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

IOTC-2025-WPEB21(DP)-10

# A review of the 2021 blue shark assessment in the Indian Ocean

21st IOTC Working Party on Ecosystems and Bycatch
28 April – 3 May 2025
Virtual Meeting

Joel Rice[1]

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

**Executive summary**

This paper presents a review of the 2021 stock assessment of blue shark in the Indian Ocean using Stock Synthesis (version 3.30.16.02 http://nft.nefsc.noaa.gov/Download.html). The blue shark assessment model is an age structured (25 years), spatially aggregated (1 region) and two sex model. The catch, effort, and size composition of catch, are grouped into 8 fisheries covering the time period from 1950 through 2019. Six indices of abundance, all from longline fisheries were considered for this analysis. This assessment considered two alternative time series of total catch. The diagnostic case model is parameterized using indices of abundance from the Portugal (2000-2019), Reunion (2007-2019) and the Japanese late (1992-2019) series, along with estimates of catch generated via a generalized additive model. The estimated abundance trend is decreasing throughout the time frame of the model, and spawning stock abundance has decreased to approximately 1.21 times SSBMSY, (80% CI is 1.08-1.36). The fishing mortality has increased over the model time frame with F2019/FMSY= 0.81 (80% CI =0.66 to 0.96).

Blue sharks are most often caught as bycatch in the Indian Ocean tuna fisheries, though some directed mixed species (sharks and tunas/billfish) fisheries do exist. Commercial reporting of landings has been minimal, as has information regarding the targeting and fate of blue sharks encountered in the fisheries. Useful data on catch and effort is mostly limited to recent years, and time series of historical catches have been estimated based on reported and observed catch rates, as well as observed ratios of blue shark to target species.

***Base Case Model Selection***

The WPEB 17 meeting chose the Stock Synthesis model with all five CPUE series as the base case model for the provision of stock status advice. Growth, natural mortality and recruitment were parameterized as for the diagnostic model, explained earlier in the main section of the document. The selected base case model tracked the population similarly to the combinations of the CPUE series, and resulted in the highest SB2019/SBMSY ratio among the variations of CPUE considered

***Stock status.*** Progress since the last Indian Ocean blue shark assessment (2017) includes the updating of the CPUE series, development of estimates of natural mortality for the Indian Ocean based on growth curves from the Indian Ocean. New reports regarding blue shark age and growth based on recent studies have corroborated previous work, and lent greater certainty to the estimates used. As is common uncertainty in data inputs and model configuration were explored through sensitivity analysis, along with expanded model diagnostics. Two stock assessment models were applied to the blue shark data in 2021, the modeling framework Just Another Bayesian Biomass Assessment (JABBA) and an integrated age-structured model (SS3) (Fig. EX 1).

All models produced similar results suggesting the stock is currently not overfished nor subject to overfishing, but with the trajectories showing consistent trends towards the overfished and subject to overfishing quadrant of the Kobe plot (Fig 1). A base case model was selected based on the best Indian Ocean biological data, consistency of CPUE standardized relative abundance series, model fits and spatial extent of the data (Fig. 1). The only major change in biological parameters since the previous stock assessment is the natural mortality, which was calculated based growth curves specific to the Indian Ocean. The major axes of uncertainties identified in the current model are catches and CPUE indices of abundance. Model results were explored with respect to their sensitivity to the CPUE series. Due to the lack of reporting to species, non reporting and potential under reporting the nominal catch was considered un-reliable. Therefore the total catch was estimated, similar to the previous assessment a ratio based approach as well as a generalized additive model was used. The ratio based estimates had very high interannual variability which did not track the other target and associated species (i.e. swordfish), therefore the GAM estimated catches were selected as the primary estimates of total blue shark catch. Under the alternative CPUE groupings investigated the stock status was somewhat less positive with lower $B_{2019}/B_{MSY}$ and slightly higher $F_{2019}/F_{MSY}$. The ecological risk assessment (ERA) conducted for the Indian Ocean by the WPEB and SC in 2018(Murua et al., 2018) consisted of a semi-quantitative risk assessment analysis to evaluate the resilience of shark species to the impact of a given fishery by combining the biological productivity of the species and its susceptibility to each fishing gear type. Blue sharks received a medium vulnerability ranking (No. 6) in the ERA rank for longline gear because it was estimated as the most productive shark species, but was also characterized by the third highest susceptibility to longline gear. Blue shark was estimated as not being susceptible thus not vulnerable to purse seine gear. The current IUCN threat status of 'Near Threatened' applies to blue sharks globally. Information available on this species has been improving in recent years. Blue sharks are commonly taken by a range of fisheries in the Indian Ocean and in some areas they are fished in their nursery grounds. Because of their life history characteristics – they live until at least 25 years, mature at 4–6 years, and have 25–50 pups every year – they are considered to be the most productive of the pelagic sharks. On the weight-of-evidence available in 2021, the stock status is determined to be not overfished and not subject to overfishing.

***Outlook.*** Increasing effort (and catch) could result in declines in biomass. The Kobe II Strategy Matrix (Table 1) provides the probability of exceeding reference levels in the short (3 years) and long term (10 years) given a range of percentage changes in catch.

***Management advice.*** Even though the blue shark in 2021 was assessed to be not overfished nor subject to overfishing, maintaining current catches is likely to result in decreasing biomass and the stock becoming overfished and subject to overfishing in the near future (Table 1). If the catches are reduced at least 10%, the probability of maintaining stock biomass above MSY reference levels ($B>B_{MSY}$) over the next 10 years will be increased (Table 1). The stock should be closely monitored. While mechanisms exist for encouraging CPCs to comply with their recording and reporting requirements (Resolution 16/06), these need to be further implemented by the Commission, so as to better inform scientific advice in the future.

The following key points should also be noted:

- **Maximum Sustainable Yield (MSY)**: estimate for the Indian Ocean stock is approximately 36,000 t. Recent catches have ranged from 42,000-60,000.
- **Reference points**: The Commission has not adopted reference points or harvest control rules for any shark species.
- **Main fishing gear** (2013–2021): Coastal longline; longline targeting swordfish; longline (deep-freezing).
- **Main fleets** (2013–2021): Indonesia; EU,Spain; Taiwan, China; Japan; EU,Portugal.

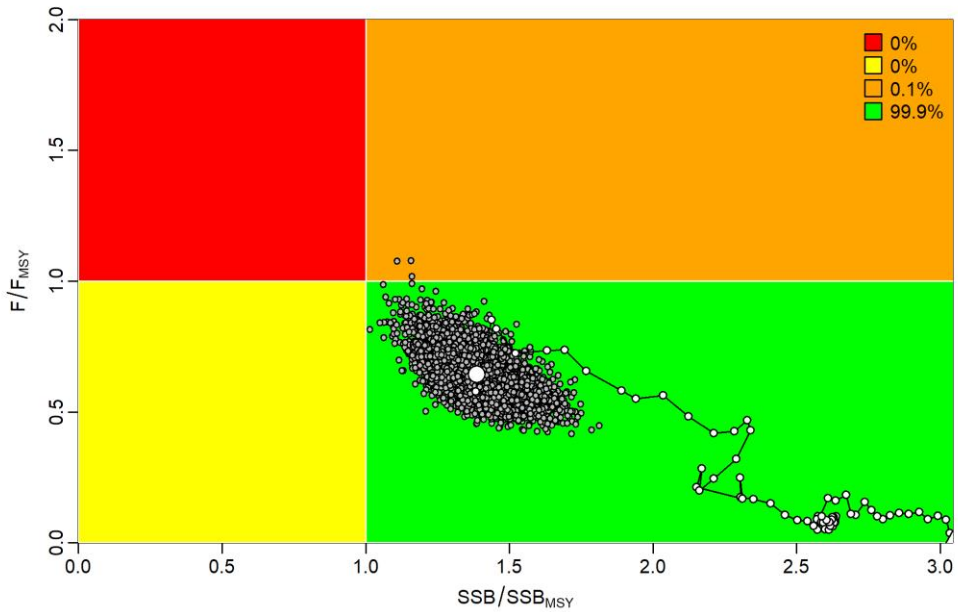



**Fig. EX 1.** Blue shark: Aggregated Indian Ocean stock assessment Kobe plot for the 2021 stock status estimate based on the base case model with trajectory and uncertainty calculated via sampling under the MVLN framework for the terminal year.

**TABLE 1.** Blue shark: Aggregated Indian Ocean assessment Kobe II Strategy Matrix. Probability (percentage) of violating the MSY-based reference points for nine constant catch projections using the base case model (catch level from 2019* (42,240), ± 10%, ± 20%, ± 30% and ± 40%) projected for 3 and 10 years.

| **Reference point and projection time frame** | **Alternative catch projections (relative to the catch level* from 2019) and probability (%) of violating MSY-based reference points** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Catch Relative to 2019 | 60% | 70% | 80% | 90% | 100% | 110% | 120% | 130% | 140% |
| Catch (t) | 25,944 | 30,267 | 34,592 | 38,916 | 43,240 | 47,564 | 51,888 | 56,212 | 60,535 |
| **$B_{2022} < B_{MSY}$** | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| **$F_{2022} > F_{MSY}$** | 0% | 0% | 0% | 0% | 0% | 1% | 5% | 16% | 36% |
| | | | | | | | | | |
| **$B_{2029} < B_{MSY}$** | 0% | 0% | 0% | 0% | 0% | 2% | 9% | 25% | 48% |
| **$F_{2029} > F_{MSY}$** | 0% | 0% | 0% | 0% | 1% | 13% | 44% | 75% | 90% |

*:catch level and respective % changes refer to the estimated catch series used in the final base case model (IOTC-2021-WPEB17(AS)-15)

# 1 Introduction

Blue shark (*Prionace glauca*) are a large pelagic species, broadly distributed throughout the Indian Ocean to a southern limit of ~50° S (Figure 1). Indian Ocean blue shark have been incidentally caught by the Japanese longline fleet since the early 1950s. The population was not heavily exploited before targeted fisheries (or bycatch rates increased) in the early 1990s. At this time the Taiwanese long line vessels began taking large numbers, initially in the SW region, followed by the other areas (Figure 1). The European longline fleet (predominantly Spanish vessels) started a targeted fishery in the 1990s, while only small numbers are reported in the driftnet fisheries, and purse seine catches are very rare.

# 2 Methods

### Data

There are many different fleets catching blue shark in the Indian Ocean, with vastly different gear types and levels of data quality (Martin et. al. 2015). This model uses the same fishing fleet structure previously used (Rice 2017, Rice and Sharma 2015), 8 fleets representing a wide variety of gears, some of which have been aggregated (e.g. F1 Miscellaneous). The number of CPUE series recommended by the WPEB data prep meeting is 6, all of which are based on longline fisheries. There is enough uncertainty about the selectivity assumptions with respect to time, and the low numbers of size composition data, that the size composition data are not expected to be very informative about year-class strength. Hence, in the assessment presented here, the length-composition data are down weighted so as to inform the selectivity but not alter the model fit to the abundance trend.

### Total catch

Catch estimates by year and fishery are shown in Figure 2. In the previous assessment (Rice 2017), estimates of total catch were produced based on generalized additive models (GAM) and the ratio of blue shark (BSH) to total target catch. While the total catch data are estimates, they are derived in large part from the industrial fleets in the Indian Ocean and are thought to be more reasonable for blue shark than for the other shark species.

The major concern identified with respect to the catch time series are that catch-and-effort for BSH are highly incomplete. Reliable data are thought to be available for a limited number of years (i.e., from the late-1990s onwards) and for a very limited number of fisheries. In the previous assessments an alternative catch series was used based on trade based estimates using the proportion of tuna caught (Clarke, 2011). This series extends from 1981-2011, and was previously extended (both earlier and later) using a ratio based approach. This method used the average ratio of the nominal to trade based estimates from the years previous to 2011 to estimate the values for the years prior to 1981 and post 2011. Because of the uncertainty in the reported nominal catches introduced by using the average ratio, this method was not repeated for this analysis.

## 2.1 Relative abundance indices

The standardized CPUE series in 2021 were similar to those from those previously submitted to the WPEB. Newly estimated CPUE series by Japan, Taiwan, Portugal, Spain, South Africa and EU France (Reunion) were used in this analysis (Figure 2). All of these are based on bycatch in the longline fisheries. Excerpts from the working papers are presented here for an overview of the CPUE series. For further information consult the working papers.

**S2 Japanese Late Series JPN (IOTC-2020-WPEB16-20)**

This paper presented revised standardized catch rates for blue shark from Japanese observer data in the Indian Ocean from 1992 to 2019, including the following abstract provided by the authors:

> *"We updated the Japanese observer data until 2019 and standardized nominal catch-per-unit-effort (CPUE) of blue shark caught by Japanese tuna longline fisheries in the Indian Ocean from 1992 to 2019. We used generalized linear model (GLM) with negative binomial error distribution to standardize the nominal CPUEs. The most parsimonious model was selected by Akaike Information Criterion (AIC) as the best model for the estimation of annual*

*CPUEs. The goodness-of-fits were diagnosed by residual plots. The 95% confidence intervals were estimated from the bootstrapping method. The annual CPUEs had a similar trend to those shown in the previous analysis except in 2000.The annual CPUE increased in 1990s and reached to the peak in 2000, and then gradually decreased with a large fluctuation until 2013. Since 2014, the annual CPUE showed an increasing trend. We suggest that the estimated annual CPUE should be utilized as one of the candidates of primary abundance indices in the next stock assessment of blue shark in the Indian Ocean scheduled in 2021 because the Japanese observer data covers a wide range of the main distribution area (temperate water) of blue shark in the Indian Ocean and a longer time period compared to the other fleets' CPUE data".*

**S3 Portuguese Longline (IOTC-2021-WPEB17(DP)-10)**

This paper presented catches and standardized CPUE of blue shark in the Indian Ocean from the Portuguese longline fleet from 2000 to 2019, including the following abstract provided by the authors:

*"The Portuguese pelagic longline fishery in the Indian Ocean started in the late 1990's, targeting mainly swordfish in the southwest region. This working document analyses catch, effort and standardized CPUE trends for blue shark captured by this fishery. Nominal annual CPUEs were calculated in biomass (kg/1000 hooks), and were standardized with Generalized Linear Mixed Models (GLMMs) using year, quarter, season and targeting as fixed effects, and vessel as random effects. The standardized CPUE trends shows a general decrease in the initial years between 2000 and 2005, followed by a more stable period with some oscillations until 2019. These results present an updated annual index of abundance for the blue shark captured by the Portuguese pelagic longline fleet in the Indian Ocean that can now be consider3d for utilization in the 2021 IOTC blue shark stock assessment.."*

**S4 Spanish Longline (IOTC-2021-WPEB17(DP)-09)**

This paper presented standardized catch rates for blue shark from the Spanish surface longline fleet from 2001 to 2019, including the following abstract provided by the authors:

*“ This paper provides an update of standardized catch rates in weight of blue shark using a Generalized Linear Model (GLM) from a total of 2,301 trips carried out by the Spanish surface longline fleet targeting swordfish in the Indian Ocean during the 2001-2019 period. The criteria used to define explanatory variables were similar to those used in previous papers. The main factors considered in the analysis were year, quarter, area, ratio, gear and the interaction quarter*area. The results indicate that the ratio factor (an indicator of target criteria of the skippers) defined as the ratio between the two most prevalent species caught -swordfish and blue shark- was the most important factor which explained the CPUE variability. The GLM results explained 80% of CPUE variability in weight. The index showed a stable trend over time.”*

**S5 Taiwanese Longline (IOTC–2021-WPEB17(DP)-07)**

This paper provided an updated and revised standardized catch rate of blue sharks caught by the Taiwanese longline fishery in the Indian Ocean, including the following abstract provided by the authors:

*"The catches and efforts of the blue shark in the Indian Ocean were estimated based on the observers’ records (2004-2019) of Taiwanese tuna longline fisheries. To cope with the large percentage of zero shark catch, the catch per unit effort (CPUE) of blue shark, as the number of fish caught per 1,000 hooks, was standardized using a two-step delta-lognormal model (DLN) that treats the proportion of positive sets and the CPUE of positive catches separately. The standardized CPUE showed a stable increasing trend for blue sharks from 2008 to 2014 (the peak), although decreased in 2015, it increased again in 2016. Overall, the standardized CPUE series of the blue shark caught by Taiwanese longline fishery showed a stable trend. The stable trend suggested that blue shark stocks in the Indian Ocean seems at the level of optimum utilization.”*

**S7 EU (France) Réunion Longline (IOTC–2021-WPEB17(DP)-08**)

This paper presented standardized CPUE of blue shark from the French swordfish longline fishery in the southwest Indian Ocean from 2007 to 2019, including the following abstract provided by the authors:

*" The blue shark Prionace glauca is the main bycatch species of the French swordfish-targeting longline fishery operating in the south-west Indian Ocean. Using observer and self-reported data collected aboard commercial longliners between 2007 and 2020, we propose a standardized CPUE series for blue shark for this fishery estimated with a lognormal generalized linear mixed model (GLMM) to be used for stock assessment. We propose to use the standardized CPUE for the period comprised between 2011 and 2020 where the monitoring effort has been consequent in comparison with previous years. Throughout 2011-2020, the standardized CPUE for the blue shark shows a significant decreasing trend.*

**S8 South Africa (IOTC-2021-WPEB17(DP)-11)**

This paper presented standardized CPUE of blue shark from the joint-venture Japanese flagged vessels fishing in the South African EEZ and includes the following abstract provided by the authors:

"*The blue shark Prionace glauca is caught as bycatch in the large pelagic longline fishery in South Africa. The fleet includes a domestic component with varying but increasing degree of observer coverage, and a foreign-flagged component of Japanese vessels that operate under joint venture agreements with South African Right Holders. Japanese flagged vessels have been operating under a mandatory 100% observer coverage since 2007. The catch and effort data include consistent records of bycatch species in numbers caught per set. We investigated blue shark abundance by standardising the Catch per Unit Effort (CPUE) in numbers from Observer data for the time series 2007 to 2019. To do this, we applied a Generalised Additive Mixed Model (GAMM) with a Poisson error distribution. Explanatory variables of the final model included year, month, grid (lat, long) with the number of blue shark caught in a set offset by the number of hooks set, so as to maintain a count distribution. Vessel was included as a random effect. Despite a period of relatively low catch rates (2009-2012) followed by a period of relatively high catch rates (2015-2017), the results indicate that blue shark CPUE in the south-*

*western IOTC area has been stable overall. Our dataset is unique in that the joint-venture Japanese flagged vessels have required 100% observer coverage since 2007. Given the increasing stricter catch regulation on shark species, our observer dataset may be the most appropriate dataset to accurately represent trends in abundance of blue sharks in the south-western IOTC region."*

### 2.2 Size composition data

As with the previous analysis sex based length-composition data collected by observers and from logsheets for the main fleets (Japan, Taiwan and Portugal) were used (Coelho et al 2017) along with additional length composition data submitted to the IOTC. In all, between 10 and 20 years of length composition data from the LL fleets were organized and used in the analysis. Some size and sex composition data of catch were available, but in many cases the data were in aggregated form covering several years, or size sampling was incomplete across fisheries. Many of the time series suffered from low sample sizes and inconsistencies across years. For this reason and because of the evidence that there was a conflict between the CPUE and the size data (see results below) lower weight was given to the size data in the model. This allowed the model to estimate selectivity, but did not allow the size data to dominate the estimates of abundance in the model. We assumed an annual effective sample size calculated as the overall (male and female) sample size divided by 40. The annual sample size was then weighted (once) by the Francis (2011 and 2014) likelihood weighting method.

