |
| 4 | 4 | 0.90 | 0.98 | |

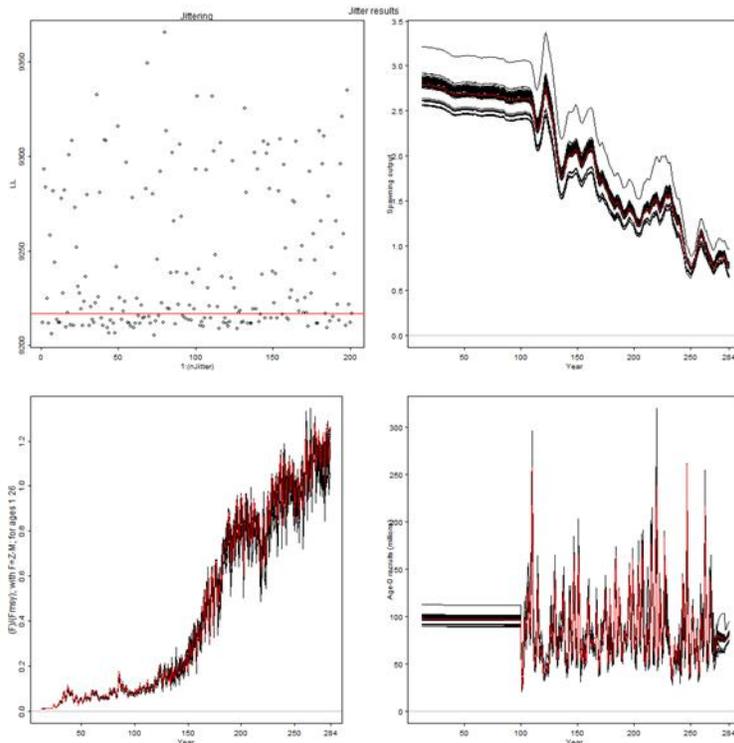


Figure 4. Jitter analysis with 200 simulations with the v3.30 model.

We analyzed and compared a 2-area model and 3-area model with the 2018 reference case.

In the spatial structure of the 2-area model, East and West are considered as two separated regions, without movement rates between them and with recruitment in both regions. So in this case, the area 1 considers the regions R1a, R1b and R2 defined in Figure 1 while the area 2 considers R3 and R4 as one region (Figure 1). In this analysis only the joint index for tropical waters of each area were

considered, because those are the regions with higher biomass within the areas defined in the model. The definition of the areas is different to the 2018 reference model, and this affect the areas where tagging occurs. Therefore, we thought that better to not include tagging data at the beginning of the process but to analyze it later as sensitivity- analysis.

In the 3-area model, the spatial structure in the West region is the same as in the 2018 model, so the temperate and tropical waters are separated into 2 areas, while in the East both regions are aggregated and considered as one area. Therefore, in this case the area 1 aggregates the regions R1a and R1b defined in Figure 1, (the same definition of area 1 in the 2018 reference model), the area 2 considers only the region R2 (the same definition as the region 2 in the 2018 reference model) and the model-area 3 aggregates the region R3 and R4. In this case fish movement rates are estimated between area 1 and 2, and between area 1 and 3. In the East following the same assumptions as in the 2-area model, only the joint index of tropical waters was considered.

When the results of the 3 models are compared (Figure 5), the 3-area and 4-area model yield very similar results, while the virgin biomass of the 2-area model (without tagging data) is higher but with similar trend. However, the jitter analysis suggests that the 2-area model is less sensitive to the initial values than the 3-area and 4-area models (Figures 6, 7 and 8). Based on this, we propose a regional structure of 2 areas for the reference case of the 2019 stock assessment.

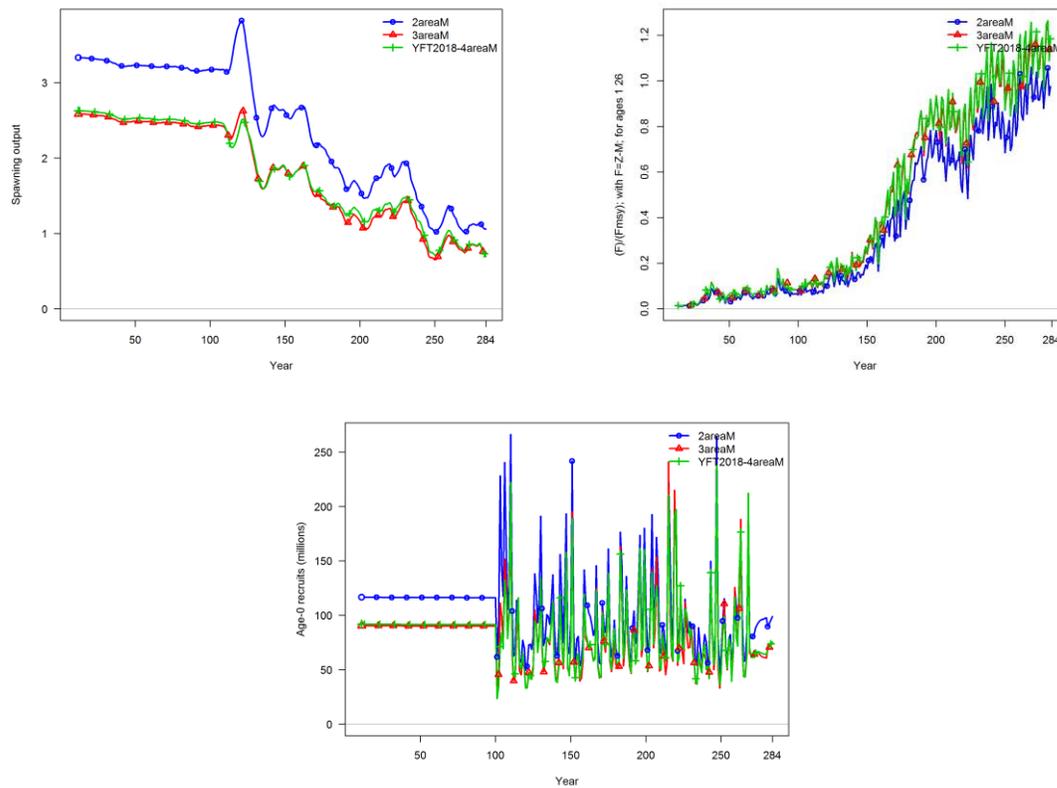


Figure 5. Comparison of the YFT 2018 model with v3.30 (4-area model), 2 area model and 3 area model.

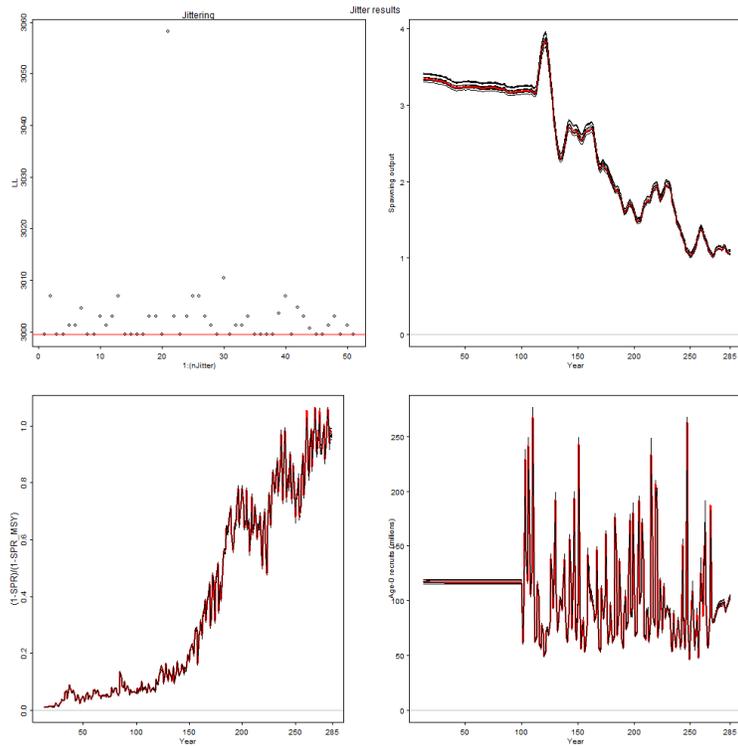


Figure 6. Jitter analysis with 50 simulations of the 2-area model.

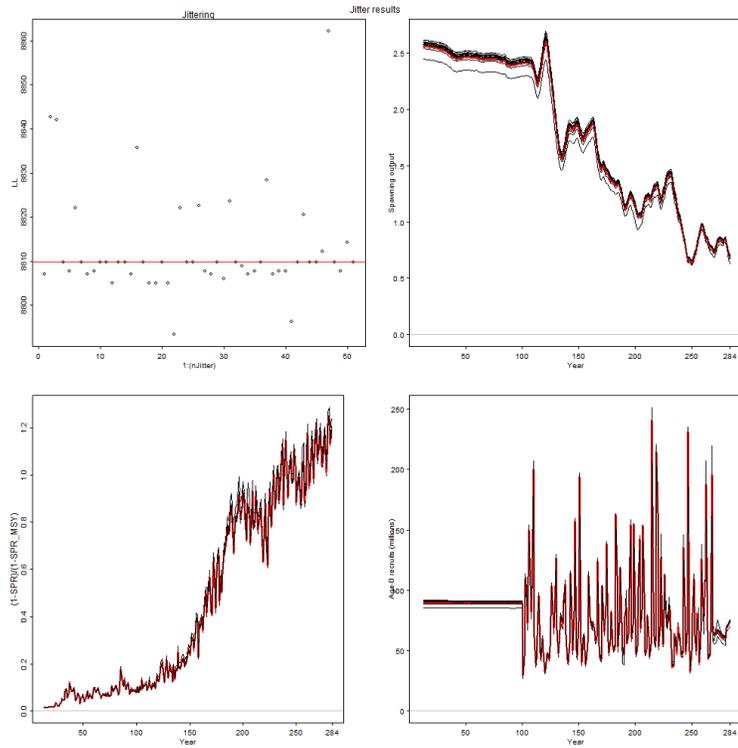


Figure 7. Jitter analysis with 50 simulations of the 3-area model.

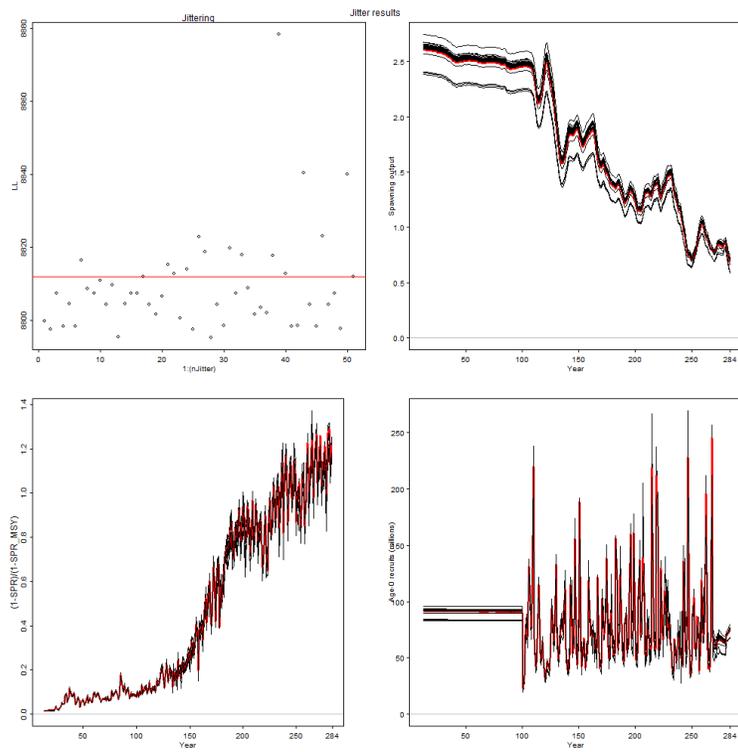


Figure 8. Jitter analysis with 50 simulations of the 4-area model.

Updated 2-area model:

Model spatial structure and movement

The 2-area model is constructed as a seasonal model with 4 seasons and with a timeframe from 1950 – 2018. The spatial structure of the model considers the temperate and tropical waters as 2 separated regions, without any movement between them, similar to the 2018 reference model estimates.

Population dynamics

The new model partitions the population into 2 regions. The population in each region is comprised of 27-quarterly age-classes with both sexes combined. The first age-class has a mean fork length of 22 cm and it is assumed to be approximately three months of age based on ageing studies of yellowfin tuna (Fonteneau, 2008). The last age-class comprises a “plus group” in which mortality and other characteristics are assumed to be constant. Insufficient sex-specific data are available to configure a two-sex population model.

The model commences in 1950 at the start of the available catch history. The initial population age structure in each region was assumed to be in an unexploited, equilibrium state.

Recruitment

Recruitment occurs in each quarterly time step of the model. Recruitment was derived from a Beverton-Holt stock recruitment relationship (SRR) and variation in recruitment was estimated as deviates from the SRR. Recruitment deviates were estimated for 1972 to 2017 (184 deviates), representing the period for which longline CPUE indices are available. Recruitment deviates were assumed to have a standard deviation (σ_R) of 0.6. For 1950-1971, recruitment was derived directly from the SRR. The base model assumed a level of steepness (h) of 0.8 for the SRR, an intermediate value within the plausible range of steepness values generally adopted in the tuna assessments by other tuna RFMOs (0.7, 0.8 and 0.9) (Harley 2011).

Recruitment was assumed to occur in both regions. This assumption was based on the temperature preference for the spawning of yellowfin tuna and a minimum temperature for larval survival of about 24°C (Suzuki 1993).

The overall proportion of the quarterly recruitment allocated to each region was estimated. The base model estimated 73% and 26% of the recruitment occurred in the respective regions. But the parameterization of recruitment is modified in comparison to the 2018 model; only the recruitment distribution parameter for one area is estimated for 1977 to 2017 and the remainder of the recruits goes to the other area (164 parameters less are estimated), and in order to get similar deviates, the deviations were increased to 1.5.

Growth and Maturation

In the reference model proposed here, the growth parameters are the same as in the reference case in 2018, fixed at values that replicated the growth curve derived by Fonteneau (2008) (Figure 9). The last year, different growth estimates were analyzed (Figure 9), however, as sensitivity analysis only the growth rate estimated by Dortel (2015) was used here (Figure 9). Dortel et al. (2015) estimated growth integrating otolith readings from mark-recapture data and mode progressions from purse seine

length frequency data. These estimates were comparable to the values currently incorporated in the assessment model. However, the estimate of the asymptotic length (L_{inf}) was higher (Figure 9).

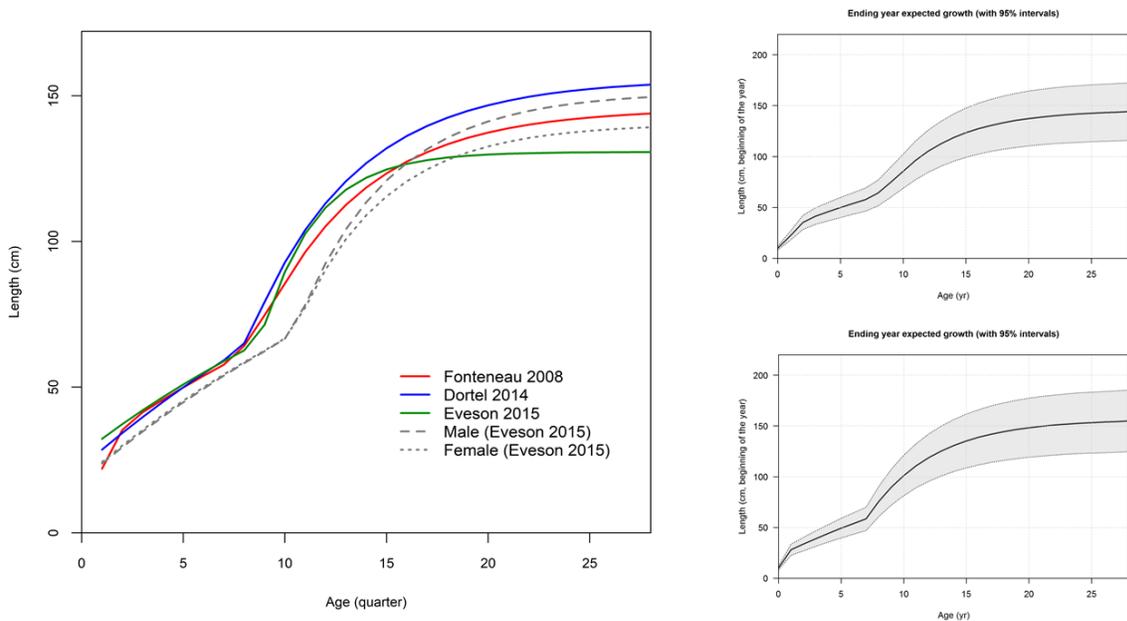


Figure 9. The growth analysed during 2018 assessment model, although in this study only the growth curve estimated by Fonteneau (2008) and by Dortel (2015) is analysed. Topright figure is the growth on the confidence interval assumed by the model with the growth estimated by Fonteneau (2008) and at the bottom assuming the growth estimated by Dortel (2015).

Length based maturity ogives for Indian Ocean yellowfin are available from Zudaire et al (2013). The paper presents two alternative maturity ogives based on either the cortical alveolar or vitellogenic stages of ovarian development. The length-based ogives were converted to age-based ogives assuming an equilibrium population age-length structure. However, in this study as well as in the 2018 assessment only the ogive based on cortical alveolar stage development was included, because the results with both methodologies were very similar. Thus, the age-based ogive was provided to the reference case as proportions of mature at age, with the onset of maturity at about age 5 quarters (about 75 cm) and full maturity at about 12 quarters (Figure 10).

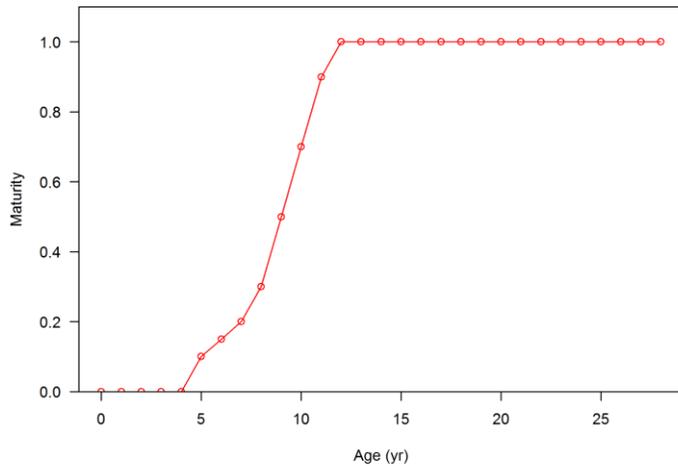


Figure 10. The age-based maturity OGIVEs for Indian Ocean yellowfin tuna (derived from Zudaire et al 2013).

Natural Mortality

Natural mortality is parameterized as in the reference case of 2018. Natural mortality is variable with age and with the relative trend in age-specific natural mortality based on the values applied in the Pacific Ocean (western and central; eastern) yellowfin tuna stock assessment.

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

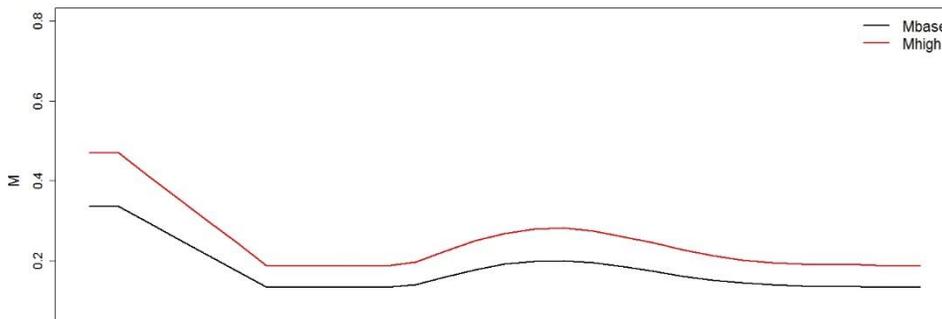


Figure 11. The age-specific natural mortality schedule assumed for the reference case model (Base) and a 40% higher age-specific M as sensitivity analysis (see text for details).

The resulting age-specific natural mortality has been used as the base level of natural mortality for the stock assessment since 2015, while a higher level of natural mortality is included in a model sensitivity (Figure 11). For the current assessment, the base level M is adopted for the reference model and a 40% higher natural mortality is included in a model sensitivity (M_{high}). However, in addition we also add another sensitivity test, where we let the model to estimate M as a constant parameter with age. And we also analyze the estimation of the parameter including tagging data (with λ 0.1, downweighted).