### 2.3 Software

The analysis was undertaken with Stock synthesis SS V3.30.16.02, 64 bit version (Methot 2000, 2009, executable available from http://nft.nefsc.noaa.gov/SS3.html), running on MS Windows™ 10). Typical function minimization of the fully disaggregated model on a 3.0 GHz personal computer required about 10 minutes. Additional simplifications and aggregations could probably reduce the minimization time further, without significant loss to the stock status inferences.

### 2.4 Model Assumptions

The most important model assumptions are described in the following sections. Standard population dynamics and statistical terms are described verbally, while equations can be found in Methot (2000, 2009). Attachment 1 is the template specification file for all of the models, and includes additional information on secondary elements of model formulation which may be omitted in the description below. All of the specification files are archived with the IOTC Secretariat. Table 2 lists the assumptions for the sensitivity runs.

### 2.5 Time Period

The model was iterated from 1950-2019 using an annual time-step, however, further analysis of seasonal processes is encouraged.

### 2.6 Biological inputs and assumptions

Blue sharks have an Indian Ocean wide distribution, and genetic evidence of distinct population structure within other oceans (e.g. Pacific) has not been found (Taguchi and Yokawa 2013), and hence was assumed to be homogenous here as well. Conventional tagging studies need to be examined in the Indian Ocean, but currently limited data exist, though some tagging effort in the Pacific shows limited movement to the western Australian EEZ. In addition to assumptions regarding stock structure, the other critical information on the biology of blue shark necessary for the stock synthesis assessment relates to sex-specific growth, natural mortality, maturity and fecundity.

### 2.7 Growth

The standard assumptions made concerning age and growth in the SS model are (i) the lengths-at-age are assumed to be normally distributed for each age-class; (ii) the mean lengths-at-age are assumed to follow a von Bertalanffy growth curve. For any specific model, it is necessary to assume the number of significant age-classes in the exploited population, with the last age-class being defined as a “plus group”, i.e. all fish of the designated age and older. For the results presented here, 25 yearly age-classes have been assumed, as age 25 approximates to the age at the theoretical maximum length of an average fish.

No attempt was made to estimate growth within the model due to the uninformative nature of the size data to track cohorts through time. The previous assessment considered the growth curves from Hsu et al. (2011) as well as specific formulations based on data from the Indian Ocean. This assessment uses new sex specific growth curves based on data from the Indian Ocean (Andrade et al 2017). A CV of 0.22 was used to model variation in length-at-age. All lengths reported from the assessment relate to fork length (FL).

### 2.8 Natural mortality

Sets of age and sex-specific natural mortality ogives were considered in the assessment based on the Peterson and Wroblewski (1984) method (Rice 2021) (Table2).

### 2.9 Maturity and fecundity

For the purpose of computing the spawning biomass, we assume a logistic maturity schedule based on length with the age-at-50% maturity for females equal to 145cm (Nakano and Seki 2003). There is no information which indicates that sex ratio differs from parity throughout the lifecycle of blue shark. Fecundity was fixed to an average of 25 pups per annual gestation period.

### 2.10 Population and fishery dynamics

The model partitions the population into 30 yearly age-classes in one region (Figure 1). The last age-class comprises a “plus group” in which mortality and other characteristics are assumed to be constant. The population is “monitored” in the model at yearly time steps, extending through a time window of 1950-2019. The main population dynamics processes are as follows: In this model “recruitment” is the appearance of age-class 1 fish (i.e. fish averaging approximately 50 cm in the population). The results presented in this report were derived using one recruitment episode per year, which is assumed to occur at the start of each year. Annual recruitment deviates from the recruitment relationship were estimated, but constrained reflecting the limited scope for compensation given estimates of fecundity. Deviations from the SRR were estimated in two parts (i) the early recruitment deviates for the 5 years prior to the model period which has the bulk of the length composition information (1966 -1970) and (ii) the main recruitment deviates that covered the model period (1971 - 2019).

There is no information which indicates that sex ratio differs from parity throughout the lifecycle of blue shark. In this assessment the term spawning biomass (SB) is a relative measure of spawning potential (the mature female population) and is a dimensionless term. It is not comparable to total biomass.

### 2.11 Initial population state

In the previous model it was assumed that the blue shark population was at an unfished state of equilibrium at the start of the model (1950) with the beginning of longline fishing occurring in the following years (at least from the 1950s onwards).

The population age structure and overall size in the first year is determined as a function of the estimate of the first years recruitment ($R_1$) offset from virgin recruitment ($R_0$), the initial 'equilibrium' fishing mortality discussed above, and the initial recruitment deviations. As the size data were found to be uninformative about initial depletion and recruitment variation only a small number (five) of initial recruitment deviates were estimated.

### 2.12 Selectivity Curves

Selectivity is fishery-specific and was assumed to be time-invariant. A double-half normal functional form was assumed for all selectivity curves except the miscellaneous fishery which was set to a logistic. An offset on the peak and scale was estimated for sex-specific differences in selectivity that were evident in the data. The selectivity function location and scale were estimated for fleets 3, 4, 6, 7 and 8 and the ascending and descending functions were fixed to a best fit when estimated independently. Only the location parameter was estimated for fleet 5 as the model failed to converge if the scale was also estimated.

### 2.13 Parameter estimation and uncertainty

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

### 2.14 Profile Likelihood

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

### 2.15 Hierarchical cluster analysis

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

The HCA used methods conducted in R using FLR (http://www.flr-project.org/). and the *diags* package. FLR provides a set of common methods for reading these data into R, plotting and summarizing them to assess the consistency in the CPUE trends. The CPUE time series along with a lowess smoother fitted to CPUE each year using a general additive model (GAM) to compare trends for the CPUEs. Hierarchical cluster analysis identified two groupings of time-series neither of which matched exactly the pervious groupings. The first group was characterized by time-series which were lightly positively correlated with each other, EU-Reunion and EU Spain,  and which had some highly negative correlations with TWN and South Africa. The second group was characterized by time-series which were less correlated with each other or were slightly negatively correlated with the CPUE series in other (positively correlated) group. This group was mad up of Taiwan, South Africa, Japan and EU Portugal. In fact EU Portugal was positively correlated with all of the other CPUE series except Taiwan. Because CPUEs with conflicting information were identified, it may be reasonable to assume that the indices reflect alternative hypotheses about states of nature and to run separate scenarios for each group. For this analysis it is recommended that we utilize the previous bas case grouping that was defined by expert opinion as well as the groupings defined through the HCA.

### 2.16 Selection of a diagnostic case and initial grid.

During the data prep meeting (April 2021) the WPEB noted that there are conflicting trends among some CPUE series and that the inclusion of conflicting data would result in a mis-specified model. This assessment follows from the 2017 assessment that used the EU Portugal, EU France (Reunion) and Japanese CPUE series, the choice of which was based in part on the hierarchal cluster analysis and in part on expert opinion to give more spread to the data. With respect to the estimated catch history the 17 WPEB data prep meeting noted that the available nominal catch data currently held in the IOTC database is likely a gross underestimate of the true catch. Given that approximately one third of the total reported sharks in the IOTC database are non species specific reports (i.e. reported as “sharks”) it is reasonable to assume that some of these reports represent blue shark given that blue shark are the most commonly caught pelagic shark. Therefore the diagnostic case assumes that the GAM estimated catches are the appropriate catch estimates.

An updated hierarchal cluster analysis showed that the most highly correlated CPUE series were EU Spain and EU France (La Reunion fleet - REU); these two series showed similar declining trends in recent years. The other grouping was made up of Taiwan China, South Africa, EU PRT and Japan. Taiwan was slightly negatively correlated with this group. Of note the EU Portugal series was positively correlated with all the other series except the Taiwan.

Sensitivity trials were run using the other CPUE time series and combinations of CPUE. During the 2017 assessment meeting the WPEB noted that the early and late Japanese CPUE series would likely have been affected by the changes in market demand for fins and blue shark meat over time. Sensitivities to the base case CPUE series groupings were run for those groups identified in Table 3. Groupings of CPUE series will be chosen by the WPEB which will seek to use the results of the HCA as well as expert opinion to extend the spatial extent and the temporal coverage of the CPUE series groupings. Initial groupings of CPUE series used in this report are A) Previous base case PRT, REU, JPN, B) EU ESP, EU REU C) EUESP, EU PRT, EU REU, South Africa, D) TWN, REU, JPN and D)all. In addition to the CPUE series groupings additional parametrizations of steepness (values of 0.75 and 0.8) and sigma r (an alternative value of 0.4) were considered.

### 2.17 Benchmark and Reference Point Methods

Benchmarks included estimates of absolute population levels and fishing mortality for the terminal year, 2019 ($F_{2019}$, $SSB_{2019}$, $B_{2019}$). These values are reported against reference points relative to MSY levels, and depletion estimates (relative to virgin levels).

### 2.18 Other Model Considerations

As explained above the length composition annual sample sizes were re-weighted by the Francis (2011) likelihood weighting method. The minimum average CV associated with the indices of abundance length likelihoods were re-weighting based on the Francis (2014) method. The life history and biology in the model are treated as constants, these parameters, along with the catch inputs influence the plausible range of population dynamics in the model.

### 2.19 Diagnostics and additional model runs

Additional model diagnostics which were carried out includes expanded analysis on the residual and hierarchal cluster analysis, runs tests and joint residual plots, likelihood profiles, and hindcasting cross-validation.

Additional model runs using an application of the generalized Bayesian State-Space Surplus Production Model framework JABBA (Just Another Bayesian Biomass Assessment) were carried out. JABBA has been previously applied and tested in assessments of Indian Ocean swordfish, South Atlantic blue shark, North Pacific blue shark, Mediterranean albacore tuna, North Atlantic shortfin mako shark, and South Atlantic swordfish.

## 3 Results

In this section we focus on the results from the diagnostic case model and the key results and diagnostics for this model. We then comment on any important differences in both outputs and model diagnostics for the sensitivity analyses, and present all results. The assessment model was implemented in Stock Synthesis version 3.30.16.02 (SS3 Methot 2013). A newer version of the model is available (version 3.30.17) but due to time constraints and the overall similarity of the model versions for the features implemented in this assessment. Stock Synthesis 3(v3.30.16.02) was implemented here as a length-based age-structured stock assessment model (Methot and Wetzel 2013; e.g., Wetzel and Punt 2011a, 2011b). Stock synthesis utilizes an integrated modeling approach (Maunder and Punt 2013) to take advantage of the many data sources available for the Indian Ocean stock of blue shark (*Prionace glauca*). An advantage of the integrated modeling approach is that the development of statistical methods that combine several sources of information into a single analysis allows for consistency in assumptions and permits the uncertainty associated with each data source to be propagated to final model outputs (Maunder and Punt 2013).

### 3.1 Diagnostic case model

The diagnostic case model choice is described in section 2.16. The choice of model parameters and data inputs reflected the input of the WPEB 17 data prep meeting and the available updated data for biology and life history. This will be further refined by input at the WPEB assessment meeting.

***Model Fits to Abundance Indices***

The model was able to fit the general trends of the indices of abundance (Figure 10). Although the CPUE series S2 and S3 had periodic increases in the CPUE that the model was unable to fit As a result, the model fitted the central tendency of each series, which for S2 the Japanese series was a slight increase in the early 1990 until 1999, after which a slight decline and levelling off is evident. The fits to S3 the Portuguese series and S7 Reunion show a modest and slight decline, respectively, throughout. The model interpreted these trends by predicting a decreasing total biomass through time. The spawning output was estimated to increase slightly in the late 1990s to the early 2000s followed by a period of decline coincident with the increase in catch (Figure 2) and decline in the CPUE series.

**Fits to the Length composition**

The differences estimated in the sex-specific selectivity curves for many of the fisheries reinforce the observations of biologists for areas of sex-segregation during the life history of blue sharks (Figure 12). With the exception of the Japanese longline fishery; all fisheries where sex specific selectivity could be estimated resulted in a lower peak selectivity (therefore catchability) for females.

The overall fit to the length data was generally good (Figure 13). Fleet specific annual length samples were often quite different, i.e. left skewed one year and bimodal the next, which accounts for the small amount of misfit in the aggregated samples. When attempting to estimate selectivity curves for fisheries with sex specific patterns the model often did not converge, therefore the sex specific offsets were fixed. Pearson residuals of the fit to the length compositions were small – on the order of 2 to -2 and did not show any temporal trend (Figures 14-15).

***Stock-recruitment Parameters***

The predicted virgin recruitment (R0; number of age 0 pups) was approximately 2,014,000 animals and the number of estimated pups was relatively constant from the early 1960 through the early 1980s, after which estimated recruitment slowly declined, and then experienced large fluctuations from 1990-2019 (Figure 16). The bias correction estimated in the model is shown in Figure 17. The corresponding estimated stock recruitment relationship and annual deviations are also shown in Figure 18.

***Fishing Mortality***

Estimated $F/F_{MSY}$ and fleet-specific instantaneous fishing mortality rates are presented in Figures 19 and 20 respectively. Fishing mortality was relatively low from the 1950 to the mid 1990s, which is in accordance with low catches and effort during that period. In the late 1990s fishing mortality increased with the advent of F1 the Miscellaneous fishery, this fishery is comprised mostly of coastal longline (>98%), with trolling, sport and artisanal fisheries contributing small percentages of the catch. Starting in the late-1990s overall fishing mortality began to increase sharply, with large fluctuations in the individual fisheries contribution to the overall fishing mortality. The overall fishing mortality has been below $F_{MSY}$ (i.e. overfishing is not occurring) for the entire time series, however, in recent years the confidence intervals have included values greater than one.

***Estimated stock status and other quantities***

The estimated equilibrium yield curve for the diagnostic case model is shown in Figure 21. The estimated MSY is approximately 33,500 MT and this is predicted to occur at 34% of the unfished biomass (Figure 21), which is less than the standard Schaefer production model ($0.5B_0$). The diagnostic case model estimates that the total biomass of the stock was at approximately 100% of the unfished level at the start of the model period (Figure 11) and steadily decreased to an estimate of $SB_{2019}/SB_{MSY} = 1.2$ that corresponds with $F_{2019}/F_{MSY} = 0.81$. Recruitment is fairly well estimated throughout the model time period (Figure 8), with recent recruitment estimated to be lower than the implied stock recruitment curve due to deviations implied by the length data. The estimates of recruitment were quite tightly constrained to the stock recruitment curve for the initial period of the model when there was no length information to inform the model. The main trends in the population dynamics can be explained through the estimated fishing mortality which was greatly increased in the 1990s and early 2000s due to the increase in catch (Figures 19 and 20). These changes in fishing mortality correspond to an overall stock status that is headed from a virgin state to the direction of overfished and overfishing (Figure 22).

**Model Uncertainty**

Stock status uncertainty was evaluated delta-Multivariate lognormal (MVLN)approximation to generate joint error distributions for $SSB_{2019}/SSB_{MSY}$ and $F_{2019}/F_{MSY}$. Figure 22 shows the estimated stock status based on the MVLN analysis for the base case model and Figure 23 shows the estimated timeseries based on the MVLN approximation. Figure 24 shows the distribution of the MLE estimates of $SSB_{2019}/SSB_{MSY}$ and $F_{2019}/F_{MSY}$.