Fleet structure

The assessment adopted the equivalent fisheries definitions used in the 2018 stock assessment. These “fisheries” represent relatively homogeneous fishing units, with similar selectivity and catchability characteristics that do not vary greatly over time. Twenty-five fisheries were defined based on the 4-area defined in the 2018 reference model, time period, fishing gear, purse seine set type, and type of vessel in the case of longline fleet (Table 4).

Table 4. Definition of fisheries for the reference case assessment model for yellowfin tuna.

| Fishery | Nationality | Gear | Region |
|-----------------------|--------------------|---------------------------|---------------|
| 1. GI 1a | All | Gillnet | 1a |
| 2. HD 1a | All | Handline | 1a |
| 3. LL 1a | All | Longline | 1a |
| 4. OT 1a | All | Other | 1a |
| 5. BB 1b | All | Baitboat | 1b |
| 6. PS FS 1b 2003-06 | All | Purse seine, school sets | 1b |
| 7. LL 1b | All | Longline | 1b |
| 8. PS LS 1b 2003-06 | All | Purse seine, log/FAD sets | 1b |
| 9. TR 1b | All | Troll | 1b |
| 10. LL 2 | All | Longline | 2 |
| 11. LL 3 | All | Longline | 3 |
| 12. GI 4 | All | Gillnet | 4 |
| 13. LL 4 | All | Longline (distant water) | 4 |
| 14. OT 4 | All | Other | 4 |
| 15. TR 4 | All | Troll | 4 |
| 16. PS FS 2 | All | Purse seine, school sets | 2 |
| 17. PS LS 2 | All | Purse seine, log/FAD sets | 2 |
| 18. TR 2 | All | Troll | 2 |
| 19. PS FS 4 | All | Purse seine, school sets | 4 |
| 20. PS LS 4 | All | Purse seine, log/FAD sets | 4 |
| 21. PS FS 1b pre 2003 | All | Purse seine, school sets | 1b |
| 22. PS LS 1b pre 2003 | All | Purse seine, log/FAD sets | 1b |

| | | | |
|------------------------|-----|---------------------------|----|
| 23. PS FS 1b post 2006 | All | Purse seine, school sets | 1b |
| 24. PS LS 1b post 2006 | All | Purse seine, log/FAD sets | 1b |
| 25. LF 4 | All | Longline (fresh tuna) | 4 |

The longline fishery was partitioned into two main components:

Freezing longline fisheries, or all those using drifting longlines for which one or more of the following three conditions apply: (i) the vessel hull is made up of steel; (ii) vessel length overall of 30 m or greater; (iii) the majority of the catches of target species are preserved frozen or deep-frozen. A composite longline fishery was defined in each region (LL 1–4) aggregating the longline catch from all freezing longline fleets (principally Japan and Taiwan).

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

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

The region 1b purse-seine fisheries (log and free-school) were divided into three time periods: pre 2003, 2003–2006 and post 2006. This temporal structure was implemented due to the apparent change in the length composition of the catch from the purse-seine fisheries during the 2000s. The length of fish caught by the FAD fishery was generally smaller from 2007 onwards, while a higher proportion of smaller fish were caught by the free-school fishery prior to 2003.

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

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

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

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

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

Selectivity

Fishery selectivity is assumed to be age-specific and time-invariant. For the longline fisheries (LL 1a, 1b, 2, 3 and 4) a single selectivity is estimated that is shared among the five fisheries. The selectivity is also shared by the two sets of LL CPUE indices. The longline selectivity was parameterised with a logistic function that constrains the older age classes to be fully selected (“flat top”). The selectivity of the fresh tuna longline fishery (LF4) was estimated using a separate logistic function.

The free-school (FS) and FAD (LS) purse seine fisheries within region 1b were divided into three time periods (pre 2003, 2003–2006 and post 2006) based on the observation that the size of fish caught differed between these periods. Earlier stock assessments had estimated separate selectivities for each time period (and fishery). However, the stock assessment results were relatively insensitive to the temporal changes in selectivity and, these changes in selectivity were associated with the tag data set and, specifically, the apparent recovery of fish at liberty for extended periods (2-3 years) from the purse-seine FAD fishery. For simplicity, a single selectivity was estimated for each method (FS and LS) for the three time periods. The corresponding purse-seine method selectivities were also shared with the purse-seine fisheries in region 2 and region 4. However, until now this selectivity was estimated based on age, but this was modified to be based on length.

The two purse seine selectivities (FS and LS) were formulated using a cubic spline interpolation with five nodes. The nodes were specified to approximate the main inflection points of the selectivity function. This formulation was sufficiently flexible to provide a reasonable representation of the modal structure of the length composition of the catch from the two purse seine methods. This selectivity was also modified to be based on length.

For the other fisheries, selectivity was parameterised using a double-normal function (Methot 2013). The baitboat and handlines fishery length data are patchily distributed, so a constant was added (0.01) to the length composition and thus the residuals pattern were improved.

No length frequency data are available for the “Other” fishery in region 1a, while limited data are available from the OT 4 fishery. Similarly, size data were available from the troll fishery in region 4, but not from the fisheries in regions 1b and 2. Due to the limitations in the data in the “other” fishery (region 1a and 4) and in the troll fishery (region 1b and 4) and considering that the selectivity estimates for both fishery were very similar in the reference model of 2018 YFT model, then we simplified the parameterization of the model and assumed a common selectivity for them.

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).

Catch

Catch data were compiled based on the fisheries definitions. A preliminary update of quarterly catches by fishery was provided by the IOTC Secretariat, including catches from 2018. The catches were compiled in the file IOTC-2019-WPTT21-DATA15_SA_0. Given the dramatic changes in fishing operations of the Taiwanese fleet in recent years, using Taiwanese fleets as proxy for estimating Indonesian catches is no longer considered appropriate (Geehan & Braham 2018). Therefore, the file does not contain size data for the Taiwanese longline fleet for the period 2002-2018 and fresh tuna longline fleet for the period 2010-2018.

For each fishery, the time series of catches were very similar to the catch series included in the 2018 assessment (Table 5, 11). Total annual catches for 2017 and 2018 included in the updated catch history are 423,814mt and 401,384 mt, respectively (Table 5). The total catch in 2017 represents a 1% increase from the 2014 catch level, or a 4% increase from 2015.

Table 5. Recent yellowfin tuna catches (mt) by fishery included in the stock assessment model. The annual catches are presented for 2014- 2018.

| Fishery | 2,014 | 2,015 | 2,016 | 2,017 | 2,018 |
|--------------------------|--------|--------|--------|--------|--------|
| 1. GI 1a | 56,735 | 59,906 | 56,317 | 68,749 | 76,033 |
| 2. HD 1a | 71,919 | 73,998 | 86,014 | 65,487 | 65,058 |
| 3. LL 1a | 449 | 342 | 463 | 338 | 357 |
| 4. OT 1a | 1,293 | 997 | 1,228 | 923 | 1,609 |
| 5. BB 1b | 20,542 | 17,642 | 12,391 | 18,370 | 20,030 |
| 6. PS FS 1b 2003-06 | 0 | 0 | 0 | 0 | 0 |
| 7. LL 1b | 7,566 | 9,023 | 11,700 | 10,237 | 12,122 |
| 8. PS LS 1b 2003-06 | 0 | 0 | 0 | 0 | 0 |
| 9. TR 1b | 1,526 | 2,401 | 4,364 | 2,846 | 4,328 |
| 10. LL 2 | 6,329 | 6,331 | 6,233 | 6,802 | 7,225 |
| 11. LL 3 | 471 | 992 | 482 | 332 | 1,081 |
| 12. GI 4 | 14,551 | 11,180 | 8,313 | 5,356 | 7,349 |
| 13. LL 4 | 795 | 1,166 | 473 | 444 | 405 |
| 14. OT 4 | 9,869 | 10,501 | 8,649 | 11,926 | 13,681 |
| 15. TR 4 | 18,967 | 12,519 | 14,472 | 7,774 | 11,808 |
| 16. PS FS 2 | 205 | 1,464 | 1,998 | 3,060 | 134 |
| 17. PS LS 2 | 461 | 4,704 | 7,134 | 7,482 | 4,563 |
| 18. TR 2 | 1,772 | 1,695 | 3,229 | 2,392 | 3,025 |
| 19. PS FS 4 | 0 | 60 | 3 | 1,248 | 5 |
| 20. PS LS 4 | 405 | 278 | 710 | 1,977 | 1,243 |
| 21. PS FS 1b pre 2003 | 0 | 0 | 0 | 0 | 0 |
| 22. PS LS 1b pre 2003 | 0 | 0 | 0 | 0 | 0 |

| | | | | | |
|---------------------------|---------|---------|---------|---------|---------|
| 23. PS FS 1b post 2006 | 47,222 | 62,439 | 47,460 | 46,392 | 14,970 |
| 24. PS LS 1b post 2006 | 85,505 | 73,413 | 91,423 | 85,021 | 110,521 |
| 25. LF 4 | 50,593 | 40,487 | 46,278 | 54,228 | 68,267 |
| Total | 397,175 | 391,538 | 409,334 | 401,384 | 423,814 |

Length composition

Available length-frequency data for each of the defined fisheries were compiled into 48 4-cm size classes (10-14 cm to 198-202 cm), the reference model of 2018 had length bin of 2 cm, so the total number of bins was reduced to a half. Each length frequency observation for purse seine fisheries represents the number of fish sampled raised to the sampling units (sets in the fish compartment) while for fisheries other than purse seine each observation consisted of the actual number of yellowfin tuna measured. A graphical representation of the availability of length samples is provided in Figure 12.

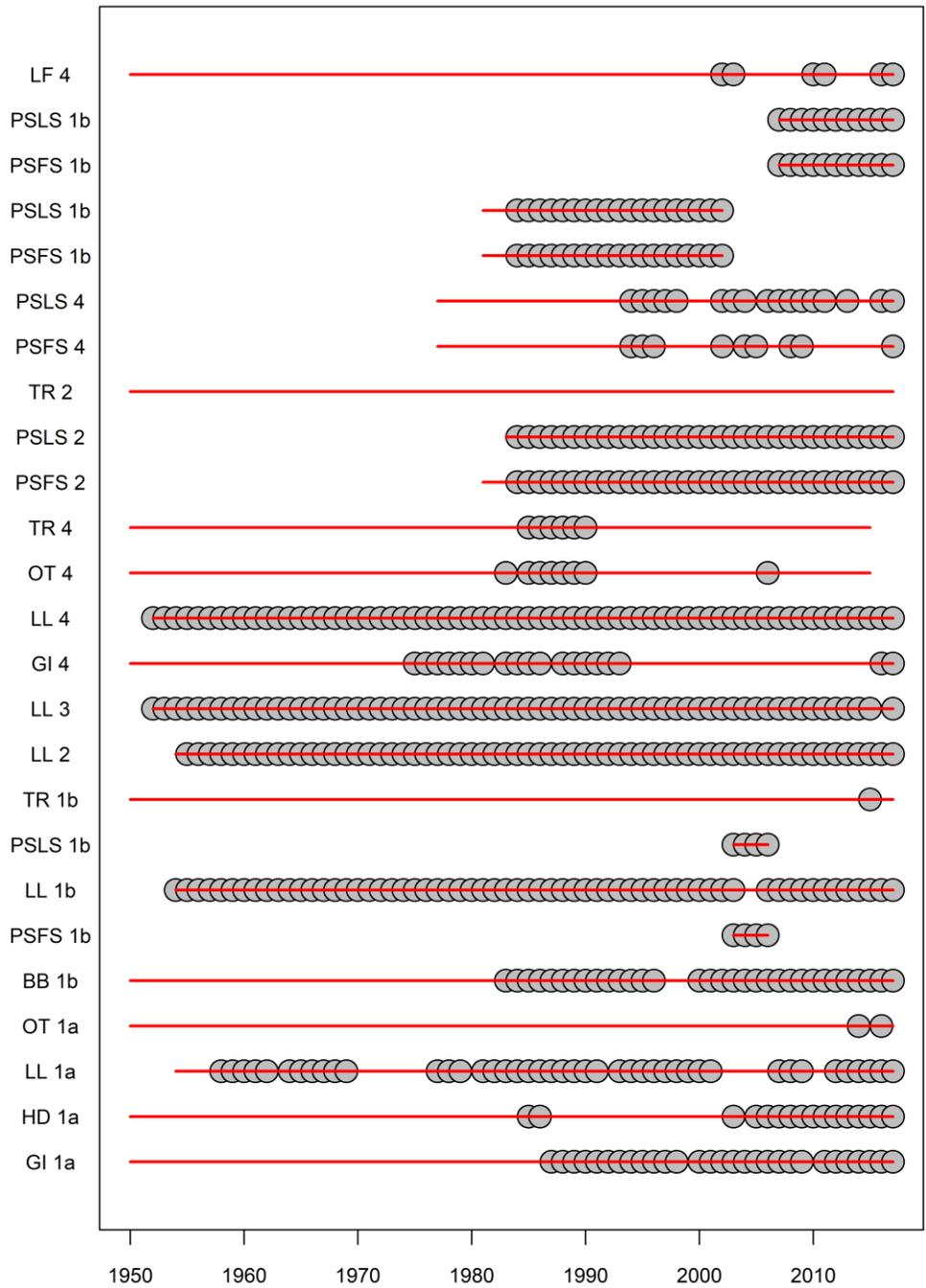


Figure 12. The availability of length sampling data from each fishery by year. The grey circles denote the presence of samples in a specific year. The red horizontal lines indicate the time period over which each fishery operated.

The length samples are not available for TR 2. The data were collected from a variety of sampling programs, which can be summarized as follows:

Purse seine: Length-frequency samples from purse seiners have been collected from a variety of port sampling programs since the mid-1980s. The samples are comprised of very large numbers of individual fish measurements. The length frequency samples are available by set type with associated sets catches typically composed of smaller fish than free school catches (Figure 13). There was a decline in the average length of fish from the FAD schools from 1985 to 2015 (Figure 14). The size composition of the catch from the free-school fishery is bimodal, being comprised of the smaller size range of yellowfin and a broad mode of larger fish (Figure 13). There is a considerable catch of smaller fish taken during free school fishing operation in the Mozambique Channel area in region 2 (Chassot 2014). The free-school fishery in region 4 appears to catch larger fish.

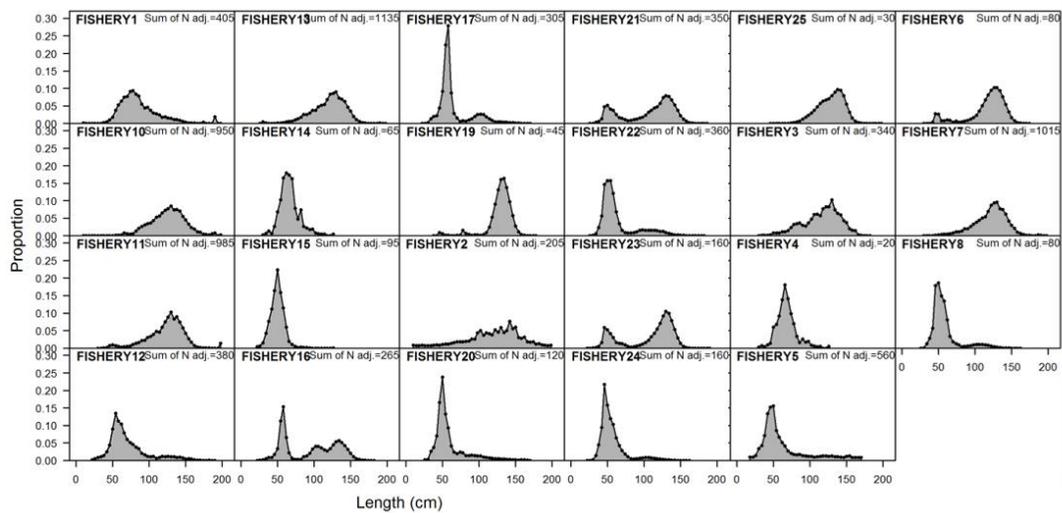


Figure 13. Length compositions of yellowfin tuna samples aggregated by fishery.

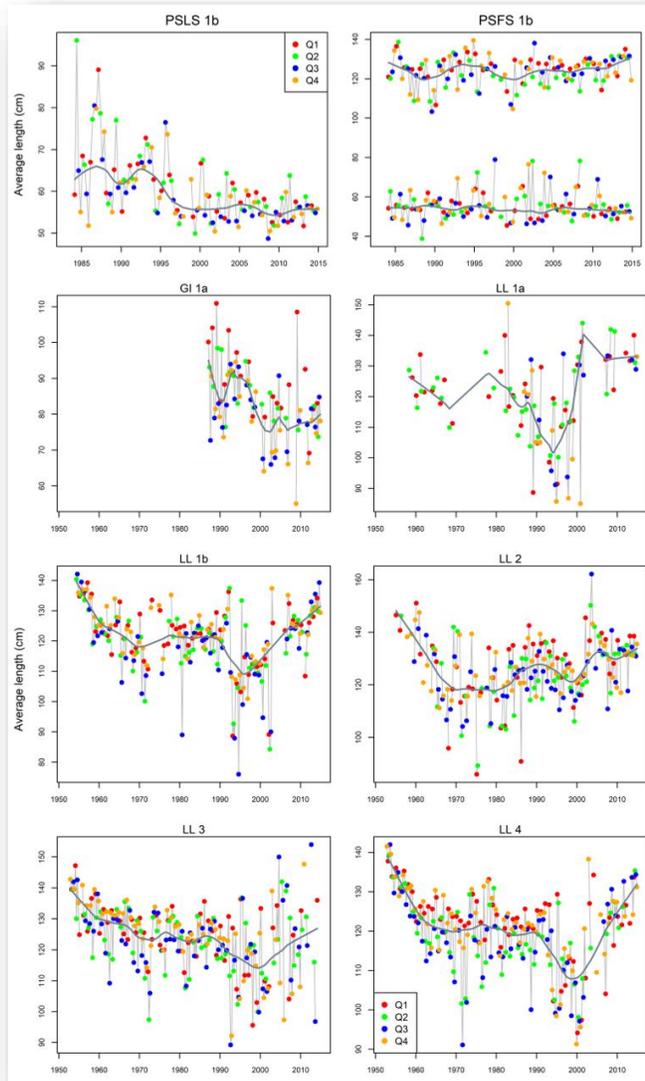


Figure 14. Mean length (fork length, cm) of yellowfin sampled from the principal fisheries (GI 1a, LL 1a-4, PSLS 1b and PSFS 1b) by year quarter. The grey line represents the fit of a lowess smoother to each data set. For PSFS 1b, the mean is calculated for fish ≤ 80 cm and > 80 cm separately. (Update)

Longline freezing: Length and weight data were collected from sampling aboard Japanese commercial, research and training vessels. Weight frequency data collected from the fleet have been converted to length frequency data via a processed weight-whole weight conversion factor and a weight-length key. Length frequency data from the Taiwanese longline fleet from 1980-2003 are also included in the length frequency data set, although data from the more recent years were excluded due to concerns regarding the reliability of these data (Geehan & Hoyle 2013). Comparisons between size data collected from Taiwanese vessels by observers and logbooks since 2003 revealed that the

vessel masters reported considerably larger fish (Simon Hoyle pers. comm.). In recent years, length data are also available from other fleets (e.g. Seychelles).

Overall, the average length of yellowfin caught by the longline fleet is generally comparable among the regions. However, there is considerable temporal variation in the length of fish caught (Figure 14). For all longline fisheries there was a marked decline in the size of fish caught during the 1950s and 1960s, while the size of fish caught stabilised during the 1970s and 1980s. The average length of yellowfin was significantly lower during the 1990s and the early 2000s in most regions, primarily due to the considerably smaller fish being sampled by the Taiwanese fleets. A quick examination of the spatial coverage of the Taiwanese samples did not reveal any apparent anomaly. Hoyle et al. (2017) suggested the substantial changes in the Taiwanese mean sizes are likely due to sampling problems rather than changes in the size composition of the population.

Longline fresh: Length and weight data were collected in port, during unloading of catches, for several landing locations and time periods, especially on fresh-tuna longline vessels flagged in Indonesia and Taiwan/China (IOTC-OFCF sampling). However, the quality of these data is highly variable. Length data from 1998-2008 were included in the previous assessment. But most samples were subsequently found to be biased (F. Fiorellato per. comm., IOTC Secretariat). For the current assessment, only four years of data are included (2002, 2003, 2010 and 2011).

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

Baitboat: Size data are available from the fishery from 1983 to 2015.

Troll: No size data are available from the TR 2 fisheries. The size data are available from the TR 1b fisheries in 2015 only. The troll fishery in region 4 was sampled during two periods: 1985-1990 (Indonesian fishery) and 1994-2004 (Sri Lankan fishery). The samples from 1994-2004 were excluded from the current assessment

Handline: Limited sampling of the handline fishery was conducted over the last decade. Samples are available for the Maldivian handline fisheries for this period.

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

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

Indices

The main index for the yellowfin stock assessment are derived from the collaborative study from multiple longline fleets (REF Hoyle et al 2019, IOTC-WPM-10-16). The standardised CPUE indices were derived using generalized linear models (GLM) from operational longline catch and effort data provided by Japan, Korea, Taiwan, China, and Seychelles (Hoyle et al 2019). Cluster analyses of species composition data by vessel-month for each fleet and number of hooks between floats were used to separate datasets into fisheries understood to target different species. Selected clusters were then combined and standardized using generalized linear models. The Seychelles data were made available and were included in the indices that used clustering, but not in the analysis of HBF because data were only available from 2009. Different scenarios were analyzed considering the Taiwanese discards data. Yellowfin catch (numbers of fish) was the dependent variable of the positive catch model (lognormal error structure), while the presence/absence of yellowfin tuna in the catch was the dependent variable in the binomial model. In addition to the year-quarter, models included covariates for vessel identity, 5° square location, number of hooks, and either cluster (for region 2 and 3) or HBF (for regions 1 and 4). The data from region 1a is not included the standardizations and the indices for region 1b is assumed to index the abundance for region 1.

During the workshop dedicated to derive this index (REF IOTC-2019-WPTT21-INFO1) different CPUE indices were derived (for each region) based on the treatment of the fishing vessel variable, the targeting strategy or Taiwanese discard data (Hoyle et al 2019). Based on the results of the statistical analysis and diagnostics of the models, the recommended estimates by the workshop of longline indices are the combined two time series with no vessel ID and with vessel ID 1952-1978 and 1979-2018 and considering Taiwanese discard data from 2005 (Table 6). However, the recommended methodology to identify the fishing strategy in the temperate and tropical region is different; cluster analysis in the temperate regions and hooks between floats (HBF) in the tropical regions (Table 6).

Table 6. The sets of CPUE indices used for each area in the reference model. The CPUEs in each region were aggregated considering the scaling or weight for each CPUE depending on the region.

| Area-Model | Weight-CPUE region | Model variables | Indices series name |
|------------|--------------------|--|--|
| 1 | 1 | No cluster, HBF, Using Taiwanese data from 2005 and accounting discards. | <i>Joint_regY_R2_dellog_novess_5279_yq.csv</i> <i>Joint_regY_R2_dellog_vessid_79nd_yq.csv</i> |
| 1 | 2 | Cluster, no HBF, Using Taiwanese data from 2005 and accounting discards | <i>Joint_regY_R3_dellog_novess_5279_yq.csv</i> <i>Joint_regY_R3_dellog_vessid_79nd_yq.csv</i> |
| 2 | 3 | No cluster, HBF, Using Taiwanese data from 2005 and accounting discards | <i>Joint_regY_R4_dellog_vessid_79nd_yq.csv</i> |
| 2 | 4 | No Cluster, no HBF, Using Taiwanese data from 2005 cluster, HBF | <i>Joint_regY_R5_dellog_novess_5279_yq.csv</i> <i>Joint_regY_R5_dellog_vessid_79nd_yq.csv</i> |

The CPUE indices from the years prior to 1972 were not included in the assessment model (as in the previous assessments in 2015 and 2016). The CPUE indices from the earlier period are considerably higher than for the remainder of the 1970s. The decline in CPUE indices during the late 1960s–early 1970s is inconsistent with the relatively low level of catch taken during this period (REF Langley 2015). At the 10th WPTT, it was agreed that the decline in the CPUE indices was unlikely to be solely due to changes in stock abundance although the reason for this seemingly excessive initial decline are still poorly understood.

The updated CPUE indices are similar to those in the previous assessment in terms of the overall trend, but with a bit smaller peaks previous to 1980 (Figure 15). Then, the two indices defined in each model region were aggregated considering the scaling factor estimated for each region.

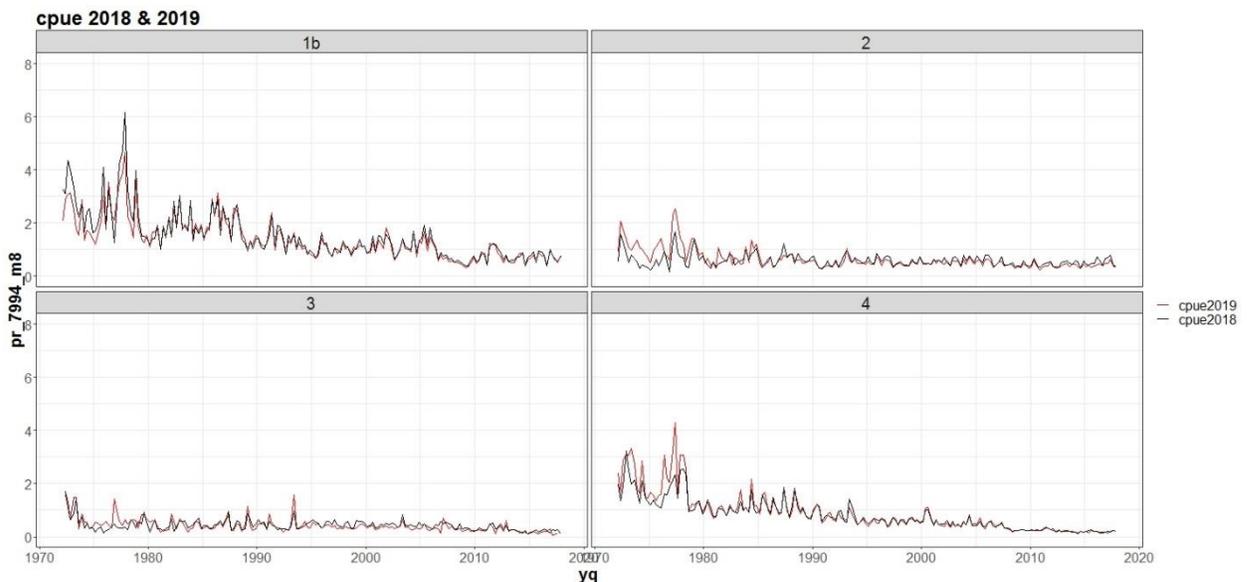


Figure 15. A comparison of the longline CPUE indices included in the 2018 stock assessment (black line) and the updated values in 2019 (red line). The final Indices for region 1 was based on data from region 1b only.

Two other additional indices were considered for inclusion as sensitivity runs; an acoustic index derived from echosounder receivers placed on FAD bouys prior to fishing (Santiago et al. 2019) and a purse seine free school index (Guery et al. 2019) that improves upon the definition of effort in the purse seine fishery. The Buoy index was linked to fishery the PS FAD fishery in each of the four quarters to inform the model on recruitment. The PS FS index was calculated separately for each quarter but since this fishery catches (and presumably the availability of YFT in the tropical region) peak in quarter 1. Hence the index of quarter 1 was the only input as the selectivity for the PSFS is for very large fish and the model has no ability to account for seasonal variation in the availability of large fish, other than through different selectivity and catchability for each quarter. As the decision was made to mirror selectivity across all four seasons, there was no strong reason to use all four quarters of the index, as the expected values of the index would be almost entirely parallel. The indices used in the current assessment are shown in Figure 16, longline for the reference case and the buoy and PS FS index for sensitivity analyses.

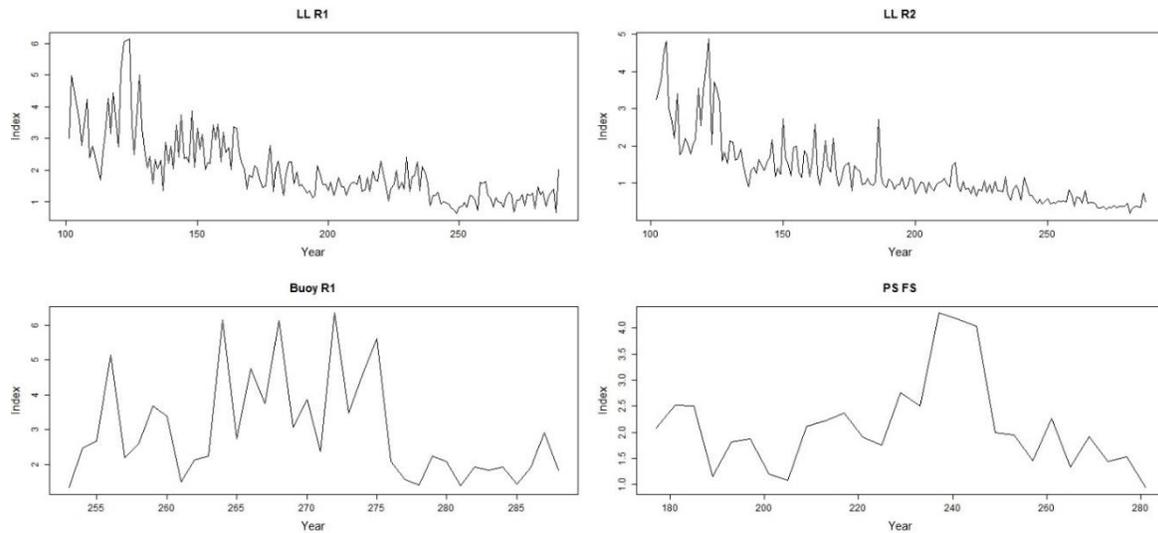


Figure 16. Longline CPUE indices and the CV-s (0.2) used in the reference case with an spatial structure of 2 areas, West (the left figure) and East (the right figure).

Regional weighting

For the regional longline fisheries, a common catchability coefficient (and selectivity) was used in the reference case in 2018 linking the respective CPUE indices among regions. This significantly increases the power of the model to estimate the relative (and absolute) level of biomass among regions. However, as CPUE indices are essentially density estimates it is necessary to scale the CPUE indices to account for the relative abundance of the stock among regions. For example, a relatively small region with a very high average catch rate may have a lower level of total biomass than a large region with a moderate level of CPUE. But then, the model can be very sensitive to scaling factor and in the case of the reference case, the model could not converge assuming the same catchability in both regions. Thus, the scaling factor is considered when the longline CPUEs are aggregated within the regions defined in the reference case but without assuming the same catchability between the reference case regions.

The approach used to determine regional scaling factors of the 4 areas of the 2018 reference case, incorporated both the size of the region and the relative catch rate to estimate the relative level of exploitable longline biomass among regions. This approach is similar to that used in the WCPO regionally disaggregated tuna assessments. During preliminary modelling, the scaling factors used in the previous assessment were considered for continuity. The scaling factors were derived from the Japanese longline CPUE data from 1963–75 by summing the average CPUE in each of the 5*5 latitude/longitude cells within a region.

For each of the principal longline fisheries, the GLM standardised CPUE index was normalised to the mean of the period for which the region scaling factors were derived (i.e. the GLM index from 1963–75). The normalised GLM index was then scaled by the respective regional scaling factor to account for the regional differences in the relative level of exploitable longline biomass among regions

However, these estimates were derived from a period in which the CPUE was considered unreliable (see the previous section). Hoyle & Langley (2018) revised the approach for estimating the regional weighing factors for IO tropical tuna species and proposed a set of alternative estimates for yellowfin based aggregated longline catch effort data. The author recommended the estimates by method '8' for the period 1979–1994 (referred to as '7994m8', see Table 2 of Hoyle (2018)) to be included in the current assessment. The relative scaling factors calculated for regions 1–4 are 1.674, 0.623, 0.455 and 1.000 respectively. The alternative sets of regional scaling factors derived from period 1975–1994 ("7595m8") and 1980 to 2000 ("8000m8") were also explored in the assessment.

A number of important trends are evident in the CPUE indices (see Hoyle et al. 2018b for more details).

- The CPUE indices in the tropical areas were characterized by very steep declines prior to 1975. From 1980–1989 the western tropical (region 1b) CPUE increased during the 1980s, then declined until 1995, increased again until 2005, and then decreased again. The low CPUE indices followed the period of exceptionally high catches from the purse seine fishery in region 1b during 2003–2005. The drop in CPUE occurred before the peak in the number of piracy incidents in the western Indian Ocean (2008–2011). After that time, it remained close to the lowest level observed.
- The eastern tropical region 4 followed a similar pattern until 1990 but then declined steadily, and by 2016 was also close to the lowest level in the time series. The recent decline in CPUE in this region is consistent with a decline in the proportion of yellowfin in the combined tuna catch from the Japanese longline fleet in the eastern Indian Ocean (see Figure 44 from Hoyle et al 2015). It is unclear whether the change in species proportion is related to a decline in the abundance of yellowfin in the region (relative to the other species) or a regional change in the targeting of the fishing fleet. However, there is an indication that there has been a differential shift towards deeper longline gear (greater HBF) in the eastern Indian Ocean since 2000 and this may indicate a shift in targeting toward bigeye tuna in this region (Hoyle pers. comm. additional JP LL analyses). Such factors may not be adequately accounted for in the standardisation of the yellowfin CPUE data.
- The CPUE indices in western temperate region 2 followed a similar pattern to the western tropical indices, with a decline until the mid-1970s followed by an increase until the late 1980s, and subsequently a slow decline with significant variability. However, the two sets of CPUE indices diverge somewhat from about 2007 with the CPUE indices from R2 being maintained at a higher level relative to R1.
- The CPUE indices from region 3 are low compared to the other three regions reflecting the low regional scaling factor. However, the overall trend in the CPUE indices is broadly comparable to the other regions. The eastern temperate region 3 the pattern was similar to the western temperate area before 1979. After 1979 catch rates increased until the mid-2000's, but then declined rapidly and reached their lowest observed levels by 2016.
- There is an exceptionally high peak in CPUE indices 1976–78 from region 1. Hoyle et al. (2017) showed this discontinuity exists in Japanese, Taiwanese and Korean data, and in multiple regions in multiple ocean. Hoyle et al (2017) suggested this is unlikely to be explained by changes to the population or catchability but may be associated with catch reporting and data management.

- The spike in the CPUE indices around 2012 in the west equatorial region (region 1) was evident for most fishing fleets. Several hypotheses has been proposed on what could have caused CPUE to have increased, including a return to fishing in areas that were most affected by piracy. However, further investigation is required.

Dynamics of tagged fish

Tag mixing

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

An analysis of the tag recovery data was undertaken to determine an appropriate mixing period for the tagging programme (Langley & Million 2012). The analysis revealed that the tag recoveries from the FAD purse-seine fishery were not adequately mixed, at least during the first 6 months following release. Conversely, the free-school tag recoveries indicate a higher degree of mixing within the fished population. Most of the tagged yellowfin were in the length classes that are not immediately selected by the free-school fishery (< 90 cm). A mixing period of about 6–12 months is of sufficient duration for most tagged fish to recruit to free-school fishery (> 90 cm) and no longer be vulnerable to the FAD fishery. However, the maximum displacements of tags reach a plateau within a few weeks of release (Figure A1 Appendix A), suggesting rapid movement of yellowfin within the tag release/recovery areas. On basis of the above, it was considered that a mixing period of three quarters was probably sufficient to allow a reasonable degree of dispersal of tagged fish amongst the yellowfin tuna population within the primary region of release. The distribution of annual RTTP tag returns from the main recovery period (2006-2009) are shown in Figure A2, Appendix A.

The release phase of the tagging programme was essentially restricted to the western equatorial region. The examination of the tag recoveries of bigeye tuna from the PSLS fishery identified considerable differences in the recovery rate (number of tags per tonne of catch) amongst latitudinal zones for tags at liberty for at least 12 months (Langley 2016b). In an attempt to account for the

incomplete mixing of tagged fish, the bigeye assessment model further partitioned the western equatorial region into two regions along the equator. A similar analysis was performed to yellowfin tag data, however, the results indicated that the recovery rate of tags after 3 quarters at liberty was similar both in trend and magnitude between latitude band 0 – 10N and 0 – 10S within the western equatorial region (Figure A3 Appendix A). This suggested a reasonable degree of mixing of tagged fish at the regional scale. Nonetheless, a sensitivity model that further partitions the western equatorial region is still considered in the exploratory modelling.

The distribution of tags throughout the wider IO appears to have been relatively limited as is evident from the low number of tag recoveries from the fisheries beyond region 1b. Tag recoveries from beyond region 1 and 2 are unlikely to significantly inform the model regarding movement rates given the lack of information concerning reporting rates of tags for these fisheries (see below).

Tag reporting

Estimates of tag reporting rates from the purse seine fishery were available from tag seeding trials. For the other fisheries, the 2018 reference case had very limited information available to indicate the tag reporting rates and fishery specific reporting rates were estimated based on uninformative priors.

Data weighting, parameters, and likelihood

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

Following the previous assessment, the weighting of the CPUE indices followed the approach of Francis (2011). A series of smoother lines were fitted to the CPUE index and the RMSE of the resulting fit to each set of CPUE indices was determined as a measure of the magnitude of the variation of each set of indices CPUE indices. The resulting RMSEs were relatively high (0.40–0.50). However, a significant proportion of this variation is related to the relatively high seasonal variation in CPUE in most regions. The analysis performed to the annualised CPUE index (Hoyle et al. 2018) resulted in considerable reduction in the RMSEs (0.15-0.2). On that basis, a CV of 0.2 was assigned to each set of CPUE indices in the base model, to ensure the stock biomass trajectories were broadly consistent with the CPUE indices while allowed for a moderate degree of variability in fitting to the indices (a CV of 0.3 was used in the previous assessment).

The CVs of purse seiner free school it as assumed the average of the all times series as well as for the indices estimated by buoy echosounder, 0.2 and 0.3 respectively.

The relative weighting of the tagging data was controlled by the magnitude of the over-dispersion parameters assigned to the individual tag release groups. In the previous assessment, the over-dispersion parameters for all tag release groups were set at 7.0 - determined iteratively from the

residuals of the fit to the tag recovery data (observed – expected number of tags recovered). The same value was used in the current assessment.

The reliability of the length composition data is variable across fisheries and over time periods. For that reason, it was considered that the length composition data should not be allowed to dominate the model likelihood and directly influence the trends in stock abundance. Following the previous assessment, an overall effective sample size (ESS) of 5 was assigned to all length composition observations (all fisheries, all time periods) following the Francis (2011) method. This essentially gave the entire length composition data set a relatively low weighting in the overall likelihood. Nonetheless, due to the magnitude of the length composition data, these data were sufficiently informative to provide reasonable estimates of fishery selectivity and provide some information regarding recruitment trends.

The weightings were applied by the values assigned to components of the likelihood of each observational dataset included in the total model likelihood. a default lambda of 1.0, represented the native weighting of the data. A lower value of Lambda would effectively downweighting the dataset relative to other observations, effectively reducing its influence on the overall model estimates. The big influence of tagging data was shown with 2018 reference case. Thus, a lambda value of 0.1 was applied to the tagging data, because the tagging data were mainly focused on the western tropical region, with few recovery rates in other regions with fleets different to purse seiners.

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

Model Diagnostics

Model convergence was assessed using several means. The first diagnostic was whether the Hessian, (i.e., the matrix of second derivatives of the likelihood with respect to the parameters) inverts. The second measure is the maximum gradient component which, ideally, should be low (<0.0001 is a standard value). The third diagnostic involved altering or jittering the starting values of the parameters to evaluate whether the model converges to a global solution, rather than a local minimum.

Other diagnostics included likelihood profiling of key parameters (steepness, and sigmaR), evaluation of fits to residuals for indices and length composition, retrospective analyses and sensitivity to different indices and compositional data inputs. Retrospective analyses are also standard diagnostic practice and was conducted on the reference case for 5 year retrospective peels.

Parameters Estimated

Overall the reference model have 204 estimated parameters, consisting of 40 selectivity parameters, 2 stock recruitment parameters, 2 catchability of the longline CPUEs and 164 recruitment deviations (Table 7). For purse seiners cubic spline parameters Beta prior distributions were used to aid in model stability. Parameter estimates, standard errors and prior distributions for reference case are shown,

(Table 7) results are similar for most parameters across the other models and are not shown here for brevity.