Stock synthesis provides estimates of the MSY-related quantities and these and other quantities of interest for management are provided in Table 4. We note that the IOTC has not yet adopted target or limit reference points for any shark species, so a suite of MSY-related quantities are presented.

***Retrospective Analysis***

As part of an analysis of model structure, retrospective analysis (sequentially deleting 1 year of data from the end of the model and re-running) was run using the base case formulation (the Portuguese, Japanese late and Reunion series and the GAM estimated catches). The estimates of spawning depletion remain very similar across all the retrospective model runs considered (Figure 25) indicating that the changes in estimates of virgin spawning biomass are based on the total catch (Figure 25 right panel). The last retrospective run (-5 years) estimated a more depleted stock that corresponds to a slightly smaller virgin recruitment (Figure 25 right panel), this is associated with higher estimated total fishing mortalities in the last 4 years. In general the retrospective analysis shows no large departures from the estimated scale, depletion, or overall trend based on the sequential deletion of the last 7 years of data.

### 3.2 Other model diagnostics

Annex 1 shows the results of the expanded analysis on the residuals, hierarchal cluster analysis, runs tests joint residual plots, likelihood profiles, age structured production model and hindcasting cross-validation are presented in Annex 1. Select details from Annex 1 are repeated here, the reader is encouraged to read Annex 1 in its entirety. Additionally the control file used for the diagnostic case model is shown in Annex 2.

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

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

Annex 3 shows the application of the generalized Bayesian State-Space Surplus Production Model framework JABBA (Just Another Bayesian Biomass Assessment) to the 2021 IOTC assessment input data for Indian Ocean blue shark. Five alternative scenarios were considered, with the groupings chosen to reflect the groupings outlined in the stock synthesis (SS3) assessment. Results indicate that the stock was not overfished nor subject to overfishing. All F/FMSY trajectories indicate that sustainable fishing mortality has never been exceeded for blue shark in the Indian Ocean, and that the probability of the stock being in the “green” quadrant of the Kobe plot is >99% in four out of five scenarios.

# 4 Conclusion

Although most pelagic sharks can be considered data poor when compared to targeted tuna and other teleosts, the information for blue shark in the Indian Ocean is relatively abundant because they are the most commonly caught pelagic shark. Although blue shark lack the traditional fisheries statistics such as landings and historic catch rates (CPUE series), blue shark have been caught in mixed target fisheries for at least the last two decades. The resulting CPUE series from these fisheries are concentrated in the most recent decade, and all come from fishery dependent longline sources. An issue of concern regarding the indices of relative abundance, is that many of them show inter-annual variability that does not seem to be compatible with the life history of the species, suggesting that the GLMs used to standardize the indices did not include all factors to help track relative abundance or that the spatial scope of sampling is too limited to allow for precise inference about stock-wide trends. The CPUE series that were used in the base case model came from on board observers and covered the majority of the southern Indian Ocean, however, the bulk of the observed effort was in the southwestern Indian Ocean in the waters from South Africa and Madagascar. In the future intersessional work to further develop the indices of abundance would be important.

Recent work has led to similar estimates with respect to age, growth, reproduction and the associated life history characteristics. As such the range of variation investigated in the previous assessment was not undertaken for this study. The parameterization of the model reflected the best available estimates. Changes to the biology and life history inputs were minor with respect to the last assessment. Changes were: steepness is now 0.8 (from a range 0.3 – 0.5); the theoretical maximum length has changed a few centimeters. These changes affect the potential productivity/resiliency of the stock in different ways but the overall characteristics of shark with moderate productivity (fecundity) and an annual long gestation period have remained.

The results of the assessment are compared across different groupings of CPUE series and show the diagnostic case parameterization resulting in estimates of $SB_{2019}$ /$SB_{MSY}$ =1.21 and $F_{2019}$/$F_{MSY}$ = 0.81 Stock status is reported in relation to MSY based reference points however the authors note that the IOTC has not yet adopted reference points for sharks. Due to the inherent unreliability of recruitment estimates in the terminal year this study defines ‘current’ as the average of the first four of the last five years (i.e. 2015-2018), and reports ratios of SB and F as current as well as with respect to 2019.

The main conclusions of this assessment are:

- The estimates of catch are highly influential in the model, but mostly in terms of scale, as the current depletion and fishing mortality indicators are approximately equal across all alternative catch estimates for a given CPUE series.

- The scale of the assessment is influenced by the CPUE series chosen, across these estimates the estimates of B0 range from approximately 1million MT to approximately 1.9 million MT.

The main drivers of this assessment are the trend in the catch and CPUE series. In particular the large increase in recent years of catch has different interpretations (within the model). based on whether the CPUE series is variable (Japanese late) or decreasing (Portuguese and Reunion Fleet). Recommended studies that would improve future analyses are:

- Develop appropriate length inputs for all fleet.
- Further investigation of CPUE series and their representativeness.
- Develop region specific biological inputs.
- Further work on developing catch histories.
- Undertake collaborative study of blue shark CPUE from multiple Indian Ocean longline fleets

## 5 Acknowledgements

The author would like to thank the IOTC staff for their help in preparing the data as well as all of the member countries that submitted data and analyses. This assessment benefited from the discussion and contribution of the WEPB17 data prep attendees.

# 6 Reference


Andrade, I., Rosa, D., Lechuga, R., and Coelho, R. 2017. Age and growth of blue shark in the Indian Ocean. IOTC–2017–WPEB13–20

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

Clarke, S. 2011. Historical Catch Estimate Reconstruction for the Indian Ocean based on Shark Fin Trade Data. IOTC–2015–WPEB11–24.

Coelho, R., Rosa, D. 2017 catch reconstruction for the Indian Ocean blue shark: an alternative hypothesis based on ratios. IOTC-2017-WPEB13-22.

Coelho R, Mejuto J, Domingo A, et al. Distribution patterns and population structure of the blue shark (Prionace glauca) in the Atlantic and Indian Oceans. Fish Fish. 2017;00:1–17

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

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

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

Hsu H H, Joung S J, Lyu G T, Liu K M, Huang C C (2011) Age and growth of the blue shark, *Prionace glauca*, in the northwest Pacific. ISC/11/SHARKWG-2/INFO02.

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

IOTC–WPEB10 2014. Report of the 10th Session of the IOTC Working Party on Ecosystems and Bycatch. Yokohama, Japan, 27–30 October 2014. IOTC–2014–WPEB10–R[E]: 94 pp.

IOTC-2015-WPEB11-DATA03 Rev_1 DATA FOR THE ASSESSMENT OF INDIAN OCEAN BLUE SHARK. Working Party on Ecosystems and Bycatch (WPEB) 11. 7-11September 2015

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

Martin et. al. 2015. IOTC-2015-WPEB11-XX Estimation of blue shark catches in the Indian Ocean. Working Party on Ecosystems and Bycatch (WPEB) 11. 7-11September 2015

Martin, S., and Rice, J. 2017. Approaches to the reconstruction of catches of Indian Ocean blue shark. IOTC–2017–WPEB13–23.

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

Methot, R. D. 2009. User manual for Stock Synthesis: Model Version 3.04 (Updated September 9, 2009), 159p.

Nakano H. 1994 Age, reproduction and migration of blue shark (*Prionace glauca*) in the North Pacific Ocean. Bulletin - National Research Institute of Far Seas Fisheries (no.31) p. 141-256

Peterson I, Wroblewski J S. 1984. Mortality Rate of Fishes in the Pelagic Ecosystem, Can. J. Fish. Aquat. Sci., 41,1117-1120.

Rice J and Semba, Y., 2014. Age and Sex Specific Natural Mortality of the Blue Shark (*Prionace glauca*) in the North Pacific Ocean. ISC/14/SHARKWG-1/0X.

Rice , J and Sharma, R. 2015. Stock assessment blue shark (*Prionace glauca*) in the Indian Ocean using Stock Synthesis. IOTC–2015–WPEB11–28 Rev_1.

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

# 7 Tables

**Table 1. Fishery definitions for the Indian Ocean Assessment**

| Fleet/ Survey Number and Short Name | Gear(s) | Selectivity |
|---|---|---|
| F1 MISC | Costal longline, trolling, sport and artisanal fisheries | Fixed logistic |
| F2 GILL | Gillnet Fisheries | Fixed logistic |
| F3 OTHER_LL | All longline fishery other than Japan, TWN, China, Korea, Portugal and Spain. | Estimated double normal |
| F4 JPN_LL | Japanese longline fishery | Estimated double normal |
| F5 KOR_LL | Korean longline fishery | Estimated double normal |
| F6 PRT_LL | Taiwanese longline fishery | Estimated double normal |
| F7 TWN_LL | Portuguese longline fishery | Estimated double normal |
| F8 ESP_LL | Spanish longline fishery | Estimated double normal |
| S2 JPN_LATE | Japan late years longline CPUE | NA |
| S3 POR | Portugal longline CPUE | NA |
| S4 ESP | Spain longline CPUE | NA |
| S5 TWN | Taiwanese longline CPUE | NA |
| S7 REU | EU-Reunion longline CPUE | NA |
| S8 ZAF | South African EEZ Longline | NA |

**Table 2: Estimates of age-specific natural mortality used in the assessment. The diagnostic case used those based on the approach of Peterson and Wroblewski (1984) method and estimates of growth from Andrade et al 2019.**

| Age | Female | Male |
|---|---|---|
| 0 | 0.453 | 0.453 |
| 1 | 0.288 | 0.277 |
| 2 | 0.222 | 0.212 |
| 3 | 0.187 | 0.179 |
| 4 | 0.165 | 0.158 |
| 5 | 0.15 | 0.144 |
| 6 | 0.139 | 0.134 |
| 7 | 0.131 | 0.126 |
| 8 | 0.124 | 0.121 |
| 9 | 0.119 | 0.116 |
| 10 | 0.115 | 0.113 |
| 11 | 0.112 | 0.11 |
| 12 | 0.109 | 0.107 |
| 13 | 0.107 | 0.105 |
| 14 | 0.105 | 0.104 |
| 15 | 0.103 | 0.102 |
| 16 | 0.102 | 0.101 |
| 17 | 0.101 | 0.1 |
| 18 | 0.1 | 0.1 |
| 19 | 0.099 | 0.099 |
| 20 | 0.098 | 0.098 |
| 21 | 0.098 | 0.098 |
| 22 | 0.097 | 0.097 |
| 23 | 0.097 | 0.097 |
| 24 | 0.096 | 0.097 |
| 25 | 0.096 | 0.097 |
| 26 | 0.095 | 0.096 |
| 27 | 0.095 | 0.096 |
| 28 | 0.095 | 0.096 |
| 29 | 0.095 | 0.096 |
| 30 | 0.095 | 0.096 |

**Table 3. Summary of SS3 specification options for the Indian Ocean blue shark assessment models. Other assumptions were constant for all models, a total 6 sensitivity runs were completed. The bold text indicates the base case configuration.**

| CPUE Grouping Number | CPUE data included | Notes |
| --- | --- | --- |
| **1** | **PRT, REU, JPN** | **Previous Base Case** |
| 2 | ESP, REU | Positively correlated Group1 |
| 3 | PRT, ESP, REU, JPN, ZAF | Positively correlated Group2 |
| 4 | TWN, REU, JPN | Minimally negatively correlated |
| 5 | All | |

| Stepness Assumptions | | Notes |
| --- | --- | --- |
| 0.75 | | Lower bound of 95% CI |
| 0.8 | | Estimate (Rosa & Coelho 2017) |
| 0.84 | | Upper bound of 95% CI |

| Recruitment deviation | | |
| --- | --- | --- |
| **0.2** | | **Low (Reference Case)** |
| 0.4 | | Hi |

**Table 4: model (in bold text) and sensitivity runs. Stock status in 2019 is in the grey shaded rows.**

| CPUE | **PRT, REU, JPN** | ESP, REU | PRT, ESP, REU, JP | TWN, REU, JPN | All | PRT, REU, JPN | PRT, REU, JPN | PRT, REU, JPN |
|---|---|---|---|---|---|---|---|---|
| Steeepness | **0.8** | 0.8 | 0.8 | 0.8 | 0.8 | 0.75 | 0.8 | 0.84 |
| SigmaR | **0.2** | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.2 |
| C2019_msy | **1.29** | 1.29 | 1.22 | 1.26 | 1.20 | 1.34 | 1.27 | 1.25 |
| Y_MSY | **33,532** | 33,600 | 35,452 | 34,212 | 36,022 | 32,251 | 34,180 | 34,525 |
| B_zero | **1,029,200** | 1,031,360 | 1,091,270 | 1,050,480 | 1,109,390 | 1,035,310 | 1,043,160 | 1,025,020 |
| B_msy | **338,377** | 339,090 | 358,411 | 345,321 | 364,293 | 354,130 | 343,616 | 325,447 |
| B_cur | **384,354** | 384,882 | 442,014 | 405,148 | 459,087 | 381,991 | 406,037 | 386,117 |
| SB_zero | **118,625** | 118,874 | 125,779 | 121,077 | 127,867 | 119,329 | 120,233 | 118,143 |
| SB_msy | **39,001** | 39,083 | 41,310 | 39,801 | 41,988 | 40,817 | 39,605 | 37,511 |
| SB_cur | **51,492** | 51,729 | 58,682 | 54,108 | 60,833 | 51,192 | 54,077 | 51,716 |
| SB_2019/SB_msy | **1.21** | 1.21 | 1.35 | 1.27 | 1.39 | 1.15 | 1.27 | 1.27 |
| SB_cur/SB_msy | **1.32** | 1.32 | 1.42 | 1.36 | 1.45 | 1.25 | 1.37 | 1.38 |
| SB_cur_init | **0.43** | 0.44 | 0.47 | 0.45 | 0.48 | 0.43 | 0.45 | 0.44 |
| Fcur | **0.28** | 0.27 | 0.24 | 0.26 | 0.23 | 0.28 | 0.26 | 0.27 |
| F_msy | **0.31** | 0.31 | 0.31 | 0.31 | 0.31 | 0.28 | 0.31 | 0.33 |
| F_2019/msy | **0.81** | 0.81 | 0.68 | 0.76 | 0.64 | 0.91 | 0.74 | 0.74 |
| F_cur/msy | **0.90** | 0.89 | 0.77 | 0.85 | 0.74 | 1.00 | 0.85 | 0.82 |
| SB_2019 | **47,315** | 47,168 | 55,724 | 50,412 | 58,244 | 46,892 | 50,120 | 47,627 |
| F_2019 | **0.25** | 0.25 | 0.21 | 0.23 | 0.20 | 0.25 | 0.23 | 0.25 |
| B_2019 | **354,936** | 353,351 | 421,637 | 379,228 | 441,293 | 351,513 | 387,050 | 357,464 |

# 8 Figures

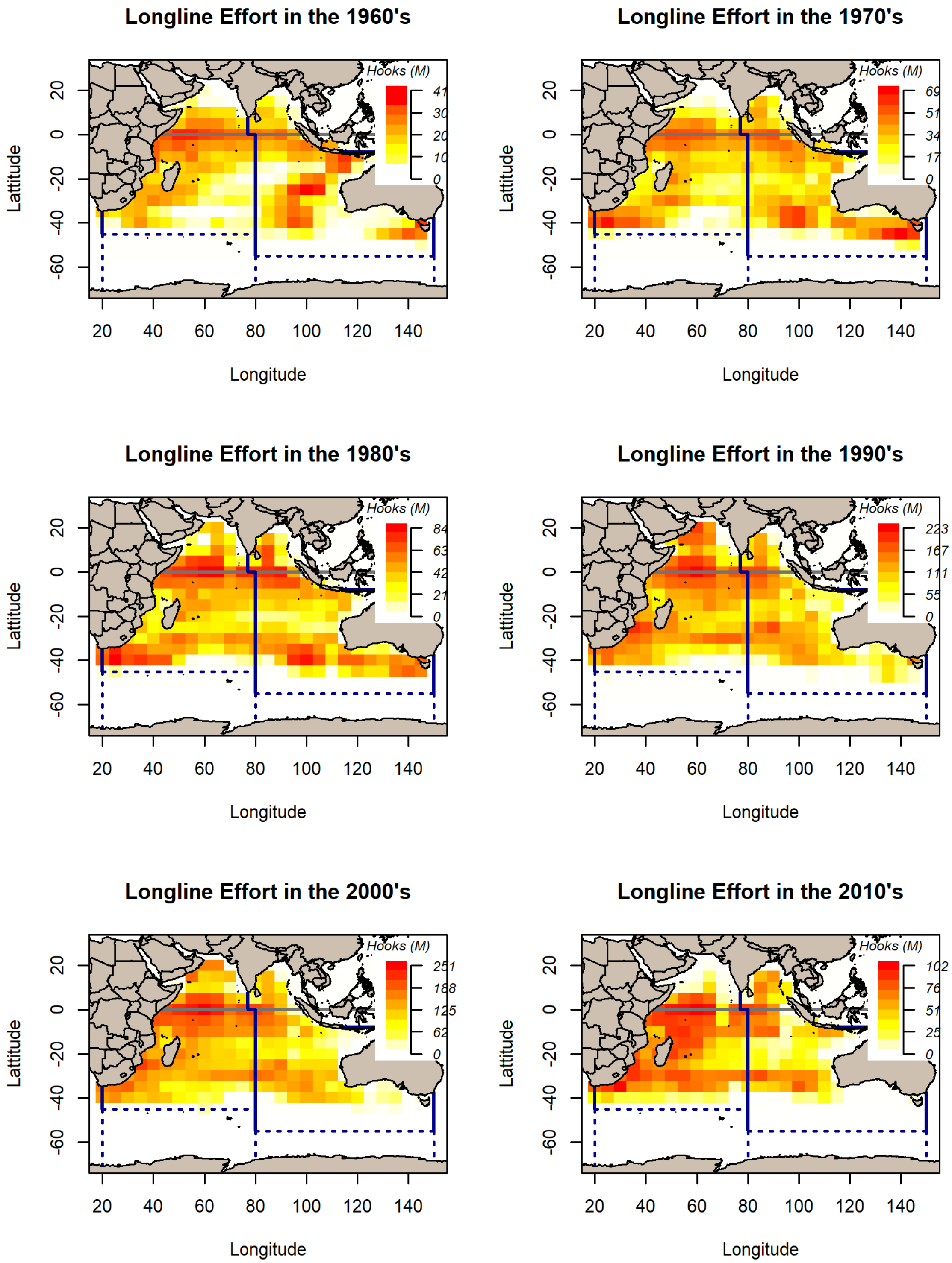
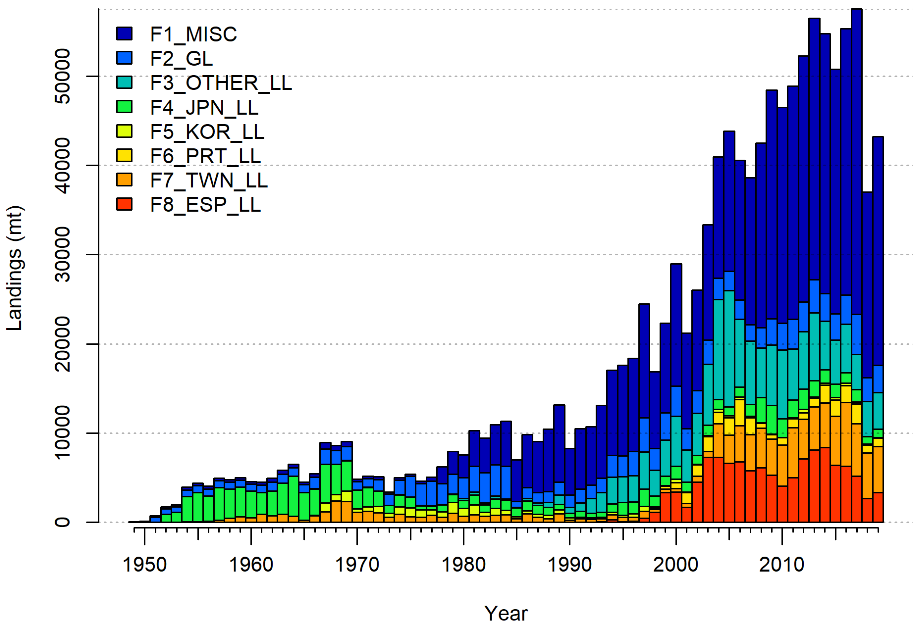



**Figure 1. Study area and reported longline effort by decade 1960s-2010s. The darker colors indicate higher effort (in millions of hooks), note the scale varies by decade.**



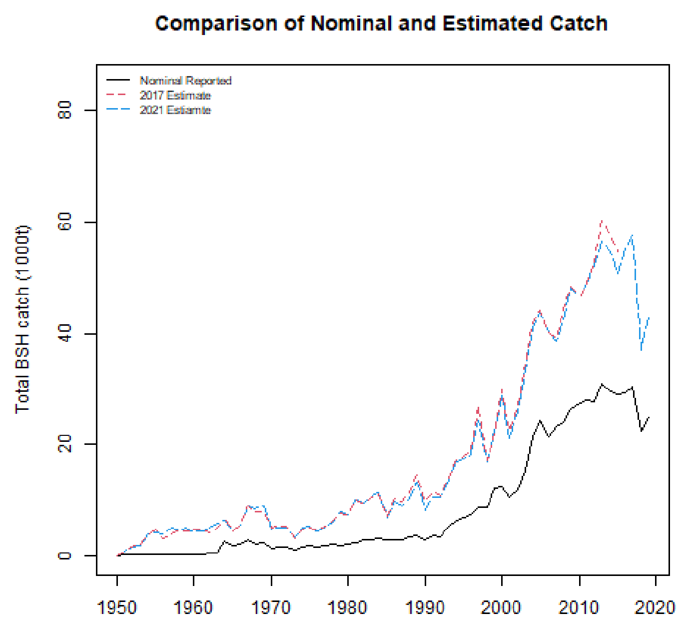




**Figure 2 Estimated total blue shark catch in mass by fishery over time for the Indian Ocean based on the GAM estimates of catch (top panel). Comparison of the 2021 estimated catch, the 2017 estimated catch and the nominal catch(bottom panel).**

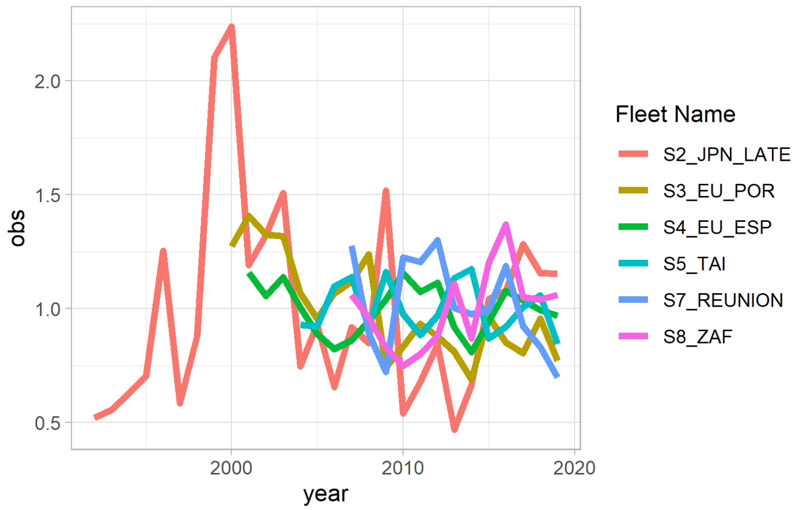



**Figure 3. Standardized CPUE for Japanese, Portuguese, Taiwanese ,Spanish, South African, and EU Reunion longline fleets based on papers submitted to WPEB-16 and 17. All series have been rescaled by their mean so that they are visually comparable for relevant periods of overlap.**

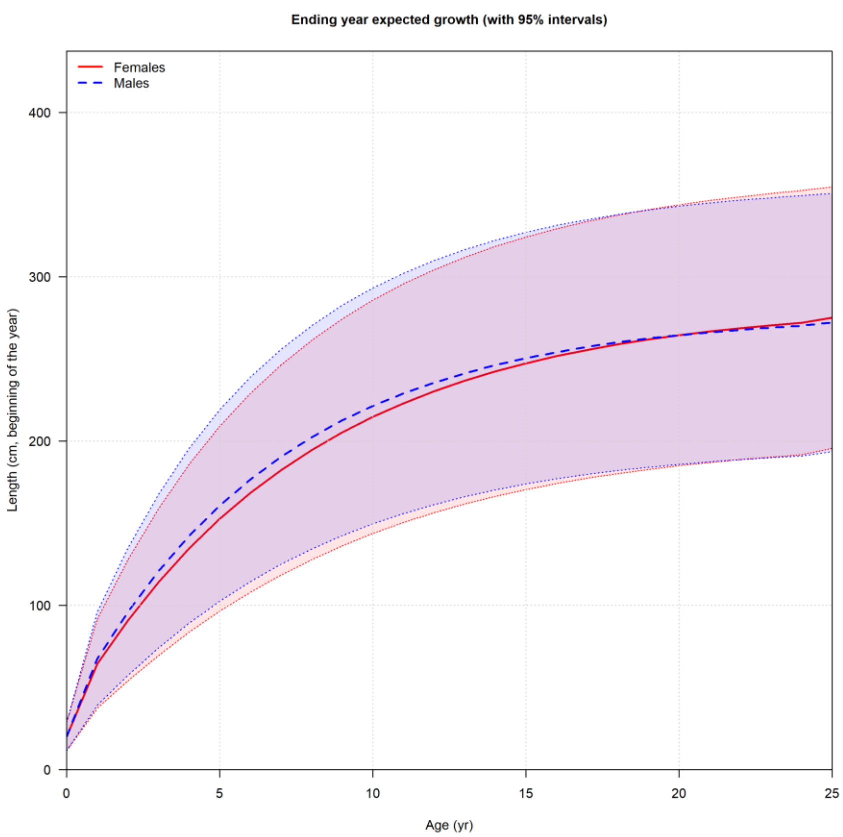



**Figure 4. Sex-specific growth curves (from Coelho et al 2017) calculated based on blue sharks in the Indian Ocean.**

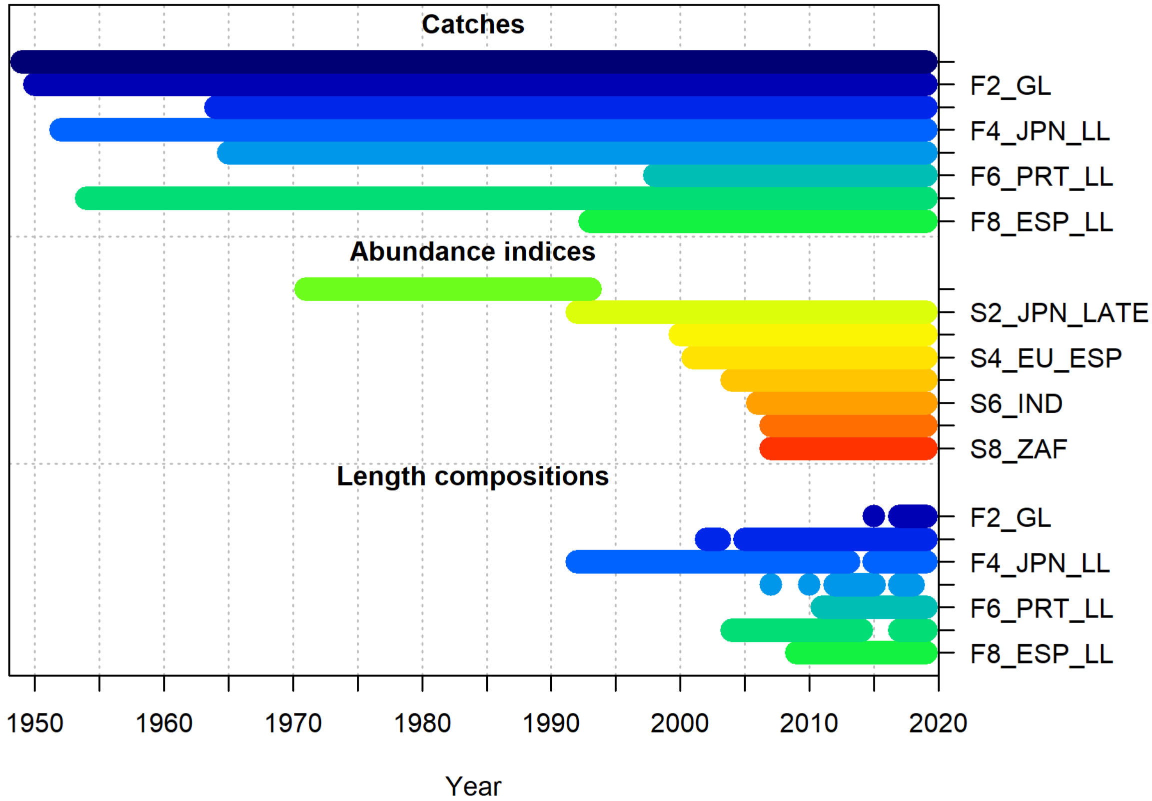



**Figure 5: Temporal data coverage for the diagnostic case model for the assessment of blue sharks in the north Pacific.**

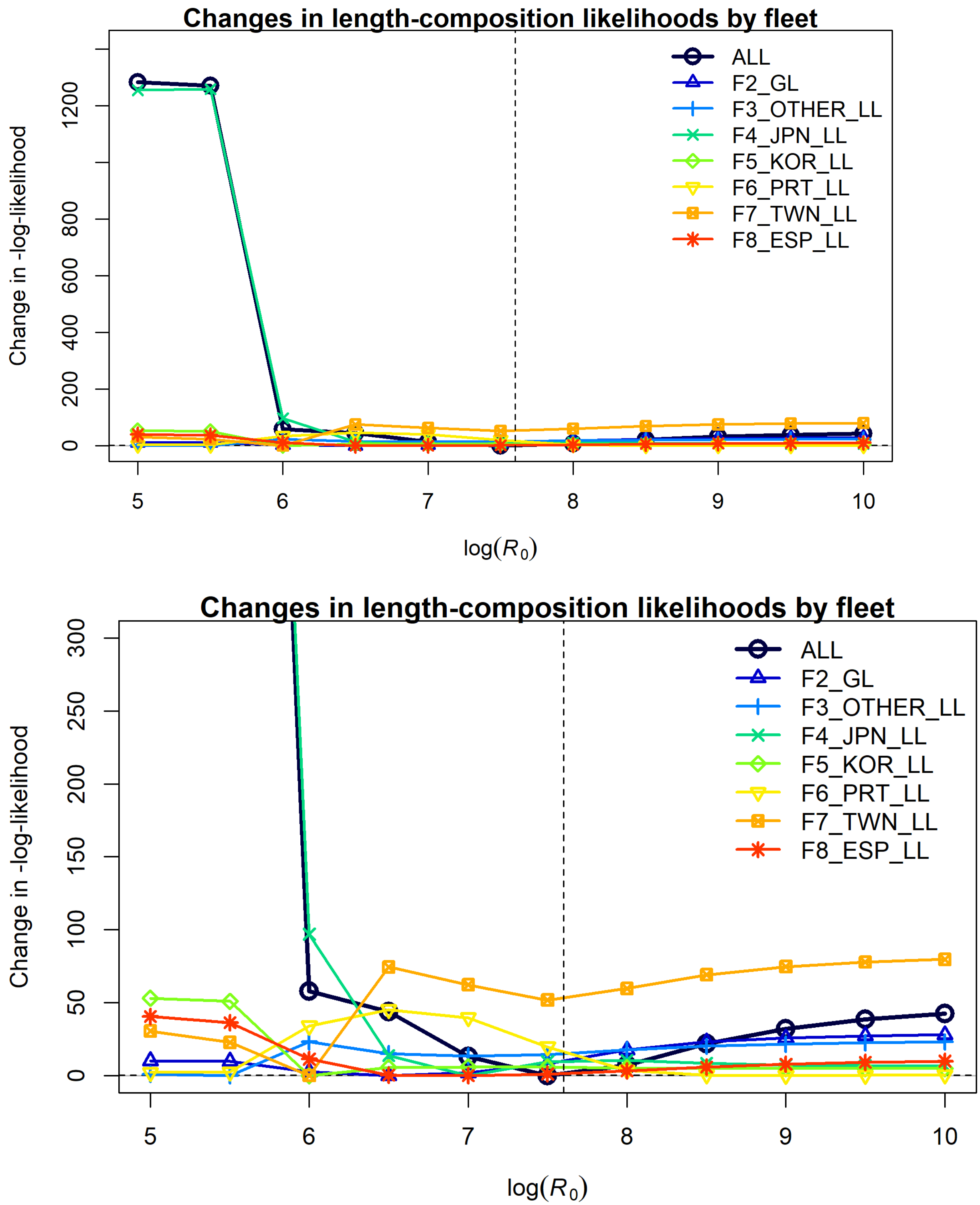
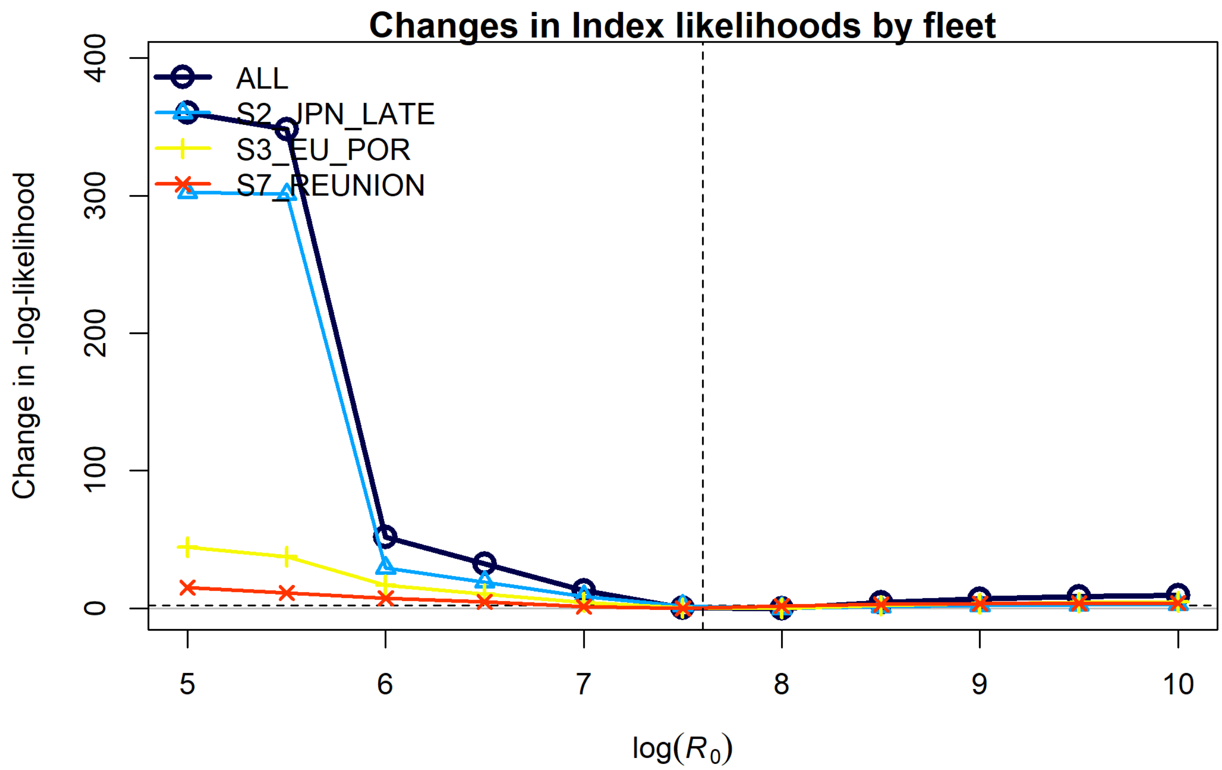