Table 7. Estimated parameters, phase of estimation, CV, gradient and priors, if used.

| | Value | Phase | Min | Max | Init | Status | Parm_StDev | Gradient | Pr_type | Prior | Pr_SD | Pr_Like | Afterbo |
|----------------------------------|-----------|-------|--------|-------|-----------|--------|------------|-------------|----------|-------|-------|-----------|---------|
| RecrDist_Area_2 | 1.015310 | 1 | 5.000 | 5.00 | 0.820000 | OK | 0.128210 | 0.00015413 | No_prior | NA | NA | NA | OK |
| SR_LN(R0) | 11.662100 | 1 | 2.000 | 25.00 | 10.000000 | OK | 0.0549498 | 0.000610901 | Normal | 10.0 | 5.000 | 0.0552522 | OK |
| LnQ_base_SURVEY1(26) | 9.368380 | 1 | 25.000 | 25.00 | 4.0181700 | OK | 0.1092300 | 0.000124573 | No_prior | NA | NA | NA | OK |
| LnQ_base_SURVEY2(27) | 8.587230 | 1 | 25.000 | 25.00 | 4.0181700 | OK | 0.1155250 | 0.00019677 | No_prior | NA | NA | NA | OK |
| SizeSpline_GradLo_FISHERY6(6) | 0.278577 | 3 | 0.001 | 1.00 | 0.270000 | OK | 0.0590104 | 3.9432e-05 | Sym_Beta | 1.0 | 0.001 | 0.0002162 | OK |
| SizeSpline_Val_1_FISHERY6(6) | 17.326400 | 2 | 35.000 | 0.00 | 15.000000 | OK | 4.9635700 | 5.98018e-05 | No_prior | NA | NA | NA | OK |
| SizeSpline_Val_2_FISHERY6(6) | 0.370912 | 2 | 10.000 | 7.00 | 0.0715442 | OK | 0.1204290 | 1.26991e-05 | No_prior | NA | NA | NA | OK |
| SizeSpline_Val_4_FISHERY6(6) | 2.253070 | 2 | 10.000 | 7.00 | 0.6151260 | OK | 0.1337590 | 6.80252e-05 | No_prior | NA | NA | NA | OK |
| SizeSpline_Val_5_FISHERY6(6) | 3.651200 | 2 | 10.000 | 7.00 | 5.1733600 | OK | 0.1517930 | 0.000127462 | No_prior | NA | NA | NA | OK |
| SizeSpline_GradLo_FISHERY8(8) | 0.390558 | 3 | 1.000 | 1.00 | 0.8956210 | OK | 0.2280470 | 0.000312244 | Sym_Beta | 1.0 | 0.001 | 0.0001653 | OK |
| SizeSpline_GradHi_FISHERY8(8) | 0.274160 | 3 | 1.000 | 0.02 | 0.1020200 | OK | 0.0575148 | 0.000239543 | No_prior | NA | NA | NA | OK |
| SizeSpline_Val_2_FISHERY8(8) | 0.353260 | 2 | 10.000 | 7.00 | 0.2916700 | OK | 0.0370511 | 6.30143e-06 | No_prior | NA | NA | NA | OK |
| SizeSpline_Val_4_FISHERY8(8) | 0.288622 | 2 | 10.000 | 7.00 | 0.5313050 | OK | 0.1181350 | 3.23266e-05 | No_prior | NA | NA | NA | OK |
| SizeSpline_Val_5_FISHERY8(8) | 1.646730 | 2 | 10.000 | 7.00 | 0.3054280 | OK | 0.4837740 | 0.000203367 | No_prior | NA | NA | NA | OK |
| Age_DbIN_peak_FISHERY1(28) | 8.975170 | 3 | 1.000 | 12.00 | 7.000000 | OK | 0.5102520 | 2.88857e-06 | Normal | 7.0 | 3.000 | 0.2167380 | OK |
| Age_DbIN_ascend_se_FISHERY1(28) | 0.333572 | 4 | 10.000 | 9.00 | 1.000000 | OK | 1.0329200 | 0.000218736 | Normal | 1.0 | 3.000 | 0.0246737 | OK |
| Age_DbIN_descend_se_FISHERY1(28) | 2.008530 | 4 | 5.000 | 9.00 | 3.000000 | OK | 0.4118760 | 6.5834e-05 | Normal | 3.0 | 1.000 | 0.4915060 | OK |

| | Value | Phase | Min | Max | Init | Status | Parm_StDev | Gradient | Pr_ty | Prior | Pr_SD | Pr_Like | Afterbo |
|-----------------------------------|-----------|-------|--------|-------|-----------|--------|------------|--------------|--------|-------|-------|-----------|---------|
| Age_DbIN_end_logit_FISHERY1(28) | 1.910180 | 5 | 9.000 | 5.00 | 2.000000 | OK | 0.3091500 | -0.000712858 | Normal | -2.0 | 1.000 | 0.0040339 | OK |
| Age_DbIN_peak_FISHERY2(29) | 20.702500 | 3 | 1.000 | 40.00 | 10.000000 | OK | 1.7775000 | 6.10608e-05 | Normal | 10.0 | 5.000 | 2.2908600 | OK |
| Age_DbIN_ascend_se_FISHERY2(29) | 3.528080 | 4 | 10.000 | 9.00 | 1.000000 | OK | 0.3271840 | 3.22858e-06 | Normal | -1.0 | 3.000 | 1.1390800 | OK |
| Age_DbIN_descend_se_FISHERY2(29) | 3.581700 | 4 | 5.000 | 9.00 | 3.000000 | OK | 0.9577330 | -7.06012e-05 | Normal | 3.0 | 1.000 | 0.1691890 | OK |
| Age_DbIN_end_logit_FISHERY2(29) | 0.847014 | 5 | 9.000 | 5.00 | 2.000000 | OK | 1.1995100 | 0.00021252 | Normal | -2.0 | 1.000 | 0.6646890 | OK |
| Age_inflection_FISHERY3(3) | 13.075900 | 3 | 8.000 | 18.00 | 14.000000 | OK | 0.1420240 | 4.58662e-06 | Normal | 14.0 | 2.000 | 0.1067420 | OK |
| Age_95%width_FISHERY3(3) | 3.565490 | 3 | 2.000 | 6.00 | 4.000000 | OK | 0.1211930 | -0.000246162 | Normal | 4.0 | 1.000 | 0.0943986 | OK |
| Age_DbIN_peak_FISHERY4(31) | 4.452490 | 3 | 1.000 | 10.00 | 2.500000 | OK | 0.2349060 | -8.36439e-05 | Normal | 2.5 | 1.000 | 1.9061200 | OK |
| Age_DbIN_ascend_se_FISHERY4(31) | 2.024040 | 4 | 7.000 | 3.00 | 2.000000 | OK | 0.9837860 | 3.4918e-06 | Normal | -2.0 | 1.000 | 0.0002890 | OK |
| Age_DbIN_descend_se_FISHERY4(31) | 2.609900 | 4 | 5.000 | 9.00 | 4.000000 | OK | 0.2235480 | -2.39179e-05 | Normal | 4.0 | 1.000 | 0.9661890 | OK |
| Age_DbIN_end_logit_FISHERY4(31) | 4.512340 | 5 | 9.000 | 7.00 | 1.000000 | OK | 0.8149120 | 3.77143e-05 | Normal | 1.0 | 2.000 | 3.7982400 | OK |
| Age_DbIN_peak_FISHERY5(32) | 3.354290 | 3 | 1.000 | 10.00 | 3.000000 | OK | 0.1739010 | -6.33613e-05 | Normal | 3.0 | 1.000 | 0.0627602 | OK |
| Age_DbIN_ascend_se_FISHERY5(32) | 2.029330 | 4 | 7.000 | 5.00 | 2.000000 | OK | 0.9788980 | -2.03334e-05 | Normal | -2.0 | 1.000 | 0.0004301 | OK |
| Age_DbIN_descend_se_FISHERY5(32) | 2.985290 | 4 | 5.000 | 9.00 | 3.000000 | OK | 0.2275740 | 0.000266532 | Normal | 3.0 | 1.000 | 0.0001082 | OK |
| Age_DbIN_end_logit_FISHERY5(32) | 3.428900 | 5 | 9.000 | 9.00 | 3.000000 | OK | 0.7259280 | -3.72618e-07 | Normal | -3.0 | 1.000 | 0.0919787 | OK |
| Age_DbIN_peak_FISHERY12(39) | 7.793600 | 3 | 1.000 | 10.00 | 5.500000 | OK | 0.4571910 | 1.45017e-05 | Normal | 5.5 | 2.000 | 0.6575720 | OK |
| Age_DbIN_ascend_se_FISHERY12(39) | 1.556180 | 4 | 10.000 | 5.00 | 2.000000 | OK | 0.3863530 | -5.21253e-05 | Normal | 2.0 | 2.000 | 0.0246224 | OK |
| Age_DbIN_descend_se_FISHERY12(39) | 0.443368 | 4 | 5.000 | 9.00 | 1.000000 | OK | 0.7582400 | -7.16927e-05 | Normal | 1.0 | 2.000 | 0.0387299 | OK |
| Age_DbIN_end_logit_FISHERY12(39) | 1.766870 | 5 | 9.000 | 9.00 | 2.000000 | OK | 0.2178700 | -5.86402e-05 | Normal | -2.0 | 1.000 | 0.0271748 | OK |
| Age_inflection_FISHERY25(25) | 14.252500 | 3 | 5.000 | 20.00 | 14.000000 | OK | 0.7953200 | 9.65403e-05 | Normal | 14.0 | 1.000 | 0.0318876 | OK |
| Age_95%width_FISHERY25(25) | 3.574100 | 3 | 2.000 | 10.00 | 4.000000 | OK | 0.7524940 | 1.4853e-05 | Normal | 4.0 | 1.000 | 0.0906973 | OK |

| | Value | Phase | Min | Max | Init | Status | Parm_StdDev | Gradient | Pr_type | Prior | Pr_SD | Pr_Like | Afterbo |
|--|-----------|-------|--------|--------|------------|--------|-------------|--------------|---------|-------|-------|-----------|---------|
| Age_DblN_peak_FISHERY1(28)_BLK1repl_213 | 7.490610 | 3 | 1.000 | 12.000 | 7.0000000 | OK | 0.5261240 | 3.82629e-05 | Normal | 7.0 | 3.000 | 0.0133724 | OK |
| Age_DblN_peak_FISHERY1(28)_BLK1repl_261 | 8.270900 | 3 | 1.000 | 12.000 | 7.0000000 | OK | 0.2627360 | 7.42141e-05 | Normal | 7.0 | 3.000 | 0.0897326 | OK |
| Age_DblN_ascend_se_FISHERY1(28)_BLK1repl_213 | 1.394560 | 4 | 10.000 | 9.000 | 1.0000000 | OK | 2.3508900 | -0.000372056 | Normal | 1.0 | 3.000 | 0.0086487 | OK |
| Age_DblN_ascend_se_FISHERY1(28)_BLK1repl_261 | 2.765200 | 4 | 10.000 | 9.000 | 1.0000000 | OK | 2.1164500 | -1.09447e-05 | Normal | 1.0 | 3.000 | 0.1731080 | OK |
| Age_DblN_peak_FISHERY2(29)_BLK2repl_225 | 12.729700 | 3 | 1.000 | 40.000 | 10.0000000 | OK | 1.6851300 | -0.000102414 | Normal | 10.0 | 5.000 | 0.1490290 | OK |
| Age_DblN_ascend_se_FISHERY2(29)_BLK2repl_225 | 1.227800 | 4 | 10.000 | 9.000 | 1.0000000 | OK | 1.3439500 | -0.000196566 | Normal | 1.0 | 3.000 | 0.2757280 | OK |

Benchmark and fishing mortality calculations

For overall fishing mortality rate, the exploitation rate in biomass was used, similar to the 2018 assessment. 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).

Uncertainty Quantification

Development of a reference case

Initially the model structure was designed to be similar to the 2018 assessment (Dan Fu et al. 2018) and a series of stepwise changes were made (Table 8). The results of the reference case are very similar to the reference case of 2018 assuming data weighting of 0.1 on tagging data (Figure 17), but when tagging data are not taken into account then the virgin biomass is higher.

Table 8. List of model changes.

| | |
|--|---|
| 1. Convert 2018 reference model from SS 3.24 to 3.30 (changes in boundaries movement rates, and survey catchability becomes a parameter. | done, get some differences, probably due to some local minimum in movement rates. |
| 2. Jitter analysis of the 2018 model in version v3.30 | The model is sensitive to the initial values |
| 3. Modify the spatial structure of the model, and compare, 2area, 3area and 4 area model | 3B model did not converge |
| <ul style="list-style-type: none"> - 2 area model without migration and without tagging data: East and West, and one survey in each region (R1b and R4 CPUE). No migration. - 3 area model (A): Region 1, Region 2, Region 3+4 (R1b, R2 and R4 CPUE) - 3 area model (B): Region 1+Region 2, Region 3,4 (R1b, and R4 CPUE, recruitment in the three regions, keep only the tagging data of PS) - 4 area model: Region 1, Region 2, Region 3+4 (R1b,R2,R3 and R4 CPUE) | done |

| | |
|---|---|
| 4. Jitter of each model; 2 area model is the most robust model and it becomes the spatial structure of the reference case | |
| 5. Catchability of each survey is estimated as different parameter | The biomass in each region changes depending if q is the same for both surveys or not. |
| 6. Add a constant of 0.01 to baitboat and handlines due to the patchy distribution | Residual pattern is improved. |
| 7. Purse seiners seem to catch not so much big fish as the selectivity estimates, try to add a knot in the oldest ages in cubic spline to reduce catches in the largest | The residual pattern not improved. |
| 8. Changes in purse seiners selectivity from based on age to based on length. | A little increase in likelihood. |
| 9. Mirror the selectivity of others in region 1 and 4 (very few data) with the troll fleet in region 1b (no data) (fixed selectivity). | The model does not fit the tagging data |
| 10. Compare the estimates of the model with the tagging data | |
| 11. Update catches | done |
| 12. Update length compositions | done |
| 13. Change the longline joint index, in the model region 1, the scaled index region 1b +region 2 (CPUE regions definition) and in the model region 2, the scaled index region 3 + region 4. | done |
| 14. Explore adding a time block in the longlines, due to the positive residual pattern in the largest the last years. | Not success, because they already catch big ones, so the residual pattern is still there. |
| 15. Remove longline length frequencies from 272 due to weird residual pattern | Done, results very sensitive to the new length composition from longlines. |
| 16. Remove the length composition from GI region 4 and fresh longline fleets region 4 | The results are very sensitive to the combination of different fleets. |
| 17. Remove all the length composition from 272 (2015-2018). | done |
| 18. Explore the priors of the fleet others in region 1a, the prior and posterior the same. | Exploring |

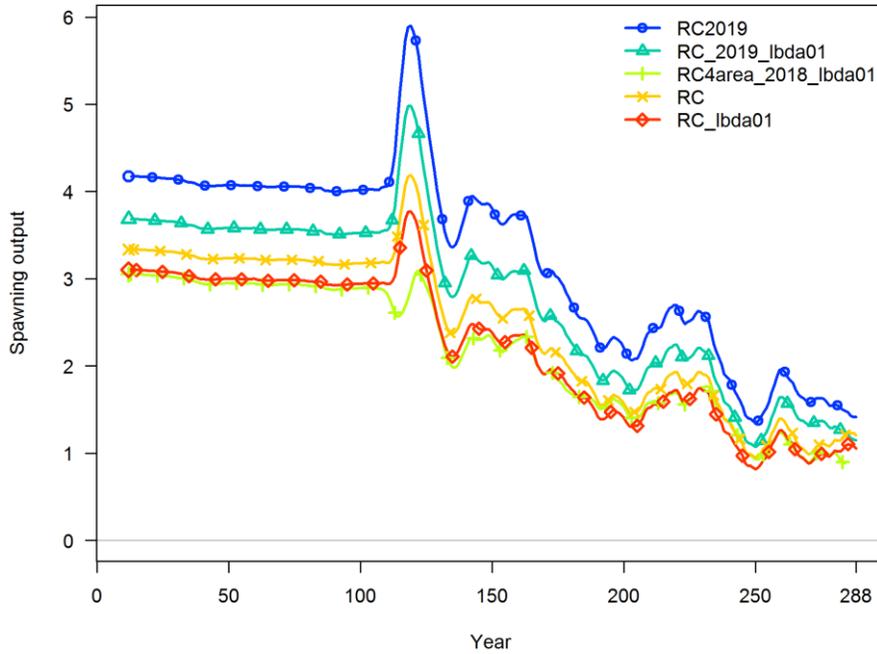


Figure 17. Spawning biomass trajectories from the reference case and the sensitivities related with the update of the model in comparison of the reference case of 2018.

Development of sensitivity runs

A series of 21 runs (Table 9) were analysed as sensitivity analysis. Subsequent sensitivity runs will undoubtedly be conducted at the assessment workshop and will be documented in an appendix to this paper.

Table 9. Table of run specifications, likelihoods and gradients.

| scenario | LL | grad | hessian | ssb0 | nparam | AIC |
|-----------------------|---------|------------|---------|---------|--------|------------|
| RC | 3018.66 | 0.00076316 | yes | 3319030 | 208 | 399.974863 |
| RC_2019 | 3526.26 | 0.00088973 | yes | 4170730 | 208 | 399.664014 |
| RC4area_2018 | 8811.88 | 0.00085372 | yes | 2625460 | 248 | 477.832288 |
| RC_2019_lbda01 | 4133.48 | 0.00072192 | yes | 3669560 | 225 | 433.34625 |
| RC4area_2018_lambda01 | 3626.34 | 0.00084233 | yes | 3050820 | 248 | 479.608042 |
| RC_lbda01 | 4133.48 | 0.00072192 | yes | 3669560 | 225 | 433.34625 |
| RC_h07 | 3017.12 | 0.00083714 | yes | 3554400 | 208 | 399.975884 |
| RC_h09 | 3020.13 | 0.00092404 | yes | 3154060 | 208 | 399.97389 |
| RC_sigmaR04 | 2982.47 | 0.00099316 | yes | 2949840 | 208 | 399.998986 |
| RC_sigmaR05 | 3001.92 | 0.00094639 | yes | 3100580 | 208 | 399.985985 |
| RC_sigmaR065 | 3026.29 | 0.0009501 | yes | 3458560 | 208 | 399.969815 |
| RC_Mp | 3005.02 | 0.00093151 | yes | 3120670 | 209 | 401.983921 |
| RC_Mp_lbda01 | 3602.37 | 0.00061198 | yes | 3006230 | 226 | 435.621306 |

| | | | | | | |
|----------------|---------|-------------|-----|---------|-----|-------------|
| RC_Mhigh | 3060.85 | 0.00082429 | yes | 2774030 | 208 | 399.947104 |
| RC_G | 3175.05 | 0.0006975 | yes | 2778230 | 208 | 399.873843 |
| RC_G_Mp | 3074.33 | 0.00087002 | yes | 1797030 | 209 | 401.938315 |
| RC_buoy | 3003.45 | 0.00074816 | yes | 3134640 | 209 | 401.984966 |
| RC_buoy_PSFS | 3022.23 | 0.00075087 | yes | 3191630 | 210 | 403.9725 |
| RC_8mix_lbda01 | 3259.1 | 0.000951863 | yes | 3169590 | 225 | 433.8215873 |
| RC_8mix_lbda1 | 5334.57 | 0.000860272 | yes | 2723880 | 225 | 432.8360729 |
| RC_lbda1 | 8904.92 | 0.000894477 | yes | 2577260 | 225 | 431.8112816 |

Results

Building the reference case and update

The 2018 reference model has a spatial structure of 4 areas and the fleets are also define considering this spatial structure: East-tropical, East-temperate, West-tropical and West-temperate. The model estimates migration between region 1 and 2, 1 and 4, and 3 and 4 with tagging and some environmental data. We made a diagnostic evaluation of the model and the jitter analysis suggested that the model was not stable, therefore, we started analyzing the spatial structure of the model and the tagging data that the model use to estimate movement. The tagging data were mainly release in region 1, and very few information is collected from the rest of the regions (Table 2), this suggest that the tagging data can be very limited source of information to estimate the movement between the 4 regions within the model. Thus, consequently the model does not find a global minimum but a local minimum and therefore, it's very sensitive to the initial values (Figure 4). Then, we analyzed different spatial structures such as 2-area model and 3-area model but keeping the fleets structure based on the 4 regions defined in the 2018 reference model. The 2-area model considers East and West as two regions without any fish movement between them (similar to the estimates of the 2018 reference model) (Table 3) while the 3-area model aggregates the region 3 and 4 of the 2018 reference model, but with similar fish movement as the 2018 reference case. The jitter analysis suggests that the 2-area model is the most stable and therefore, we started the reference case with the spatial structure of two areas (Figure 6). Afterwards different modifications were done in order to improve the parameterization of the model (Table 7). So the 4-area model and the 2-area model give very similar results considering the same data as in 2018 reference model (Figure15).