**Figure 6: Likelihood profiles for length composition, the bottom panel is a close up version of the top..**



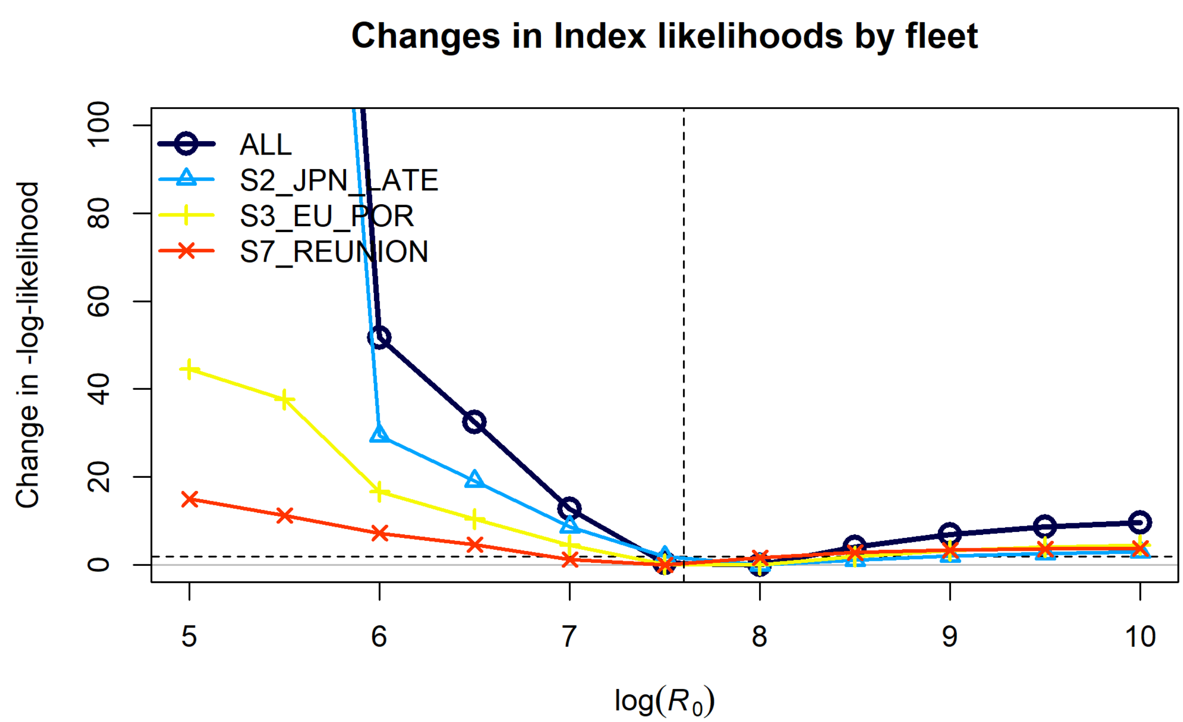




**Figure 7**: Likelihood profiles for the CPUE components for the reference run, The Bottom panel is a close up of the top.

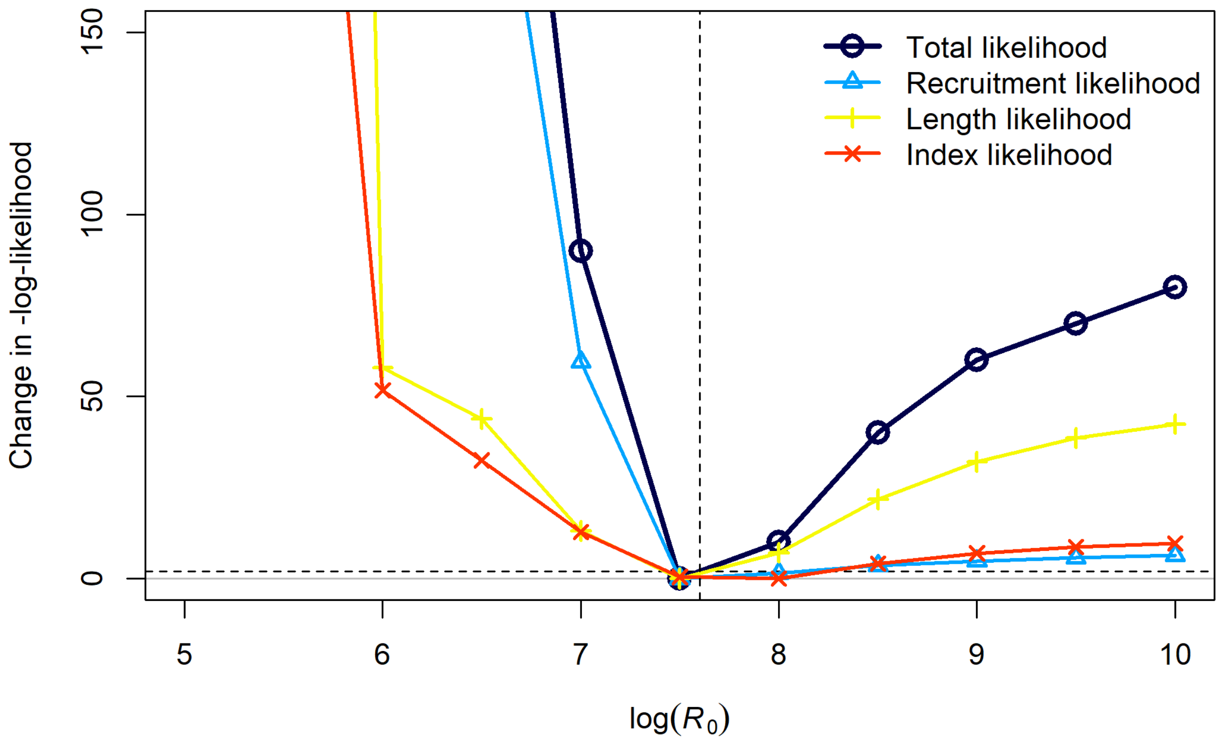



**Figure 8** Likelihood profile for the total likelihood, based on the diagnostic run.

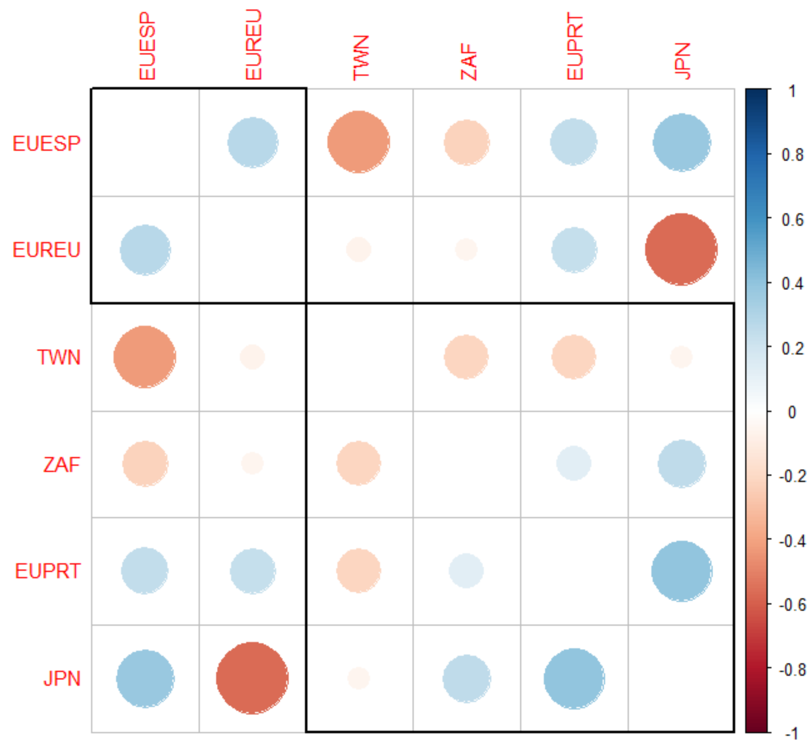



**Figure 9.** Correlation matrix for CPUE indices available for the Indian Ocean blue shark. Blue indicates positive and red negative correlations. The order of the indices and the rectangular boxes are chosen based on a hierarchical cluster analysis using a set of dissimilarities.

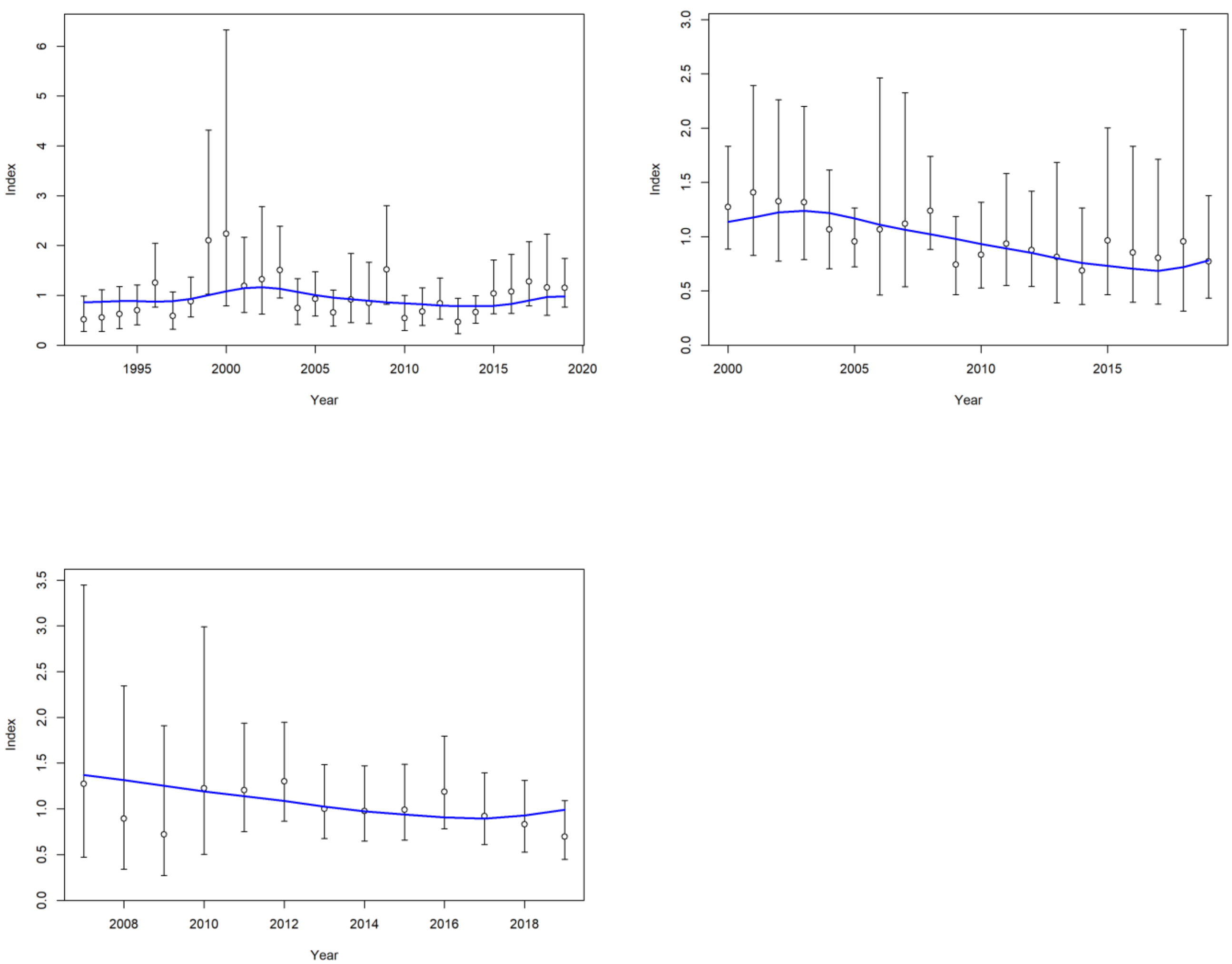



**Figure 10: Diagnostic case fit to the CPUE series, presented on a log scale. The top left panel is the Japanese late series (S2) the top right is the Portuguese series (S3) and the bottom is S7 Reunion.**

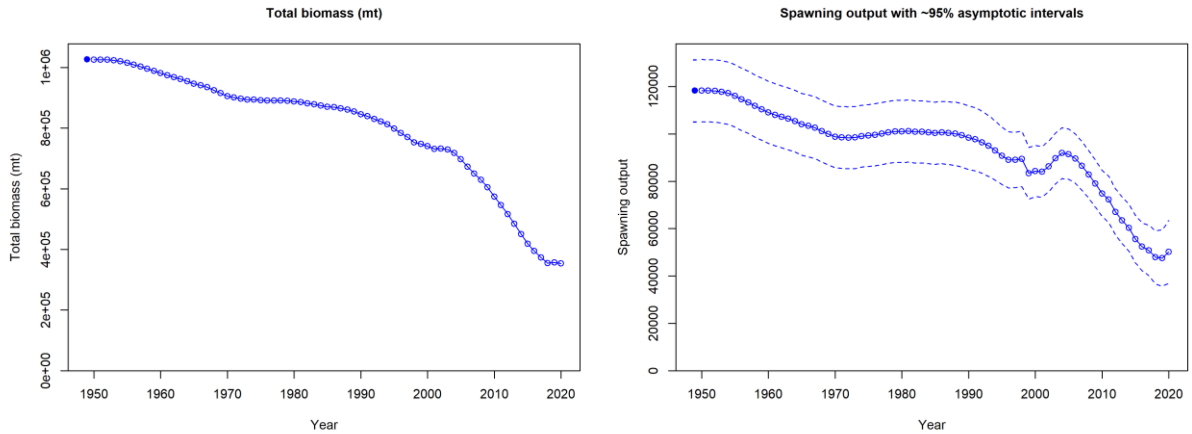



**Figure 11: Total biomass (left) and spawning potential (output) for the diagnostic case parameterization model. The filled dot represents the pre-model estimate of unfished biomass.**

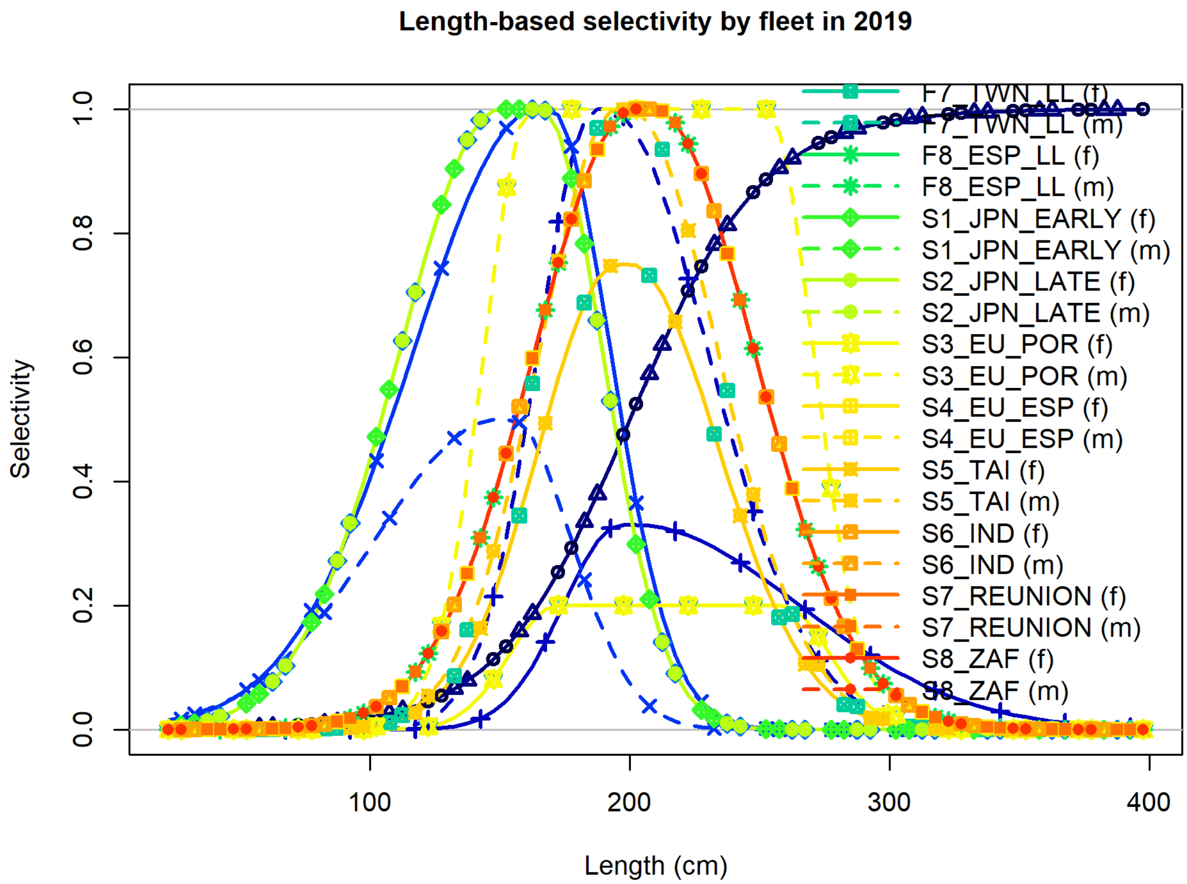



**Figure 12: Selectivity curves estimated for female and male from the diagnostic case model for the assessment of blue sharks in the Indian Ocean.**

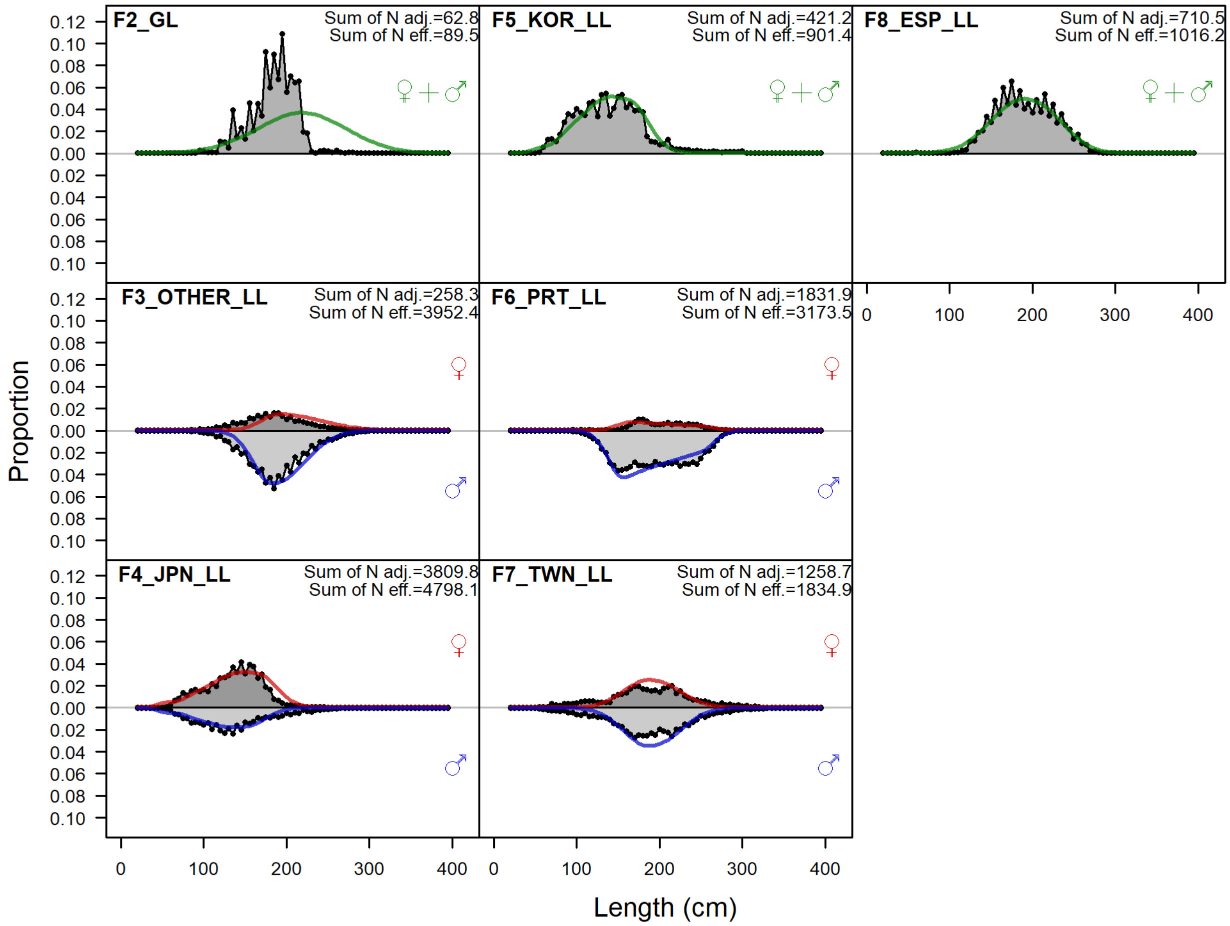



**Figure 13. Fits to the length frequency data for the diagnostic case model for the assessment of blue sharks in the Indian Ocean.**

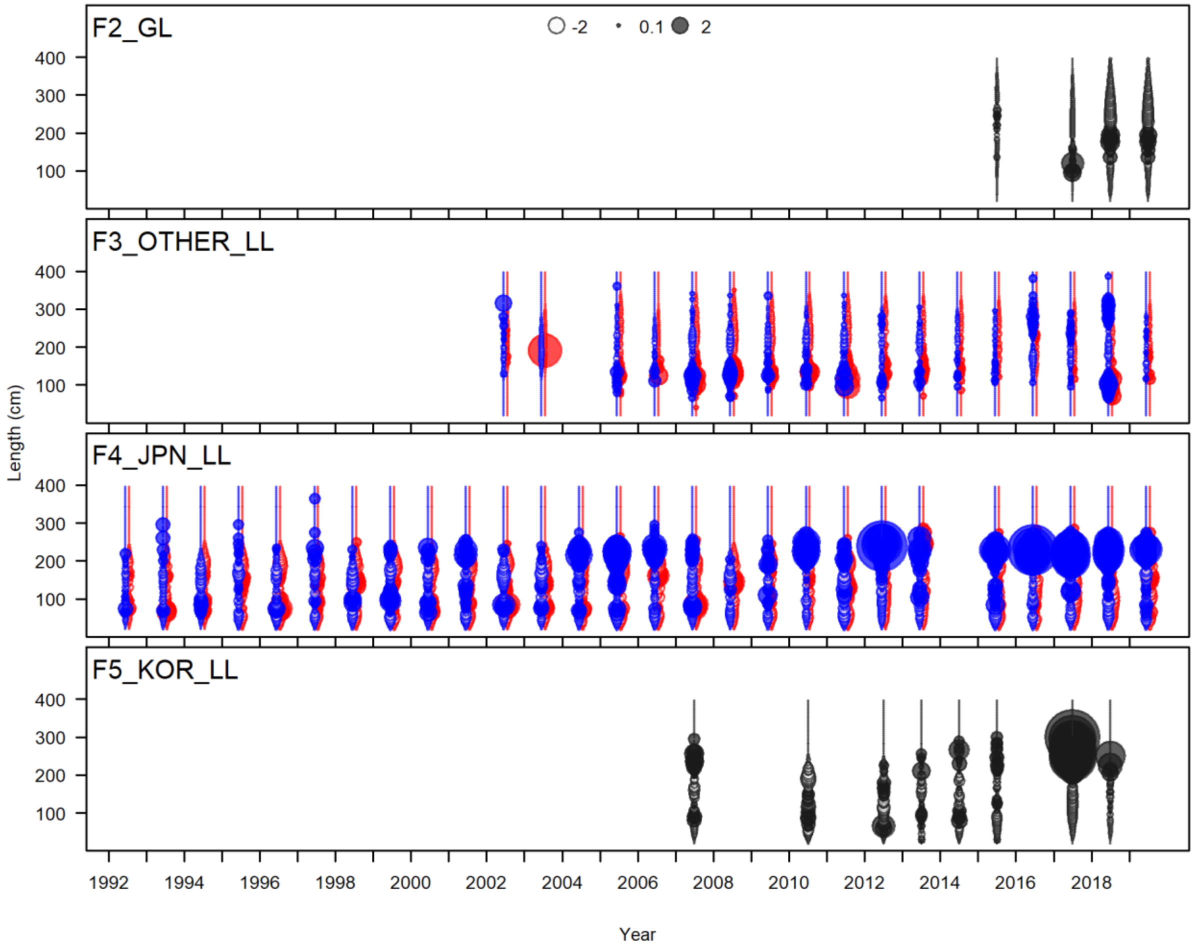



**Figure 14. Residuals from the fits to the length frequency data for the diagnostic case model for the assessment of blue sharks in the Indian Ocean, Fleets 2-5. combined sex length data is shown in black, female length data in red and male length data in blue. Closed bubbles are positive residuals and open bubbles are negative residuals, bubble sizes are scaled to maximum within each panel. Thus, comparisons across panels should focus on patterns, not bubble sizes**

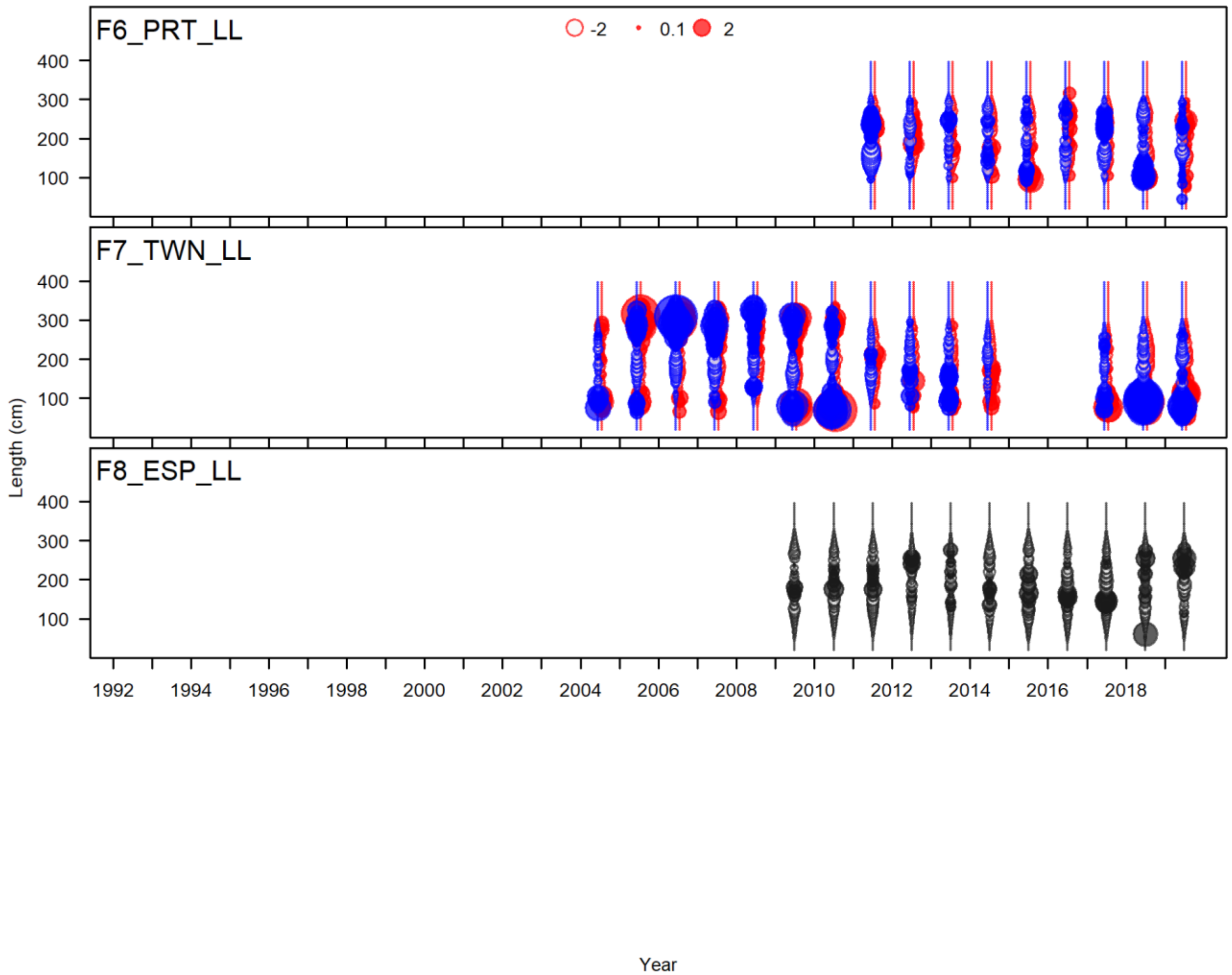



**Figure 15 Residuals from the fit to the length frequency data for the diagnostic case model for the assessment of blue sharks in the Indian Ocean, Fleets 6-8. Combined sex length data is shown in black, female length data in red and male length data in blue. Closed bubbles are positive residuals and open bubbles are negative residuals, bubble sizes are scaled to maximum within each panel. Thus, comparisons across panels should focus on patterns, not bubble sizes.**

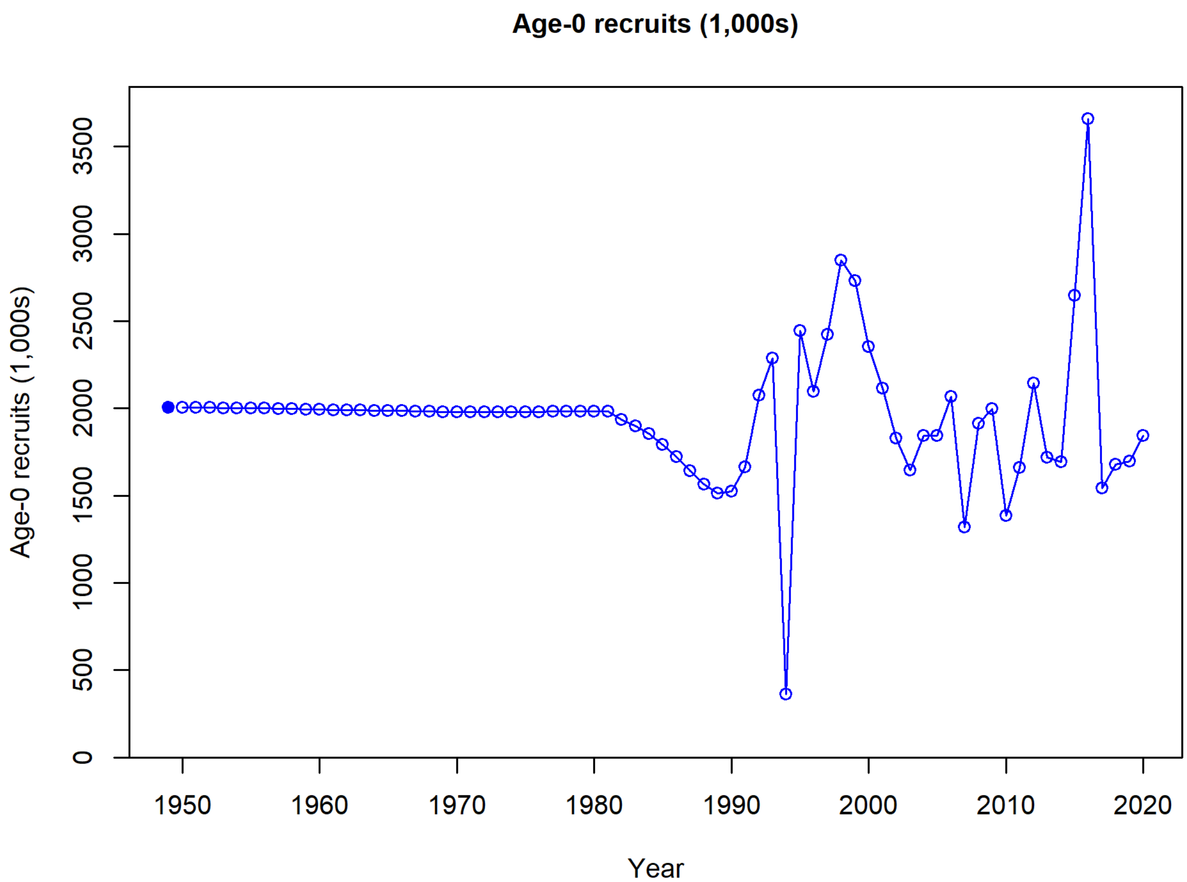



**Figure 16 .Estimated recruitment including the estimate of virgin recruitment (filled circle at the start of the time series) for the diagnostic case model for the assessment of blue sharks in the Indian Ocean.**