Later the model was updated with catch, length composition and the new joint indices. However, the model was very sensitive to the new data with a very different virgin biomass to the 2018 reference model, where only length composition data until 2014 were considered. The model shows a big residual pattern in most of the fleets (Figure 18) and different fleets were removed in order to understand the source of the difference between both estimates. And although the model was very sensitive to the longline indices (not shown here), it was not the only fleets influencing the results and therefore in the reference model only data until 2014 were considered as the 2018 reference model.

Initial diagnostic performance for initial reference model and selected sensitivity runs

All of the sensitivity runs had positive definite Hessians and maximum gradient components less than 0.001 (Table 10). Most parameters (only estimates for reference case shown in Table 9) had relatively low standard errors (<1) except for gillnets and handlines in region 1a, and purse seiners LS. Plots of the parameter prior distribution and maximum likelihood estimates are included in each of the run folders and are more informative about parameter estimability. There were some highly correlated selectivity parameters in the reference case (Table 11) with a few notable exceptions being Q of the joint longline index in the region 1 of the model and the initial recruitment parameter. None of the sensitivity models had any bounded.

Table 10. Likelihood components of each sensitivity analysis and the reference case.

| Scenario | TOTAL | Survey | Length_com p | Tag_com p | Tag_negbi n | Recruitmen t | Parm_prior s | Parm_softbound s |
|---------------------------|---------|---------|-----------------|--------------|----------------|-----------------|-----------------|---------------------|
| RC | 3018.66 | -332.16 | 3379.2 | 0 | 0 | -58.3817 | 14.2466 | 0.00427246 |
| RC_2019 | 3526.26 | -309.51 | 3861.51 | 0 | 0 | -58.5116 | 15.1757 | 0.0042117 |
| RC4area_2018 | 8811.88 | -309.53 | 3397.33 | 4028.08 | 1694.5 | -48.4641 | 35.1444 | 0.00461399 |
| RC_2019_lbda01 | 4133.48 | -313.39 | 3868.72 | 421.601 | 179.825 | -58.4813 | 16.7954 | 0.00416009 |
| RC4area_2018_lamb da01 | 3626.34 | -313.59 | 3368.35 | 413.898 | 173.678 | -55.8466 | 23.5252 | 0.0042268 |
| RC_lbda01 | 4133.48 | -313.39 | 3868.72 | 421.601 | 179.825 | -58.4813 | 16.7954 | 0.00416009 |
| RC_h07 | 3017.12 | -333.03 | 3378.14 | 0 | 0 | -57.8469 | 14.3143 | 0.00426801 |
| RC_h09 | 3020.13 | -331.37 | 3380.04 | 0 | 0 | -58.6704 | 14.1916 | 0.00427571 |
| RC_sigmaR04 | 2982.47 | -326.4 | 3403.38 | 0 | 0 | -125.463 | 14.1089 | 0.00428587 |
| RC_sigmaR05 | 3001.92 | -329.72 | 3389.61 | 0 | 0 | -88.4713 | 14.1589 | 0.00427759 |
| RC_sigmaR065 | 3026.29 | -333.22 | 3374.87 | 0 | 0 | -45.1097 | 14.2999 | 0.0042707 |
| RC_Mp | 3005.02 | -332.18 | 3366.93 | 0 | 0 | -58.1752 | 12.0201 | 0.00467517 |
| RC_Mp_lbda01 | 3602.37 | -333.62 | 3362.54 | 422.11 | 175.146 | -56.0077 | 15.9897 | 0.00454192 |
| RC_Mhigh | 3060.85 | -330.22 | 3424.22 | 0 | 0 | -60.1063 | 8.99174 | 0.00508561 |
| RC_G | 3175.05 | -353.2 | 3535.08 | 0 | 0 | -40.8746 | 19.3278 | 0.00512885 |
| RC_G_Mp | 3074.33 | -340.96 | 3436.23 | 0 | 0 | -53.2397 | 12.4612 | 0.00482738 |
| RC_buoy | 3003.45 | -351.39 | 3380.32 | 0 | 0 | -54.2959 | 14.1419 | 0.00428326 |
| RC_buoy_PSFS_q1 | 3022.23 | -328.73 | 3378.91 | 0 | 0 | -57.2862 | 14.2217 | 0.0042058 |
| RC_8mix_lbda01 | 3259.1 | -333.15 | 3386.73 | 144.107 | 90.3836 | -57.5829 | 12.6527 | 0.00418886 |
| RC_8mix_lbda1 | 5334.57 | -333.22 | 3390.89 | 1415.89 | 879.341 | -52.6252 | 16.7492 | 0.0041654 |
| RC_lbda1 | 8904.92 | -331.69 | 3419.22 | 4154.8 | 1673.05 | -50.8628 | 22.4628 | 0.00426373 |

Table 11. Correlation between parameters.

| label.i | label.j | corr |
|---------------------------------|-------------------------------|-----------|
| LnQ_base_SURVEY1(26) | SR_LN(R0) | -0.913149 |
| SizeSpline_Val_2_FISHERY6(6) | SizeSpline_Val_1_FISHERY6(6) | 0.92064 |
| SizeSpline_Val_5_FISHERY6(6) | SizeSpline_Val_4_FISHERY6(6) | 0.726609 |
| SizeSpline_Val_5_FISHERY8(8) | SizeSpline_GradHi_FISHERY8(8) | 0.897619 |
| Age_DblN_ascend_se_FISHERY1(28) | Age_DblN_peak_FISHERY1(28) | 0.960995 |
| Age_DblN_ascend_se_FISHERY2(29) | Age_DblN_peak_FISHERY2(29) | 0.877216 |

| | | |
|--|---|-----------|
| Age_95%width_FISHERY3(3) | Age_inflection_FISHERY3(3) | 0.792925 |
| Age_DblN_ascend_se_FISHERY4(31) | Age_DblN_peak_FISHERY4(31) | 0.882682 |
| Age_DblN_ascend_se_FISHERY5(32) | Age_DblN_peak_FISHERY5(32) | 0.943432 |
| Age_DblN_ascend_se_FISHERY12(39) | Age_DblN_peak_FISHERY12(39) | 0.925004 |
| Age_DblN_descend_se_FISHERY12(39) | Age_DblN_peak_FISHERY12(39) | -0.869416 |
| Age_DblN_descend_se_FISHERY12(39) | Age_DblN_ascend_se_FISHERY12(39) | -0.753311 |
| Age_DblN_ascend_se_FISHERY1(28)_BLK1repl_213 | Age_DblN_peak_FISHERY1(28)_BLK1repl_213 | 0.972 |
| Age_DblN_ascend_se_FISHERY1(28)_BLK1repl_261 | Age_DblN_peak_FISHERY1(28)_BLK1repl_261 | 0.945977 |
| Age_DblN_ascend_se_FISHERY2(29)_BLK2repl_225 | Age_DblN_peak_FISHERY2(29)_BLK2repl_225 | 0.931484 |

We conducted full diagnostic evaluation at this point on the reference case. Initial diagnostic performance based on jitters indicates that the reference case is quite stable (Figure 6).

The model cannot converge if steepness, R_0 and σ_R are estimated, but with fixed steepness and fixed σ_R , the model can estimate R_0 and converge. The reference model assumes steepness of 0.8 and a σ_R of 0.6, similar to the 2018 reference case. However, the likelihood increases at higher steepness of 0.9. At higher steepness the model estimates lower R_0 (Figure 19), with a little increase on length and survey's likelihood (Table 10). The σ_R of the reference case is quite high = 0.6, but still the likelihood also increases at higher $\sigma_R=0.65$ (Table 10). The model estimates higher R_0 , and the likelihood increases because the fit to the recruitment is improved, although the likelihood on length and survey gets a bit lower (Figure 20).

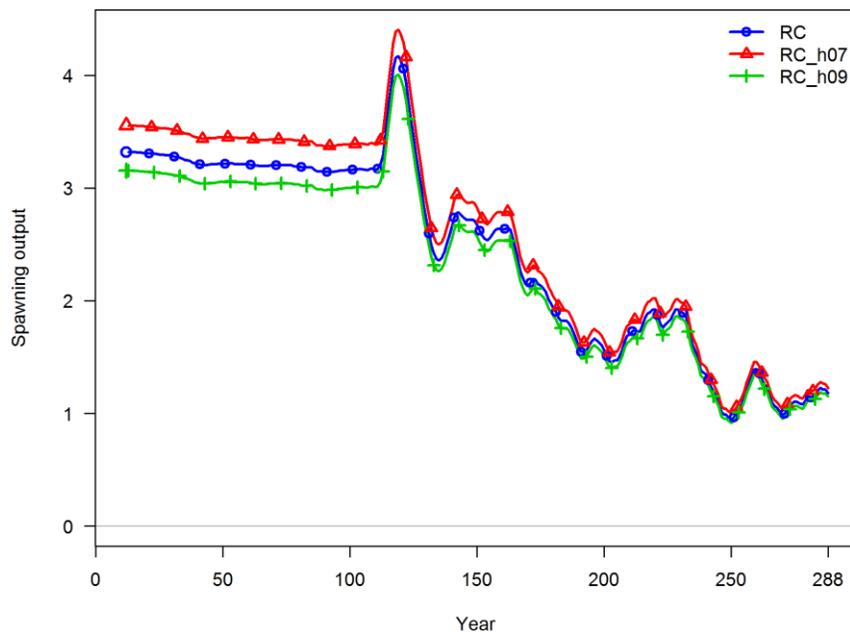


Figure 19. Spawning biomass trajectories from the reference case and the sensitivities related with the steepness of the model.

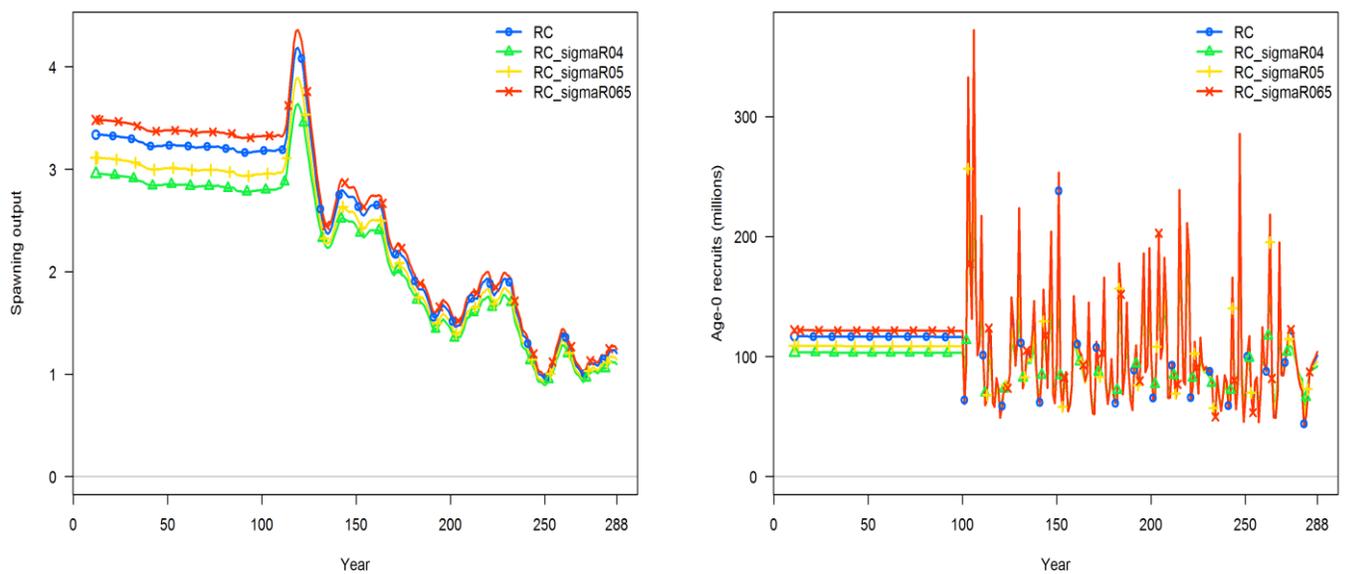


Figure 20. Spawning biomass trajectories from the reference case and the sensitivities related sigmaR.

The model does not have enough information to estimate M at each age, therefore, we analyzed the scenario estimating a unique M . In the reference case the model estimates an M of 0.67, and this would mean in a yearly time step a natural mortality of 0.17. The model suggests a much lower M , than for Atlantic YFT where the model estimates being around 0.4. The likelihood is lower than in the reference case with a bit lower likelihood in length composition (Table 10), and with a lower virgin biomass (Table 9, Figure 21). When tagging data are considered in the model (with $\lambda=0.1$) the model estimates a very similar $M=0.63$. We also analyzed the model with a different growth parameterization estimated by Dortel (2015), and let the model to estimate M . In this case the model estimate an M of 1.2, in a yearly time step it would mean 0.3, and this value is much closer to the value of Atlantic YFT. In this case, the fit to the survey gets worst but the fit to the length composition is improved (Table 10, Figure 22). However, the likelihood of the model is higher assuming the M of the reference case and the growth of Dortel (2015). This model gives the highest likelihood of all the models analyzed here (without considering tagging data), and also the lowest likelihood in the surveys and virgin biomass value (Table 9, Figure 22).

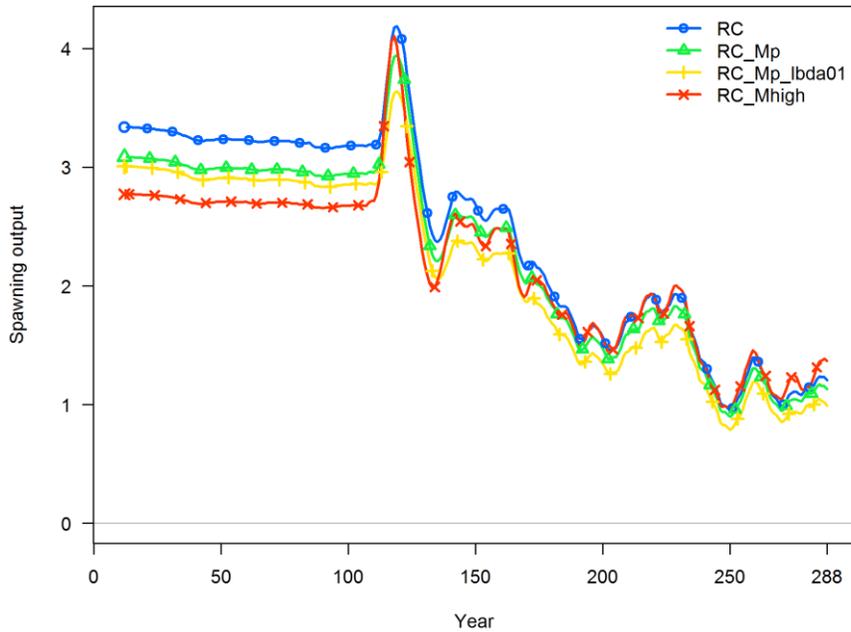


Figure 21. Spawning biomass trajectories from the reference case and the sensitivities related with natural mortality.

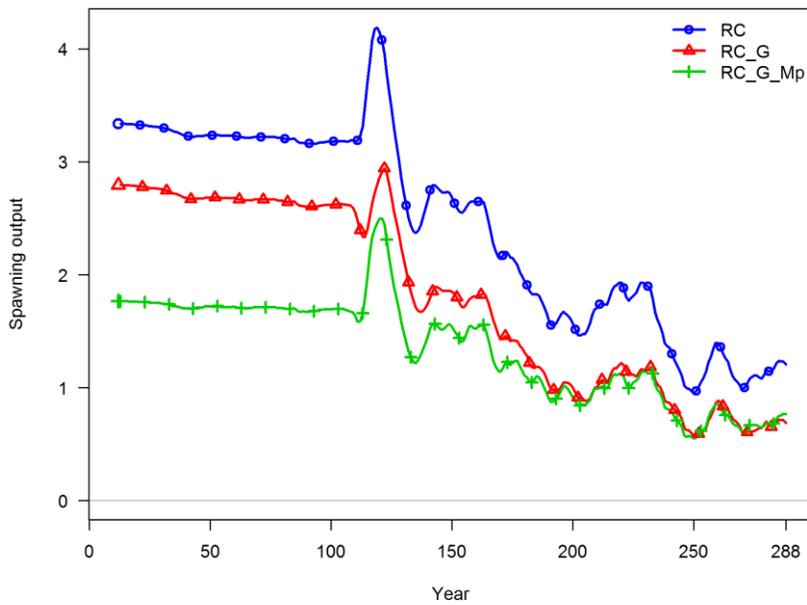


Figure 22. Spawning biomass trajectories from the reference case and the sensitivities related with growth.

Retrospective performance indicates some bias in the reference case and mainly the bias is on F value, where the mohns rho value is higher than 0.2 (Figure 23, Table 12). The length frequency data consider in the reference case are until 2014, the same as in the 2018 reference case, where also was observed a trend in the retrospective pattern.

Table 12. Mohn's rho values estimated with the reference case.

| | B/Bmsy | F/Fmsy | SSB | REC | SPRratio | F |
|------------------|-------------|------------|-------------|----------|------------|------------|
| RC2area_2019_272 | -0.02568354 | 0.17406121 | -0.08155841 | 0.018405 | 0.07360533 | 0.29472103 |

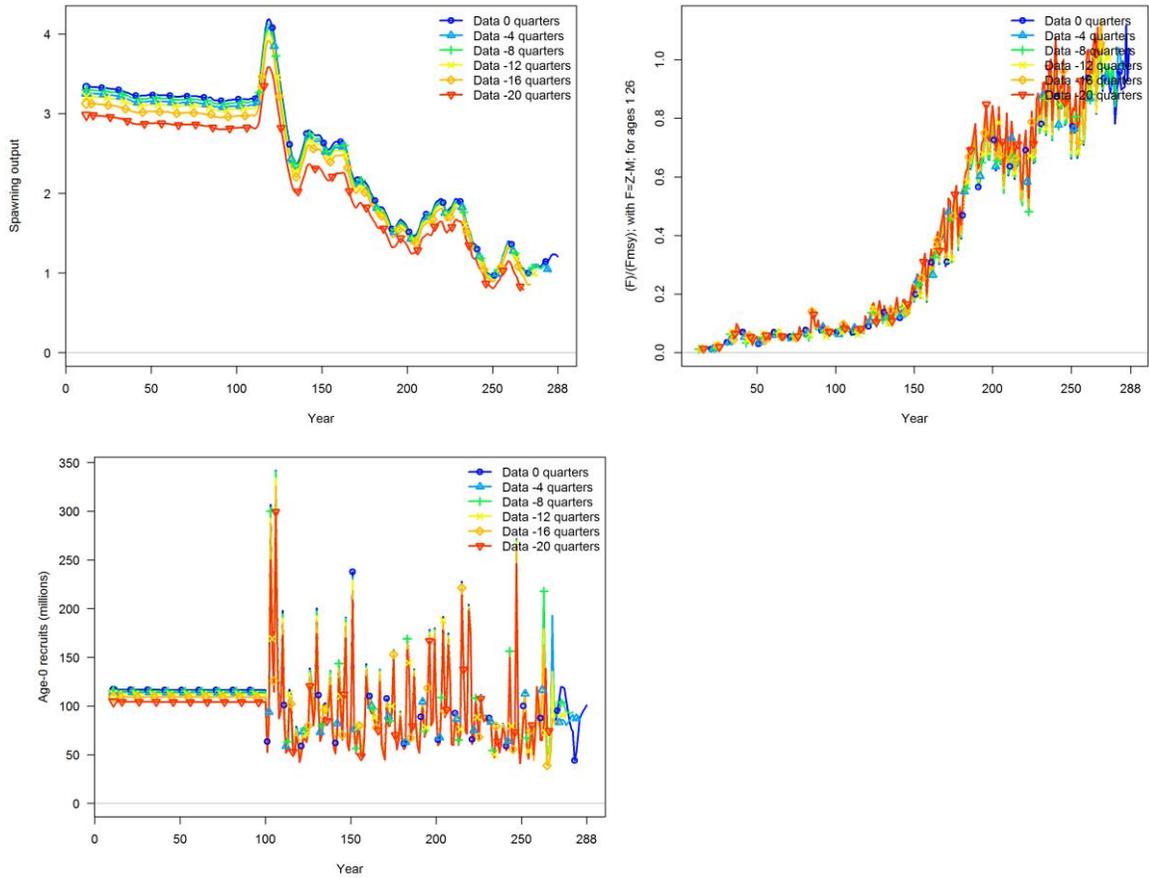


Figure 23. Retrospective pattern of the reference case .