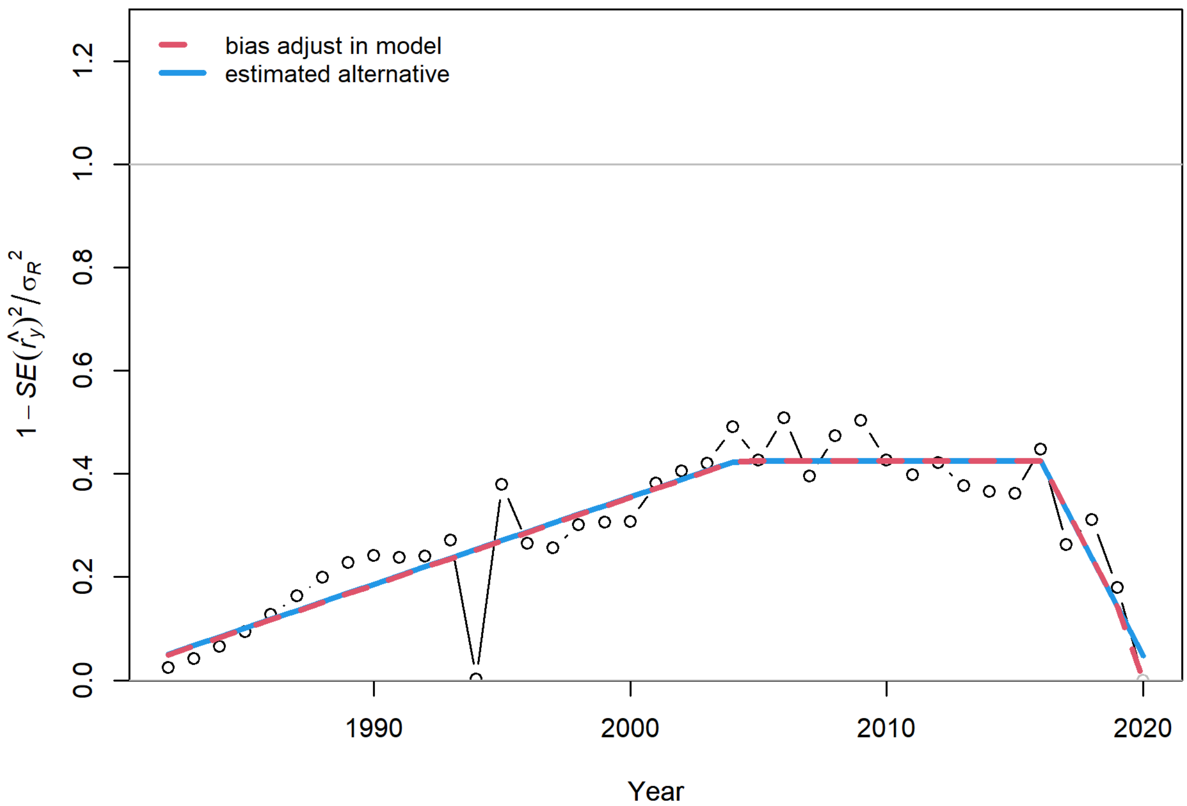



**Figure 17 .Estimated bias adjustment in the model.**

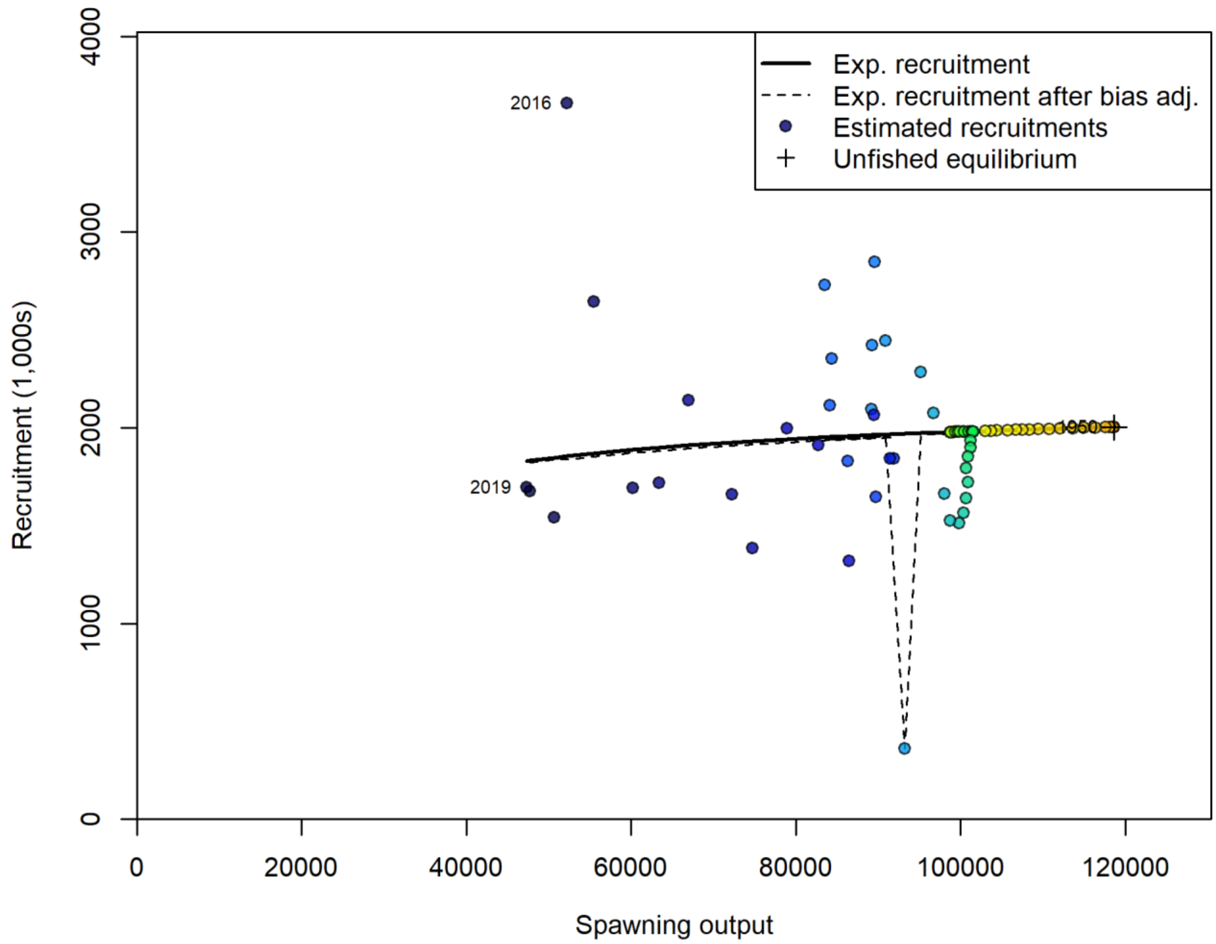



**Figure 18 Stock recruitment curve used in the assessment and time series of estimates of recruitment deviations (colored points).**

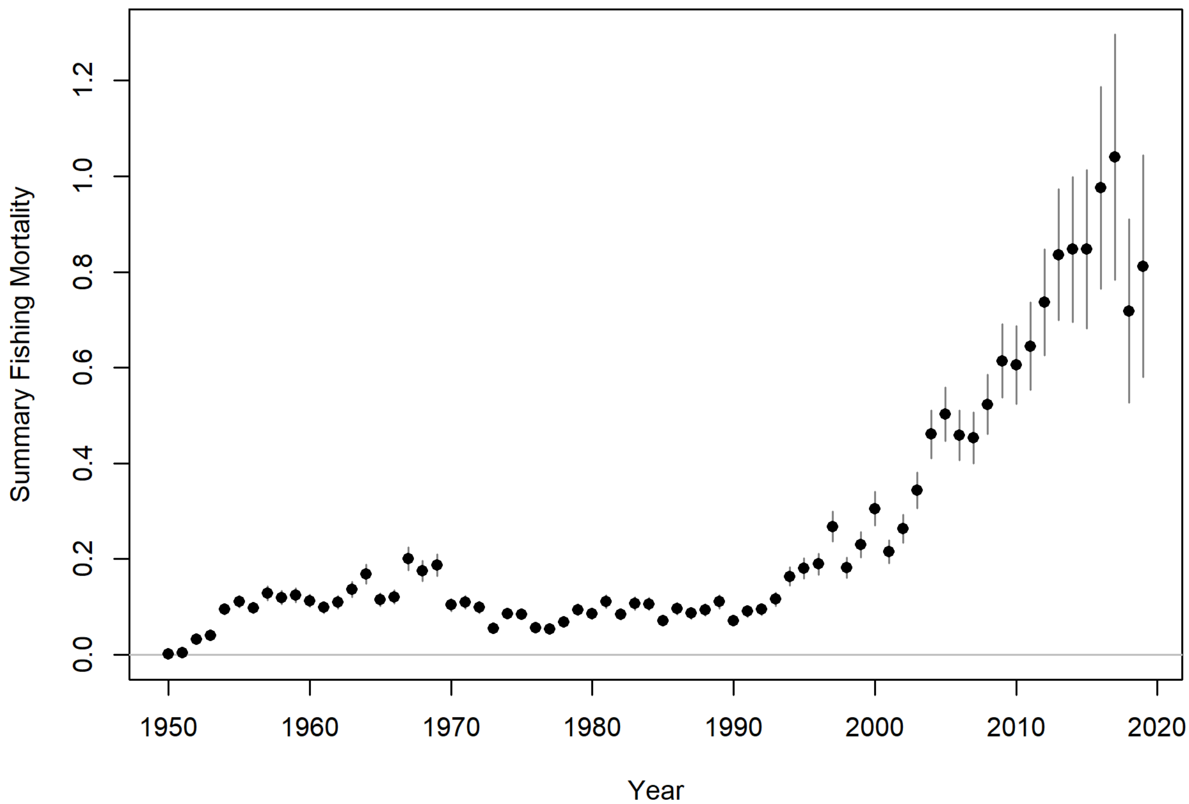



**Figure 19 Estimated total fishing mortality/FMSY**.

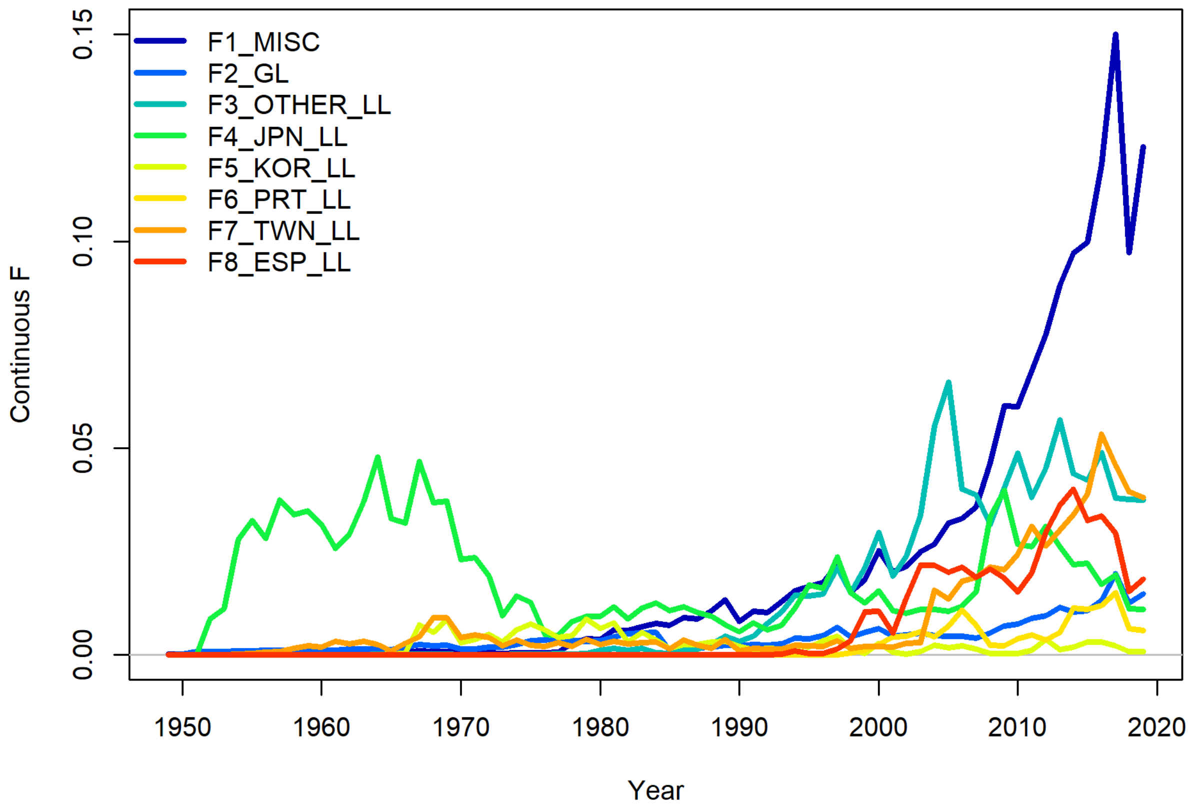



**Figure 20. Estimated fleet specific fishing mortality by year for the base case model configuration.**

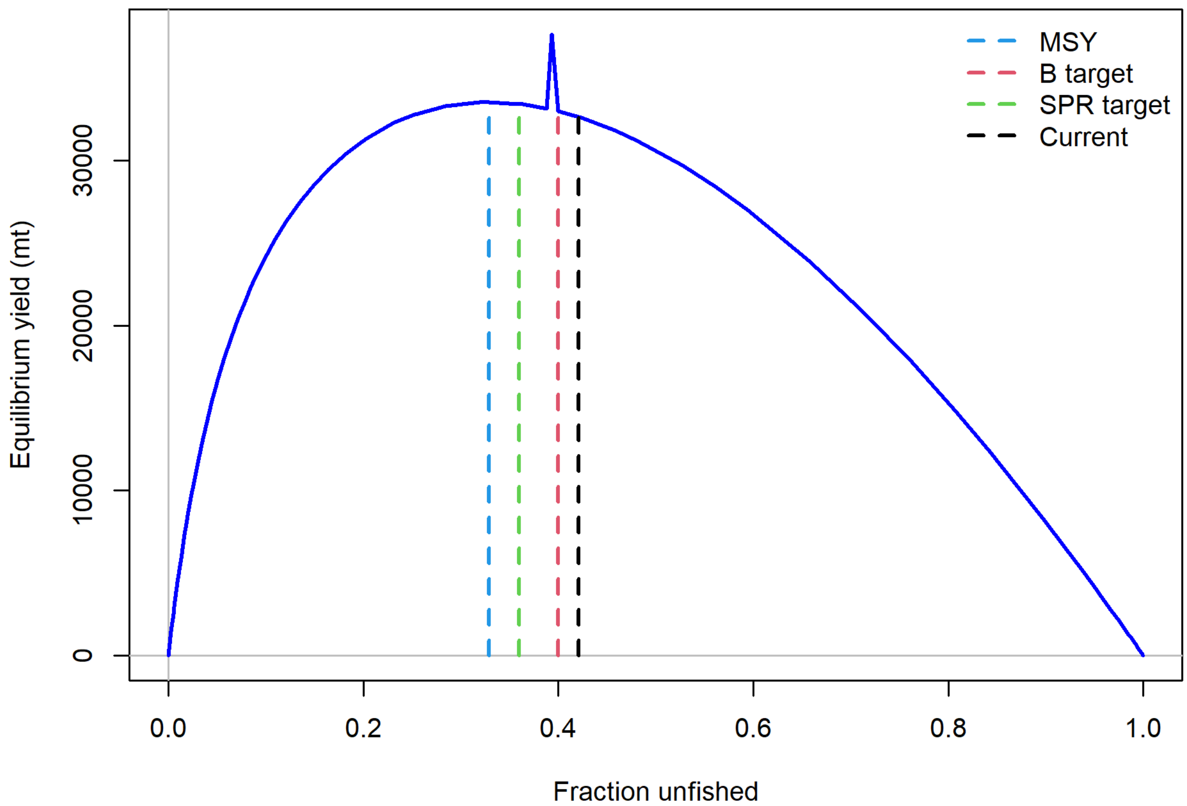



**Figure 21. Equilibrium yield curve for the diagnostic case model for the assessment of blue sharks in the Indian Ocean.**