Fit to the joint index show some residual patterns for the indices in both regions (Figures 24), but particularly at the start of the time series.

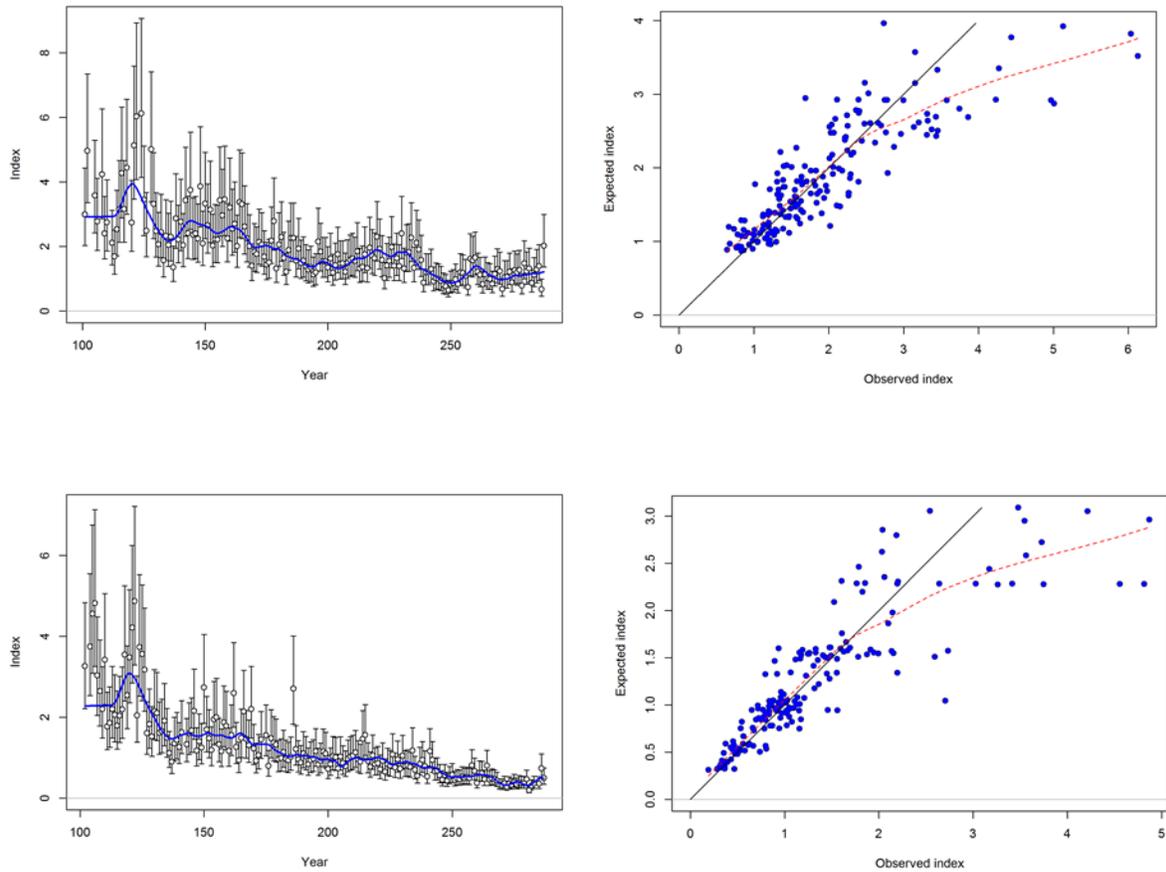


Figure 24. Residuals of the fits of joint index in the region 1 and region 2 of the ss-model.

The residual pattern of the length composition was shown above as a basis to only consider length composition until 2014 (Figure 17). The aggregated length composition by fleet showed three general patterns (Figures 25 and 26): clear selectivity on the smallest fish (gillnets, baiboat troll and other fleets)(Figure 27), bimodal selectivity (PS LS) (Figure 28) and clear selectivity on the largest (PS FS, longlines, fresh longlines and handlines) (Figure 29).

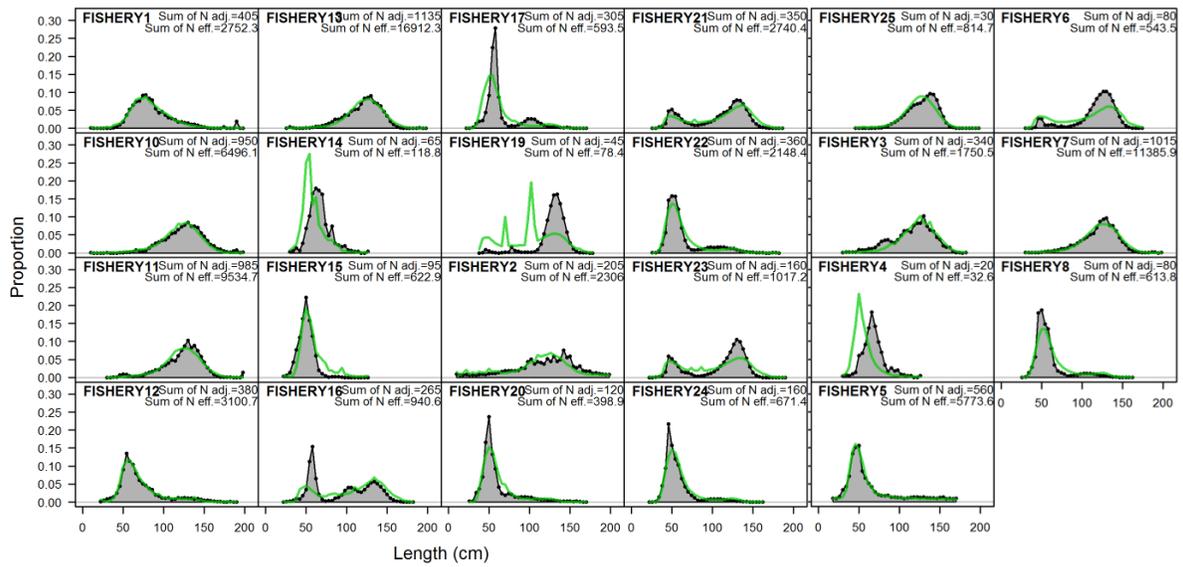


Figure 25. The length composition and the fit of the reference case by fleet.

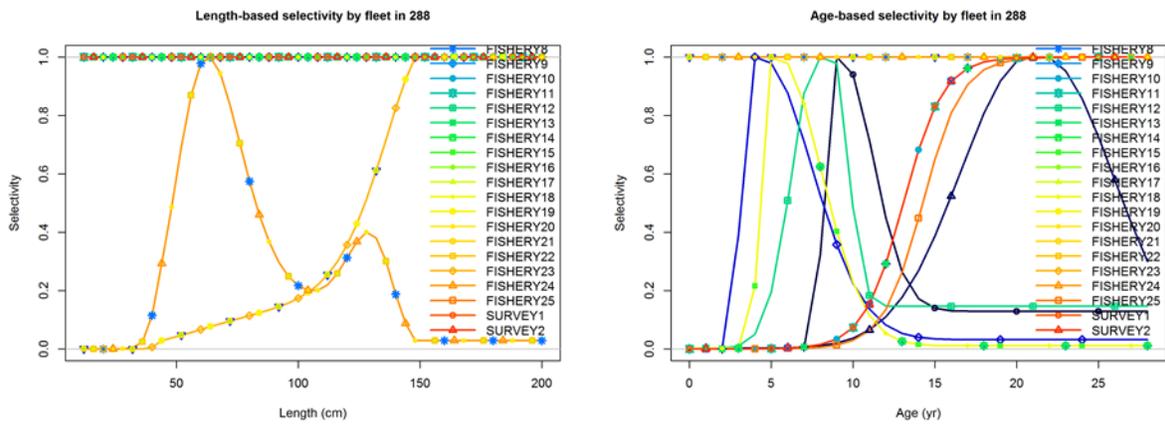


Figure 26. Estimated selectivity by fleet based on length and based on age.

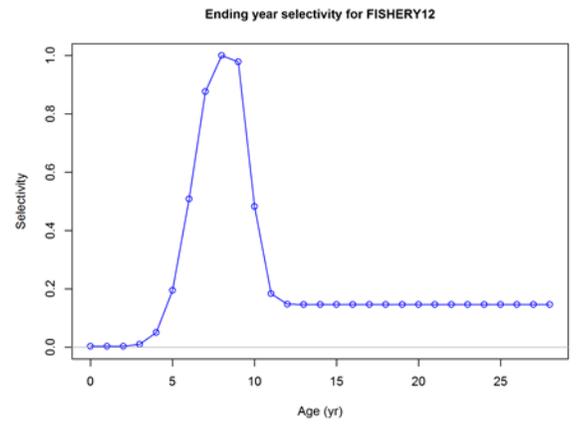
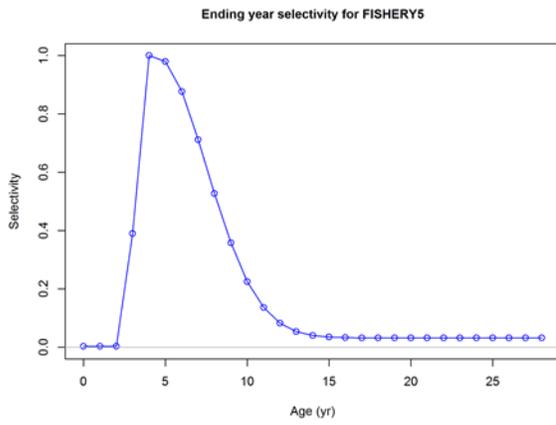
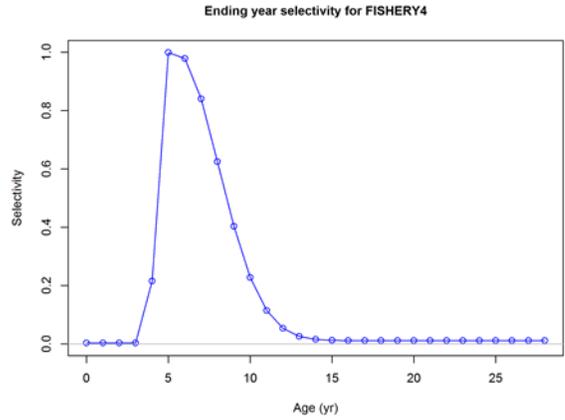
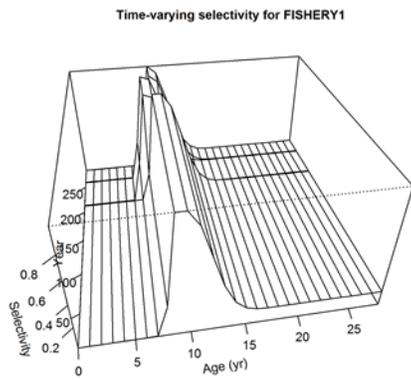


Figure 27. Estimated selectivity for gillnets in region 1a (Fishery1) and region 4 (Fishery12), baitboat (Fishery 5-region 1b) troll and other fleets in region 1 and 4 (Fishery 4).

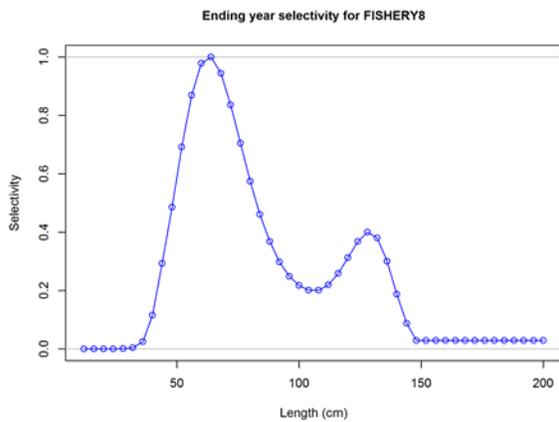


Figure 28. Estimated selectivity for purse seine LS fishing in region 1, 2 and 4 (Fishery6).

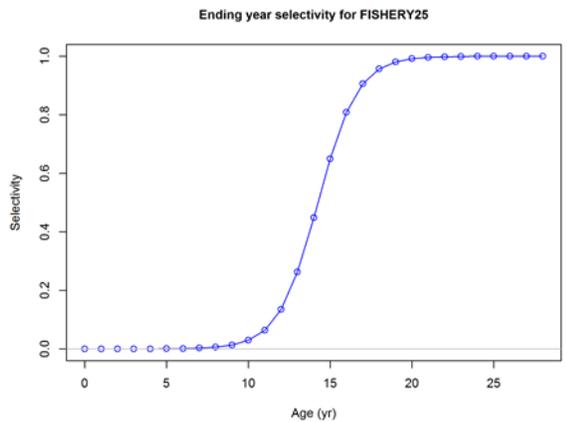
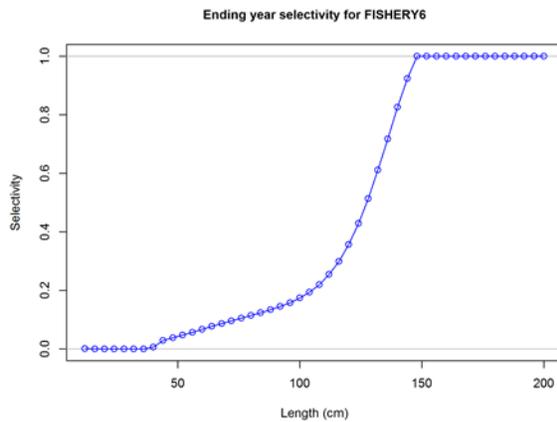
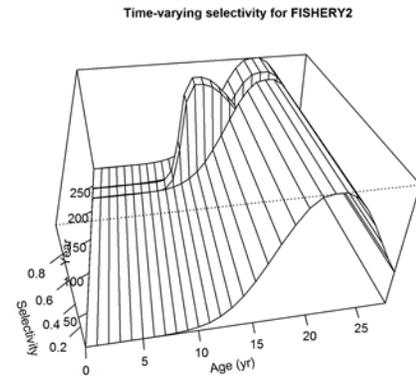
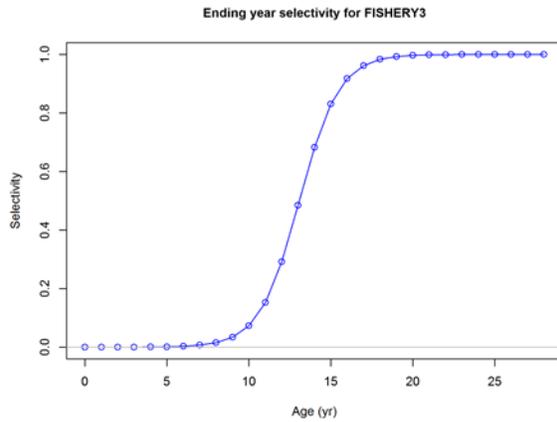


Figure 29. Estimated selectivity for longlines in region 1, 2, 3 and 4 (Fishery3) and region 4 (Fishery12), handline (Fishery 5-region 1b) troll and other fleets in region 1 and 4 (Fishery 4).

The estimated stock recruitment relationship shows little evidence of a relationship between SSB and recruits, however, the residuals show a negative trend at the highest values of SSB (Figure 30). The effect of using the low value of the maximum bias correction is that there is very little difference in absolute magnitude of the expected recruitment with or without the bias adjustment.

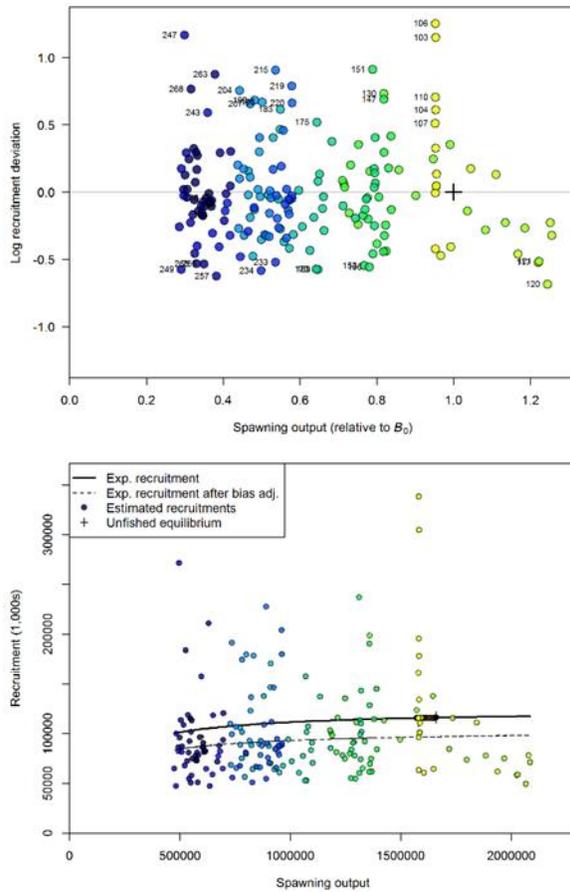


Figure 30. Estimated Beverton-Holt Spawner-recruit relationship and recruitment (age 0) estimates (with darker colors for more recent values) for Run 1. Dashed line is the bias-adjusted recruitment level during the period where recruitment deviations are estimated. The level of the adjustment, or reduction in recruitment level is determined by the bias correction factor that makes the mean recruitment level during the recruitment deviation estimation period equal to R_0 . Steepness is fixed at 0.8.

The recruitment deviates indicate recruitment varies seasonally (Figure 31). Recruitment deviates were low during 2004–2006 (229–236, model year), especially during 2005 (233–236 model year). Recruitment deviates declined between 2011 and 2017 (257–284 model year) and very close to 0 during 2018 (285–288 model year).

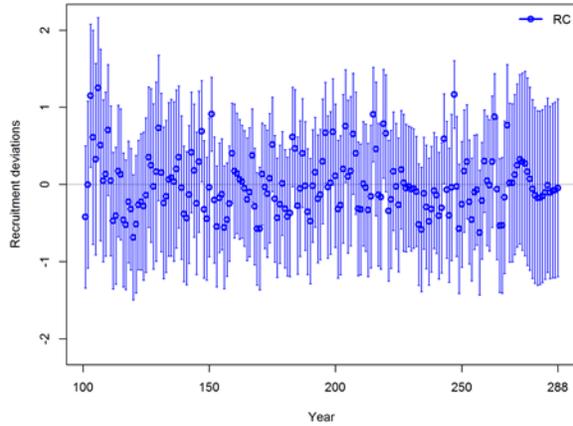


Figure 31. Recruitment deviates from the SRR and the associated 95% confidence interval.

Recruitment is parameterised to occur in region 1 and 2 with an estimated proportion of recruitment of 73% and 26 % in the respective regions. The proportion of total recruitment assigned to either region varies temporally during the estimation period (1977–2017) and overall the proportion of recruitment allocated to region 1 during the estimation period is higher than the base level (and vice versa for region 2) (Figure 32). Recruitment within both regions fluctuate around the equilibrium level with some very high values sporadically mainly in region 1. But the model estimates very clear low recruitment during 2004-2006 (229-240) related with the sharp decline in the CPUE index the late 2000s (Fu et al. 2018). Since 2014 (270 model year) the model estimates recruitment closer to the equilibrium level without any high deviates (Figure 33).

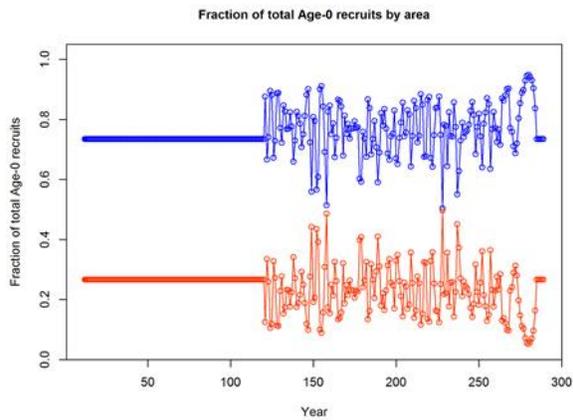


Figure 32. Proportion of the total quarterly recruitment assigned to region 1 (red) and region 4 (blue).

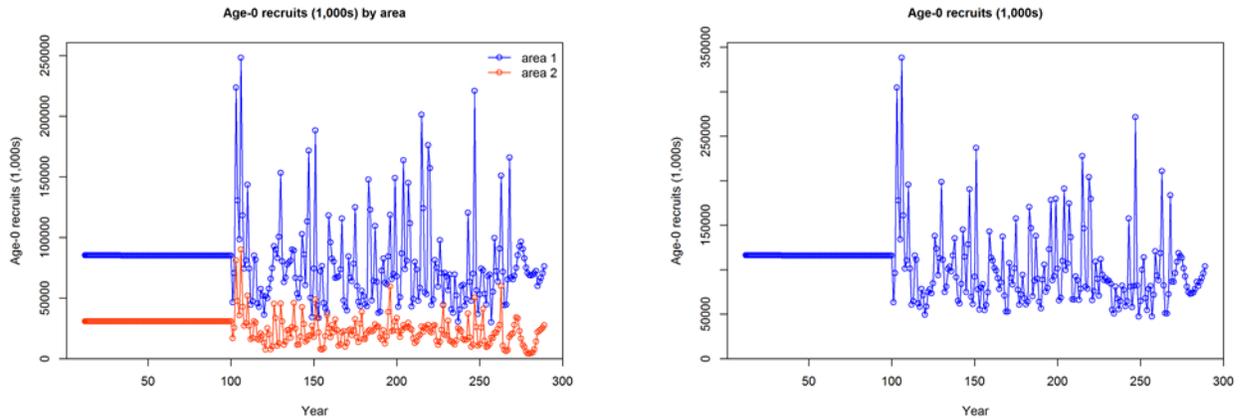


Figure 33. Estimated quarterly recruitment by region and for the entire IO with the reference case.

Total spawning biomass for the IO stock is estimated to have remained relatively high throughout the 1950s, 1960s and early 1970s (until around 120 model year) (Figure 34) corresponding with the relatively low levels of catch during the period and the assumption of equilibrium recruitment. Total spawning biomass declined rapidly during the late 1980s to mid-1990s, recovered slightly during the late 1990s (around 200 model year) and early 2000s (around 230) before declining to a low level in 2008–2009 (around 250 model year). Total spawning biomass recovered slightly during 2009–2011 (around 250-260 model year) and and since then it remains on those low levels. Current (2018) total spawning biomass is estimated to be close to the historically low level.

The confidence intervals are quite narrow associated with the time-series of total spawning biomass (Figure 34). The high level of precision is likely to be a function of the key assumptions of the model, especially constant catchability and selectivity associated with the LL CPUE indices and the fixed biological parameters.

Relative trends in spawning biomass are comparable for both model regions (Figure 34), although the overall magnitude of the decline in biomass is substantially higher in Region 2. The biomass in this region declined steadily throughout the 1990s and 2000s following the trend in the regional LL CPUE indices. For the most recent years, region 2 biomass is estimated to be at a very low level (Figure 34). The trend mainly follows the same pattern as the CPUE indices, and as sensitivity analysis, other two indices were also introduced in model region 1: acoustic biomass estimates from FAD buoys deployed prior to fishing (Santiago et al. 2019) and a purse seine free school index (Guéry et al. 2019) (Figure 35). The estimated pattern in spawning output considering the buoy indices, is very similar to only consider joint longline indices (Figure 36), with a bit lower virgin biomass, a little decrease in spawning output the last year not observed in the reference case and with a bit higher likelihood in recruitment (Table 10). However, the model is not able to fit the highest observed values by the index (Figure 36). But when the PS FS index is introduced in the model, the model shows some pattern in the residual of the lowest observed values and at the highest. The model does not estimate a high index during 2006-2008, when the longline indices decrease and it estimates very similar spawning biomass to the reference case (Figure 37).

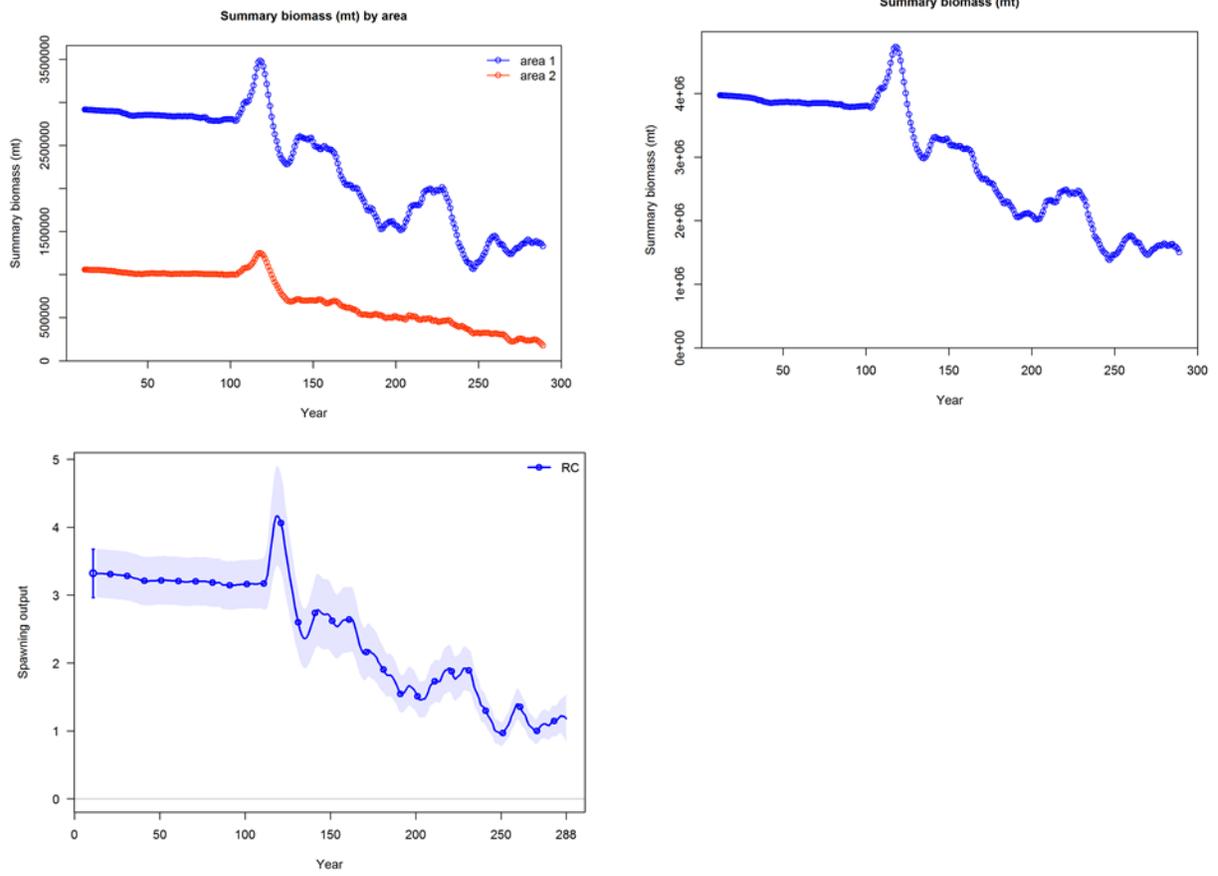


Figure 34. Total biomass (thousand mt) by region and spawning output for the IO with the reference case.

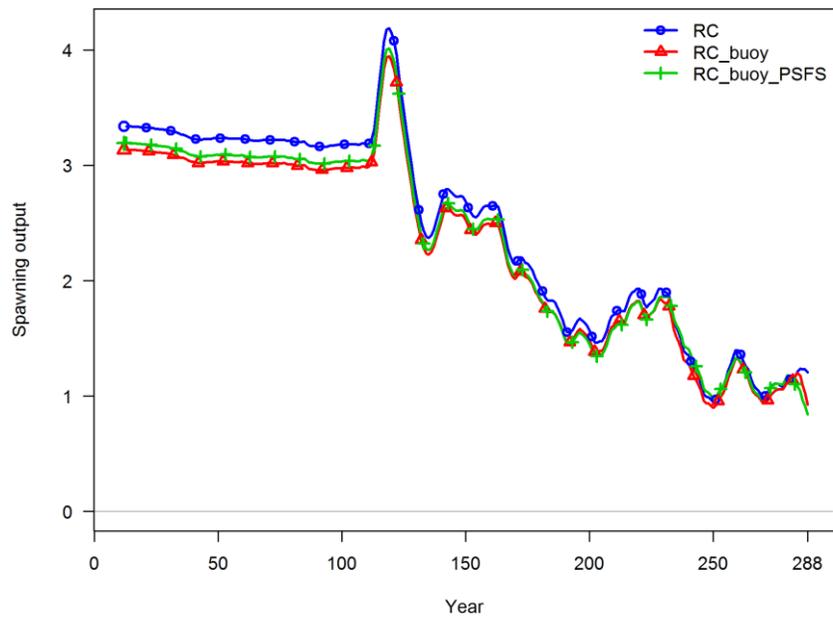


Figure 35. Spawning biomass trajectories from the reference case and the sensitivities related with the CPUE index.

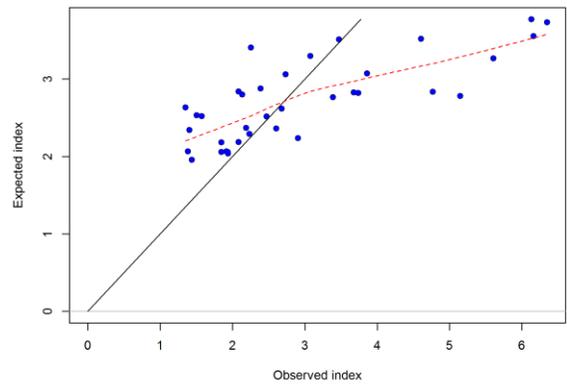
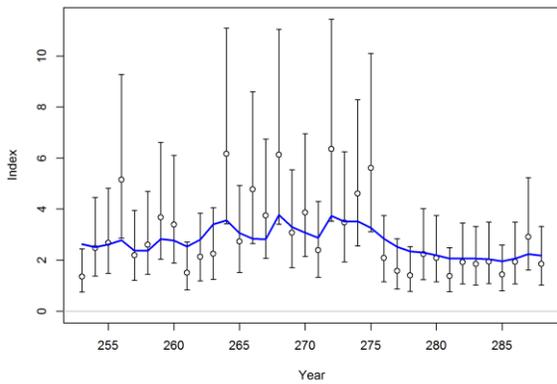
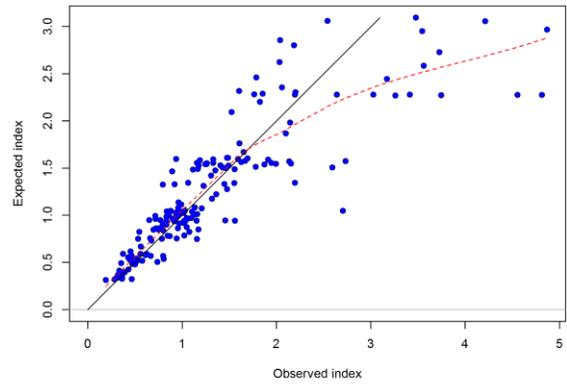
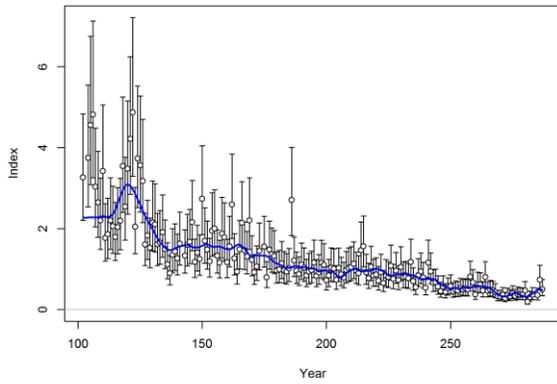
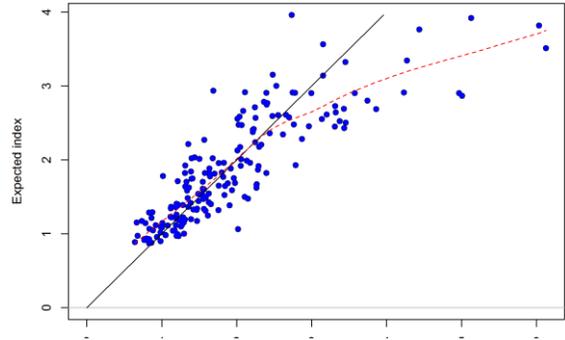
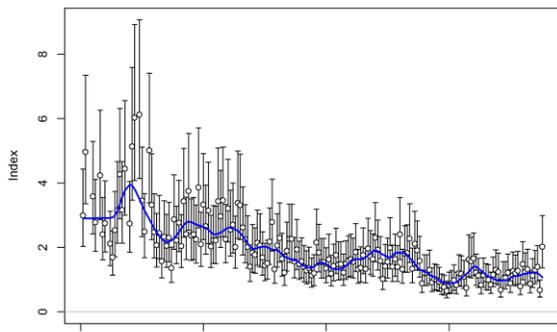


Figure 36. Residuals of the fits of longline (region 1 and 2) and buoy index in region 1 (model RC_buoy).

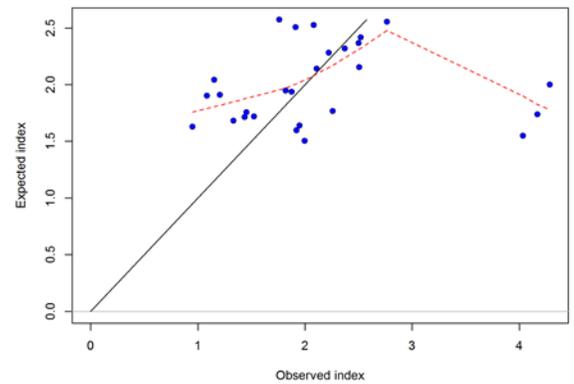
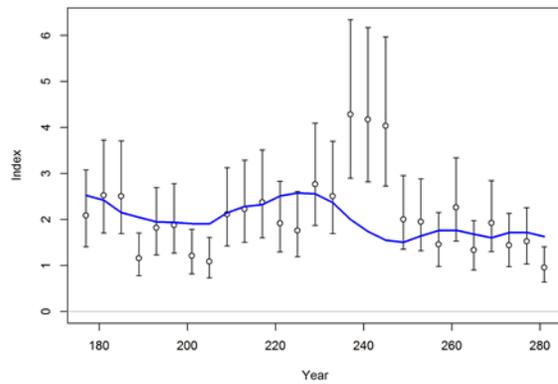
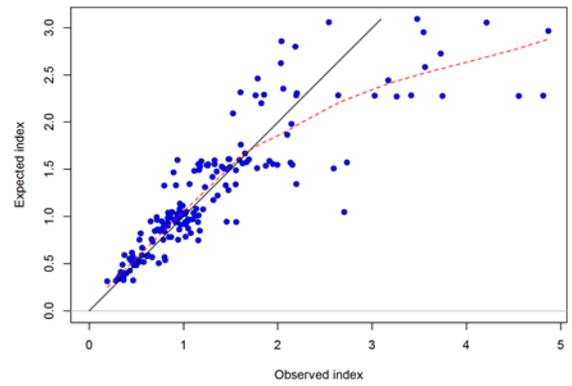
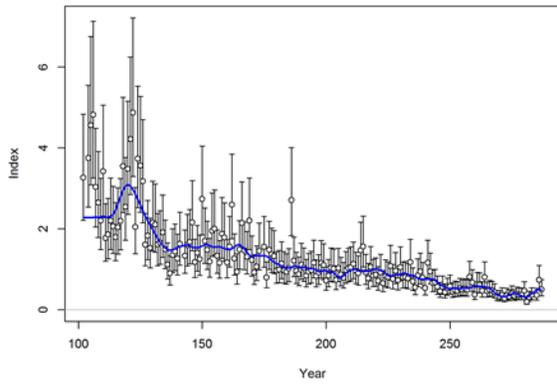
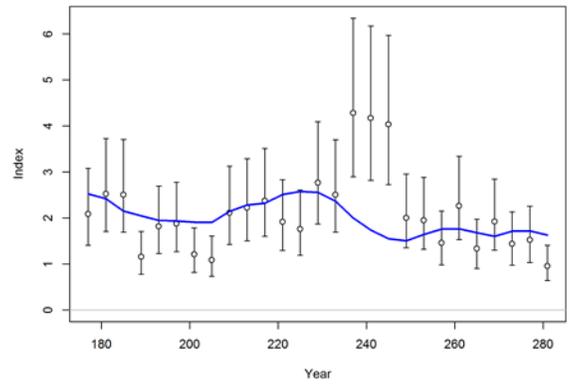
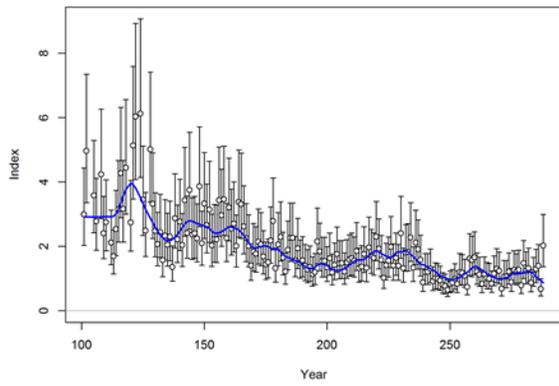


Figure 37. Residuals of the fits of longline (region 1 and 2), buoy index in region 1 and PS FS index in region 1 (model RC_buoy_PSFs).

Analysis of the available tagging data using the reference case

The observed recaptures of tagging data per quarter and year released, excluding the three-quarter mixing period and 7-quarters period are shown in Figure 38. In the case of assuming the 3-quarter mixing period 48% of the recapture data are used, while in the case of assuming 7-quarter of mixing period only 8% of the recoveries are used. The residuals of the fit to the tagging data by group shows that the model overestimates the tags from groups 1 to 20, from region 1. On the contrary, in the groups from 120 to 131, from region 2, the model underestimates the recoveries of those groups. In the rest of the groups, the residuals show a clear pattern on the first quarter after the mixing period, where the model underestimates the number of tags recovered. This could be a signal showing that after 3 quarters the tags are not completely mixed (Figure 39). However, when the model assumes 7-quarters as mixing period, the residuals are smoother in comparison to the 3 quarters assumption (Figure 39). In the case of the fleets, the purse seiners residual show a clear positive residual trend (Fleets 6, 8, 23 and 24), but when 7-quarters of mixing period is assumed, the high residuals of fleets 6 and 8 disappears because these fleets are only defined between 2006 and 2008 (Figure 40). The analysis of the different assumptions on tagging data, such as, mixing period or the weight on the likelihood of tagging data, shows that the results are very sensitive to the treatment of the tagging data (Figure 41). The general patterns are similar, but the virgin biomass changes assuming different weights on tagging data; it decreases when tagging data are considered and increases if tagging data are considered but downweighed to 0.1. The virgin biomass decreases also assuming a mixing period of 7-quarters, however, the impact of the mixing period is lower than the impact the weight of tagging data on the likelihood. The likelihood of the length composition increases when the tagging data are considered assuming 4 or 7-quarters of mixing period, with a higher increase with three-quarter mixing period (Table 10). These figures suggest that the tagged fish recovered by the purse seine fleets have not been fully mixed in the fishing area and therefore. Should these data be used in the assessment as a factor of the uncertainty grid, these results suggest that their influence should be downweighed and the mixing period should be extended.

Figure 38: The recaptured tagging data per quarter and year released, excluding the three-quarter mixing period (left) and excluding 7-quarter period (right), is shown in Figure 1. Bubble size shows the number of tag recaptured.

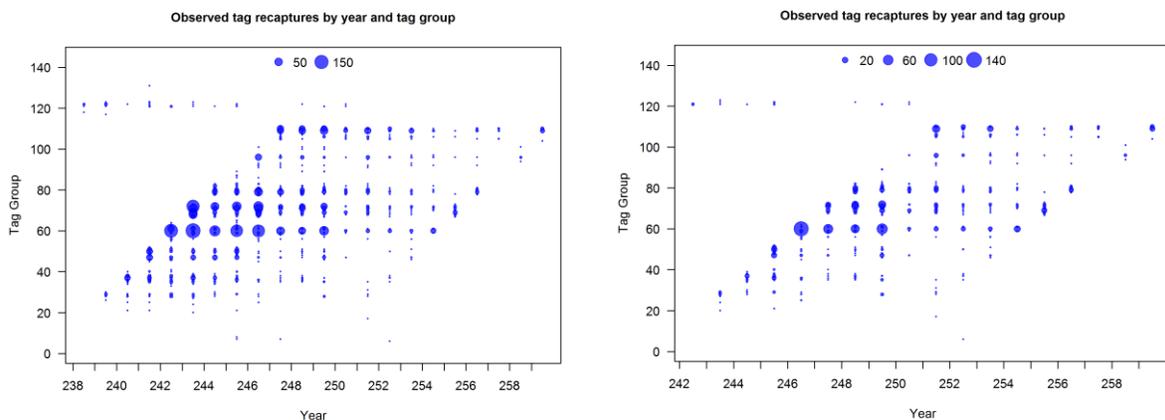


Figure 39: The residuals of the model fit of tagging recaptured data by tagging group assuming three-quarters of mixing period (top figure), and a downweighting in likelihood of lambda 0.1 (left) and no downweighting (right), and bottom figures assuming 7-quarters of mixing period.

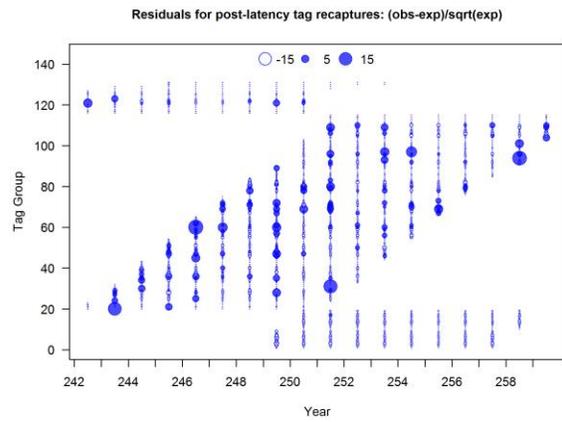
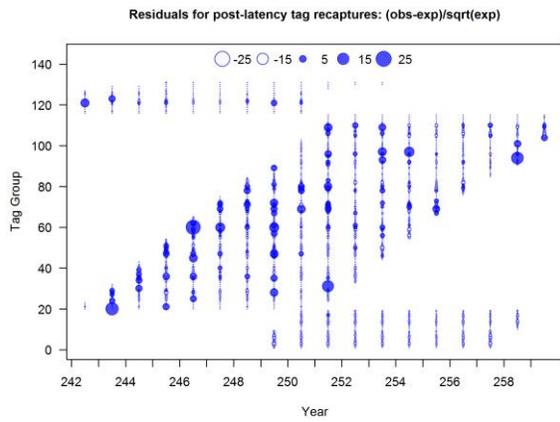
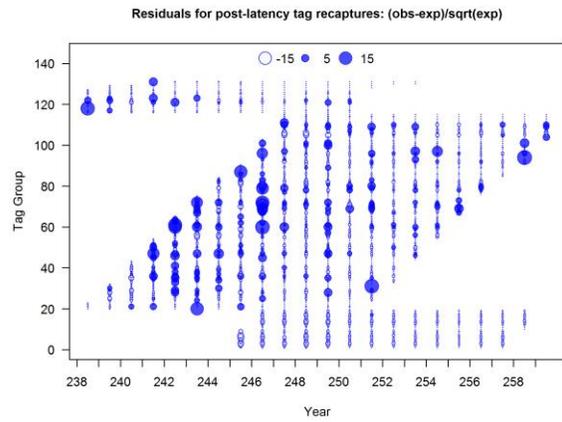
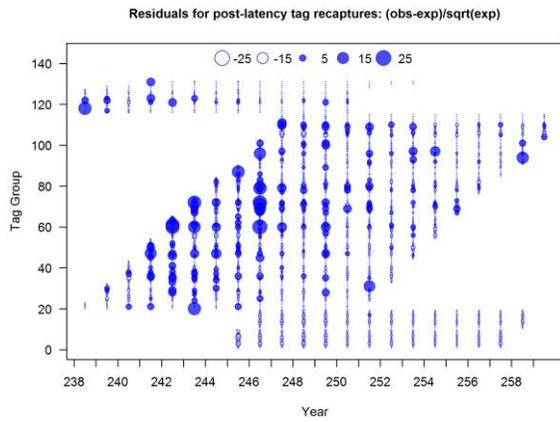


Figure 40: The residuals of the model fit of tagging recaptured data by fleet assuming three-quarters of mixing period (left figure) and assuming mixing period of 7-quarters and both downweighting by 0.1 the likelihood of the tagging data.

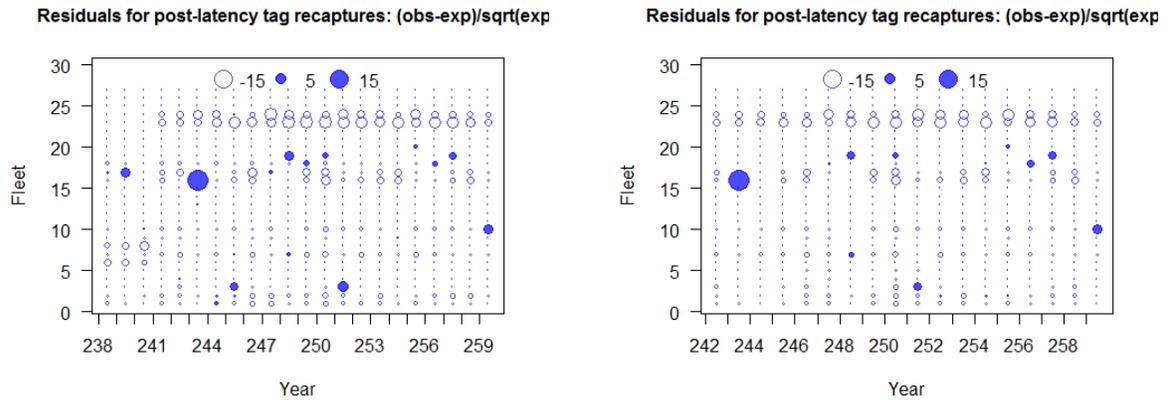
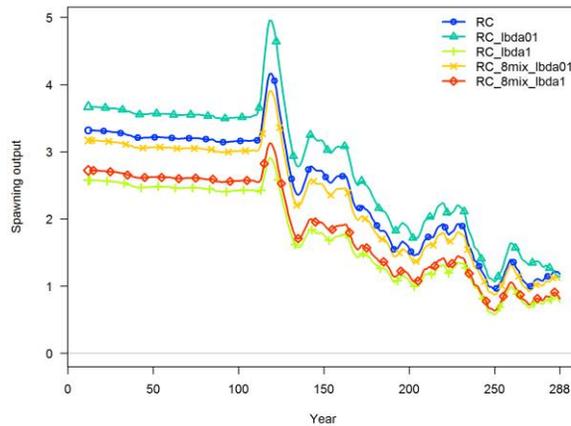


Figure 41: The comparison of the reference case on spawning output with sensitivity analysis on tagging data: assuming tagging data without downweighting and downweighting 0.1 and assuming 7-quarters of mixing period.



References

- Chassot, E. 2014. Are there some small yellowfin tunas caught in free-swimming schools? IOTC-2014-WPDCS-10-INF05.
- Fonteneau, A. 2005. An overview of yellowfin (*Thunnus albacares*) tuna stocks fisheries and stock status worldwide. IOTC-2005-WPTT-21.
- Fonteneau, A. 2008. A working proposal for a Yellowfin growth curve to be used during the 2008 yellowfin stock assessment. IOTC-2008-WPTT-4.
- Fu D., Langley A., Merino G. Urtizbera A. Preliminary Indian Ocean Yellowfin tuna stock assessment 1950-2017 (Stock Synthesis) (2018). IOTC-2018-WPTT20-33.
- Guéry L., Kaplan D., Marsac F., Floch L., Báez JC and Gaertner D. Accounting for fishing days without set, fishing concentration and piracy in the CPUE standardization of yellowfin tuna in free schools for the EU purse seine fleet operating in the Indian Ocean during 1991-2017 period. IOTC-2019-XXX
- Harley, S.J. 2011. Preliminary examination of steepness in tunas based on stock assessment results. WCPFC SC7 SA IP-8, Pohnpei, Federated States of Micronesia, 9-17 August 2011.
- Hoyle, S.D., Okamoto, H., Yeh, Y., Kim, Z., Lee, S.4 and Sharma, R. (2015) IOTC-CPUEWS-02 2015: Report of the Second IOTC CPUE Workshop on Longline Fisheries, April 30th-May 2nd, 2015. IOTC-2015-CPUEWS02-R[E]: 128pp.
- Hoyle, S.D., Satoh, K., Matsumoto, T. 2017. Selectivity changes and spatial size patterns of bigeye and yellowfin tuna in the early years of the Japanese longline fishery. IOTC-2017-WPTT19-34.
- Hoyle, S.D., Langley, A. 2018. Indian Ocean tropical tuna regional scaling factors that allow for seasonality and cell areas. IOTC-2018-WPM09-13.

Hoyle, S.D., Chassot, E., Fu, D., Kim, D.N., Lee, S.I., Matsumoto, T., Satoh, K., Wang, S.P., Yeh, Y.M., Kitakado, T. 2018b. Collaborative study of yellowfin tuna CPUE from multiple Indian Ocean longline fleets in 2018. IOTC-2018-WPM09-12.

Hoyle, S.D., Merino, G., Murua, H., Yeh, Y.M., Chang, S.T., Matsumoto, T., Kim, D.N., Lee, S.I., Herrera, M. and Fu, D. IOTC-CPUEWS-06 2019: Report of the Sixth IOTC CPUE Workshop on Longline Fisheries, April 28th–May 3rd, 2019. IOTC-2019-CPUEWS06-R[E]: 28 pp.

IOTC 2008b. Report of the 10th session of the IOTC Working Party on Tropical Tunas, Bangkok, Thailand, 23 to 31 October 2008. IOTC-2008-WPTT-R[E].

Langley, A. 2012. An investigation of the sensitivity of the Indian Ocean MFCL yellowfin tuna stock assessment to key model assumptions. IOTC-2012-WPTT-14-37.

Langley, A., Million, J. 2012. Determining an appropriate tag mixing period for the Indian Ocean yellowfin tuna stock assessment. IOTC-2012-WPTT-14-31.

Langley, A. 2015. Stock assessment of yellowfin tuna in the Indian Ocean using Stock Synthesis. IOTC-2012-WPTT-17-30.

Langley, A. 2015. Stock assessment of yellowfin tuna in the Indian Ocean using Stock Synthesis. IOTC-2012-WPTT-17-30.

Langley, A. 2016a. An update of the 2015 Indian Ocean Yellowfin Tuna stock assessment for 2016. IOTC-2016-WPTT18-27.

Langley, A. 2016b. Stock assessment of bigeye tuna in the Indian Ocean for 2016 — model development and evaluation. IOTC-2016-WPTT18-20.

Maunder, M.N., Aires-da-Silva, A. 2012. A review and evaluation of natural mortality for the assessment and management of yellowfin tuna in the eastern Pacific Ocean. External review of IATTC yellowfin tuna assessment. La Jolla, California. 15-19 October 2012. Document YFT-01-07.

Methot, R.D. 2000. Technical description of the Stock Synthesis assessment program. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-43, 46 p.

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

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

Santiago J., Uranga J., Quincoces I, Orue B., Grande M., Murua H., Merino G., Urtizberea A., Pascual P., Boyra G. A novel index of abundance of juvenile yellowfin tuna in the Indian Ocean derived from echosounder buoys. IOTC-2019-WPTT19-XX.

Suzuki, Z. 1993. A review of the biology and fisheries for yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. In Shomura, R.S.; Majkowski, J.; Langi S. (eds). *Interactions of Pacific tuna fisheries. Proceedings of the first FAO Expert Consultation on Interactions of Pacific Tuna Fisheries*. 3–11 December 1991. Noumea, New Caledonia. Volume 2: papers on biology and fisheries. FAO Fisheries Technical Paper. No. 336, Vol.2. Rome, FAO. 1993. 439p.

APPENDIX A: ANALYSIS OF TAG RECAPTURE DATA FROM THE RTTP-IO PROGRAM

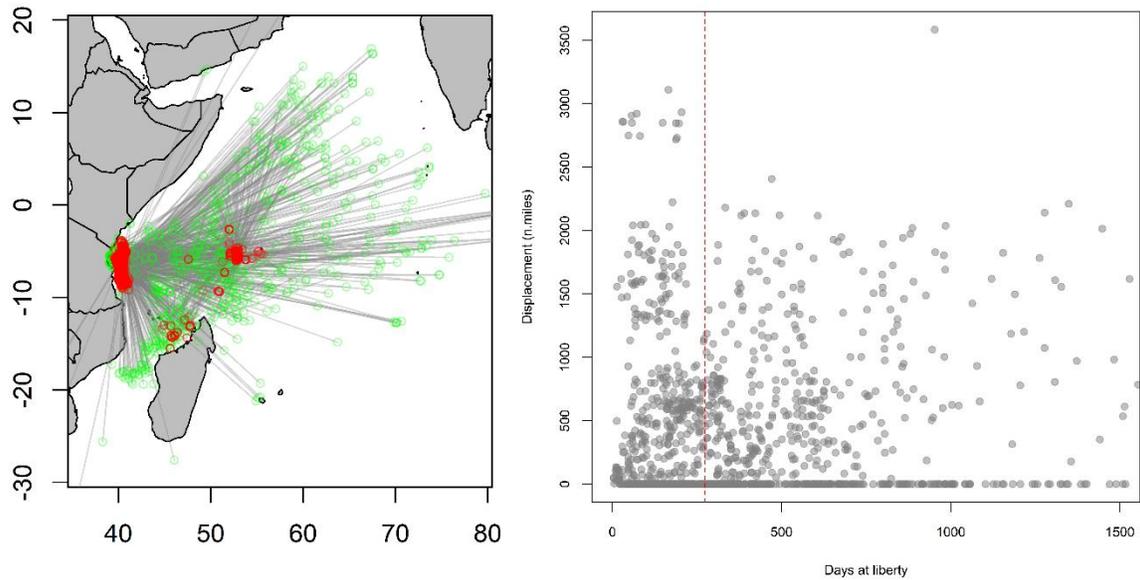


Figure A1. Net movement of tags between release and recapture (left) and displacement vs. days at liberty for a subset of tag recaptures from the RTTO-IO program. Only tag recaptures that have different (directional) bearing (and maximum net displacement for those of the same bearing) are included. Red circles indicate releases and green circles indicate recaptures.

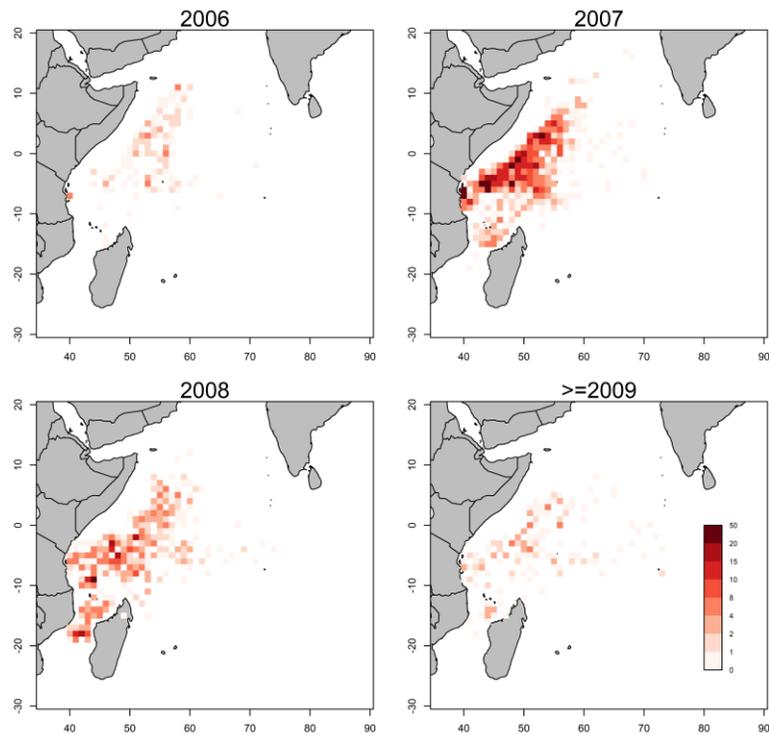


Figure A2: Spatial distribution (1 degree cell) of number of yellowfin RTTP tag recoveries of fish at liberty for at least 3 quarters, from the purse seine fisheries in the western tropical region.

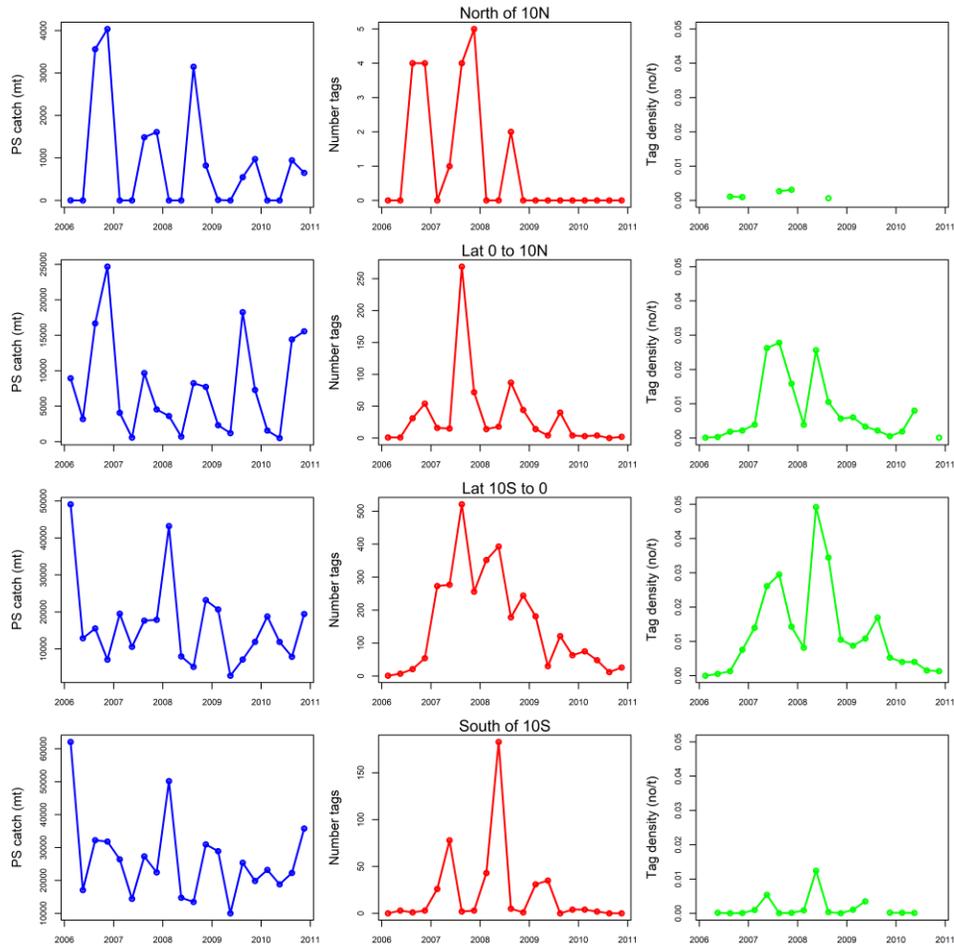


Figure A3: Quarterly yellowfin tuna catch (t) and number of tags recovered by the purse seine fishery in region 1 by latitudinal band. Only tags at liberty for at least 3 quarter mixing periods are included. The tag recovery density (tags/catch) is also presented for each latitudinal band.

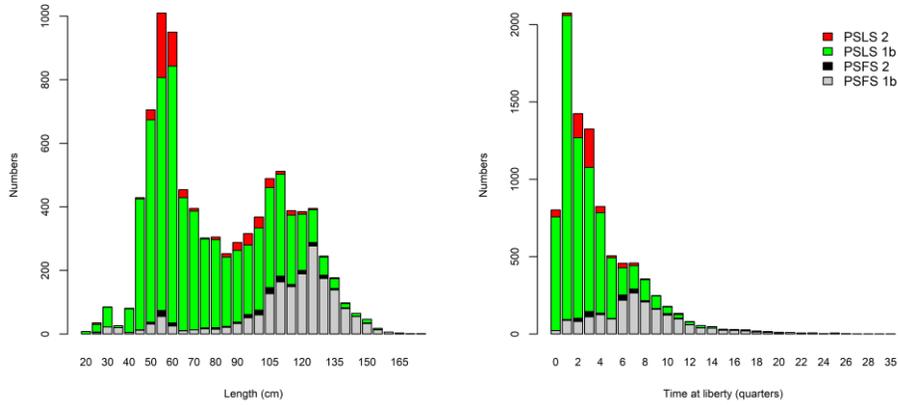


Figure A4: Distribution of yellowfin tag recoveries by length (left) and by time-at-liberty (right) for the purse seine free school and purse sein FAD schools in region 1b and 2. Purse seine tag recoveries have been corrected for reporting rate.