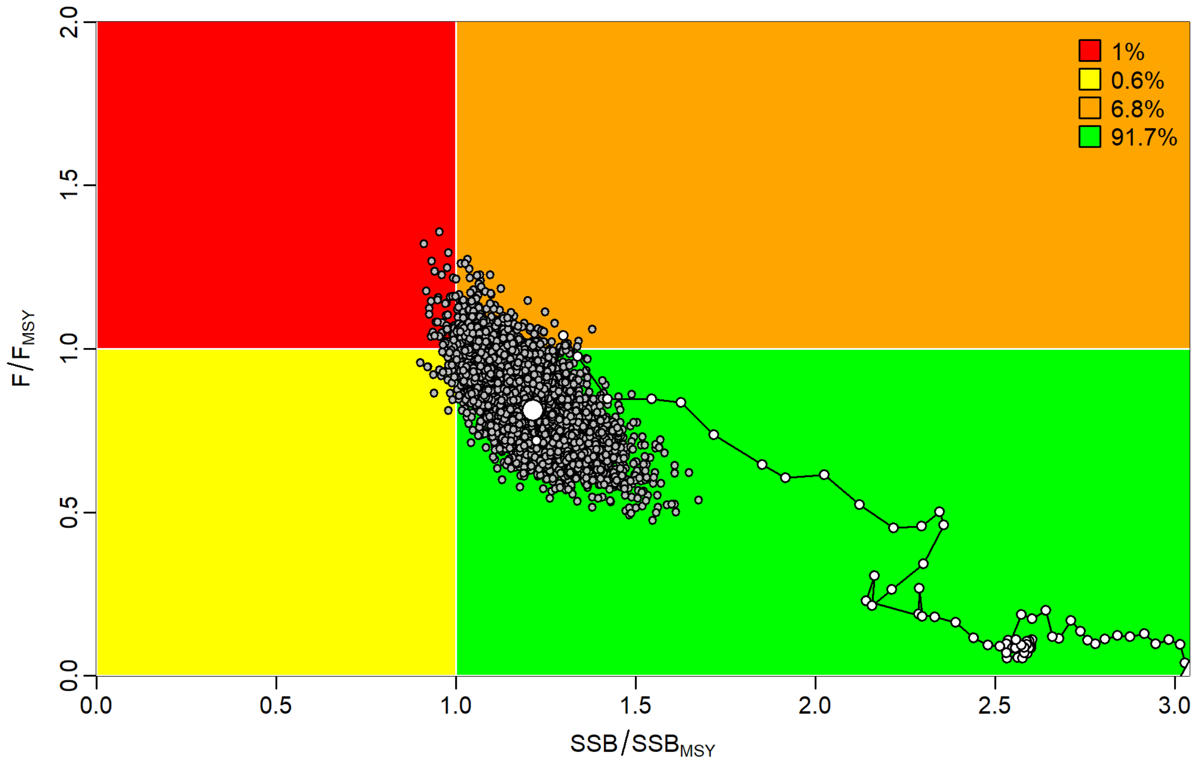



**Figure 22. Kobe plot of the annual stock status**

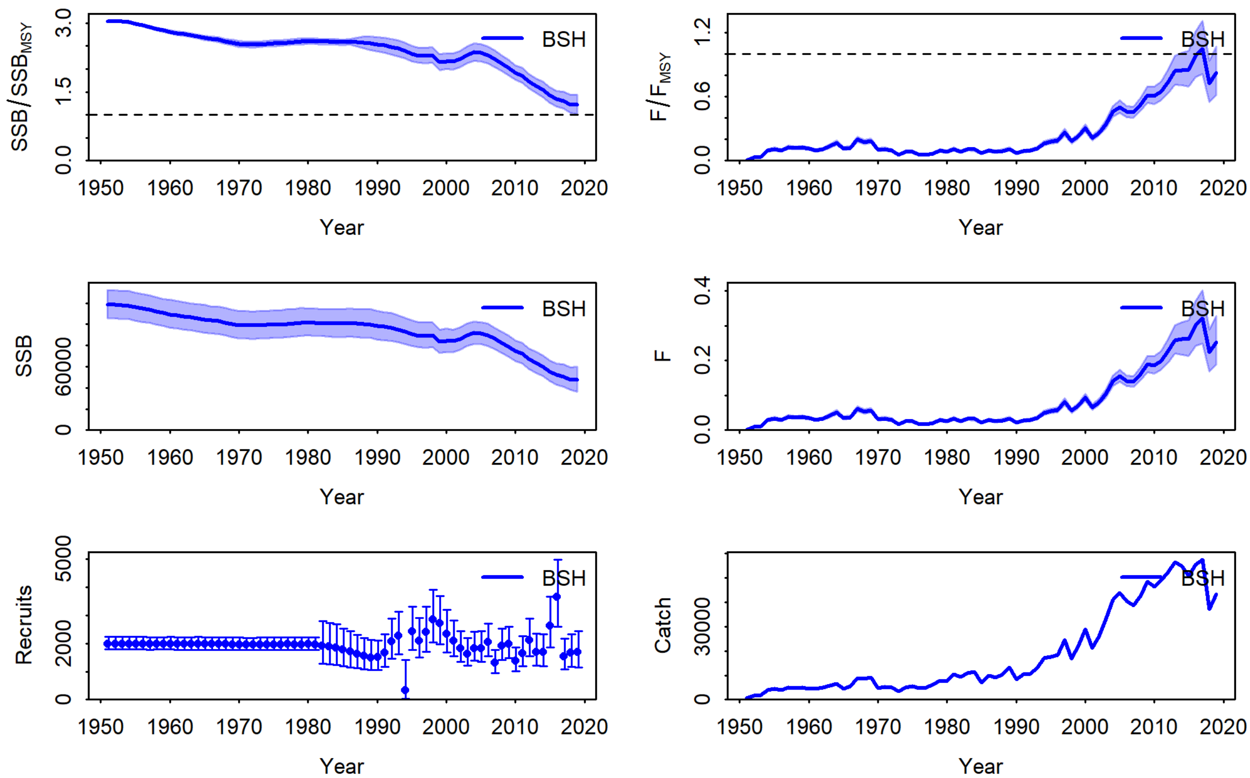



**Figure 23. Estimated timeseries based on the MVLN approximation**

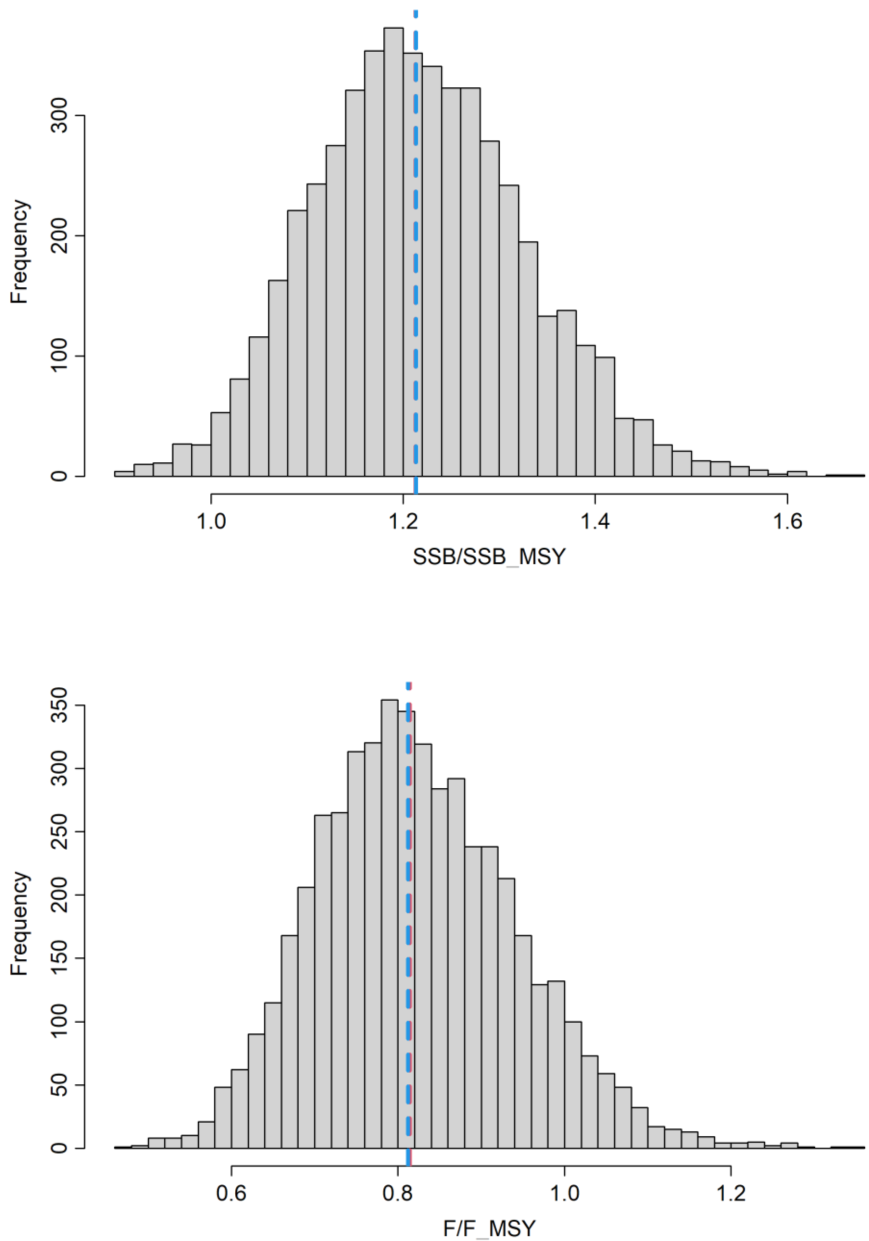



**Figure 24. Estimated spawning biomass in 2019 relative to MSY (SSB2019/SSBMSY, top panel) and estimated total fishing mortality in 2019 relative to MSY (F2019/FMSY, bottom panel) for the base case model configuration, dashed lines indicate the 50$^{th}$ quantile.**

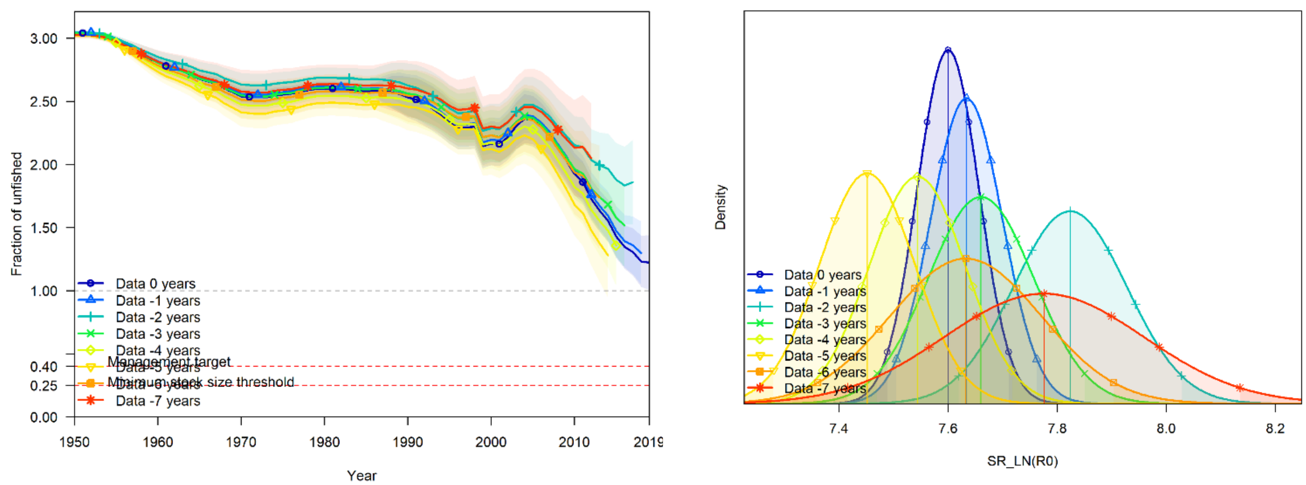



**Figure 25.** Estimated spawning biomass relative to MSY ($SB/SB_{msy}$, left panel) by year along with 95% asymptotic uncertainty (shaded areas) and the maximum likelihood estimate (MLE, vertical lines) and asymptotic uncertainty (bell shaped curves) of the natural log of virgin recruitment size (right panel) for each of the retrospective model runs conducted for the base case model configuration.

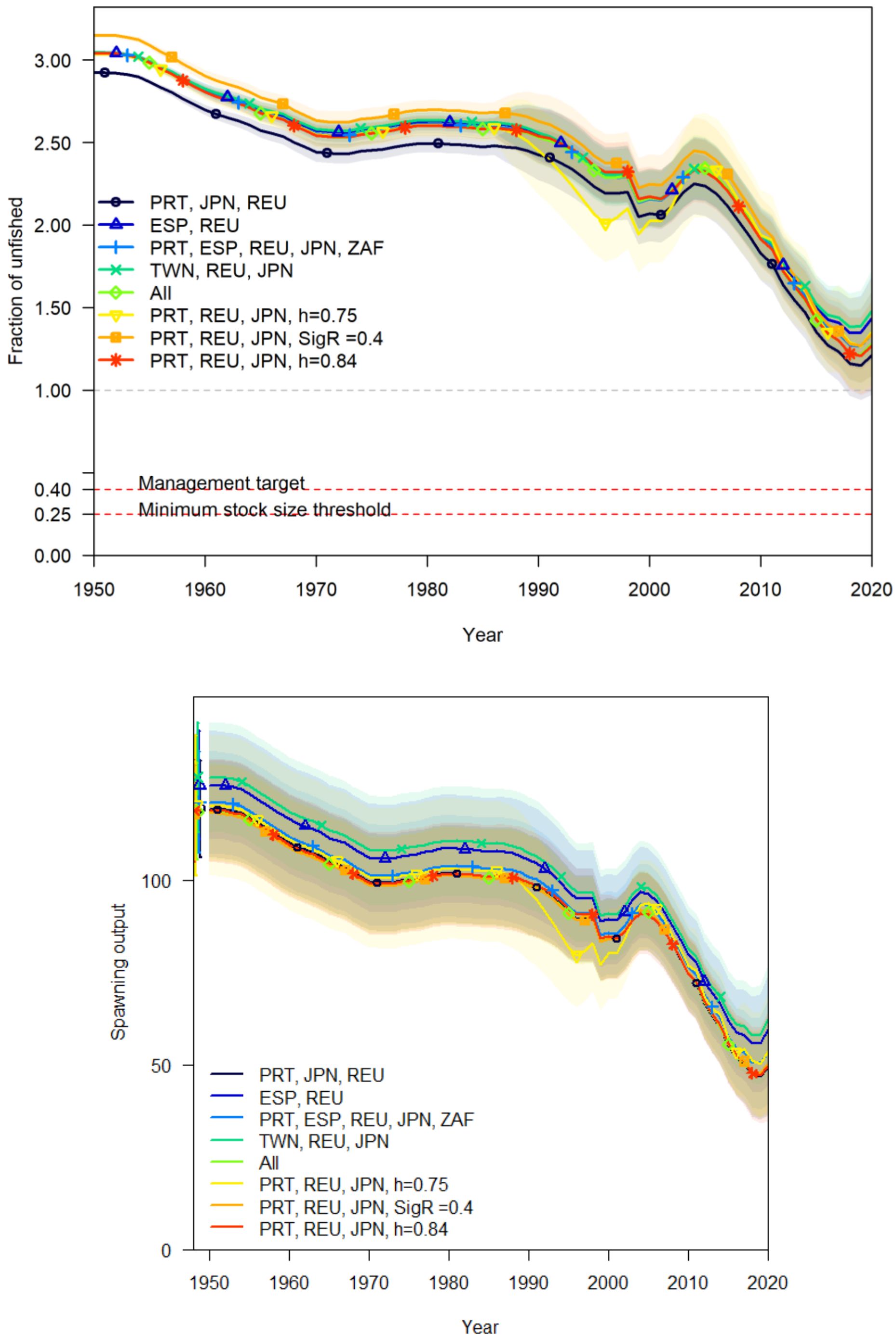



**Figure 26 Spawning biomass depletion for the sensitivity The top panel shows the depletion (from $B_{MSY}$) and the bottom panel shows the estimated spawning output.**

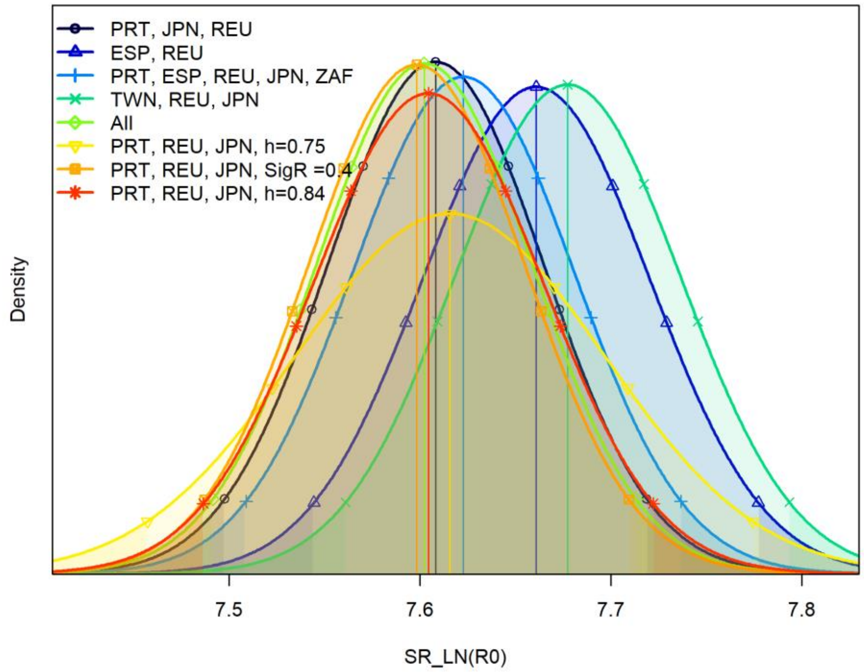



**Figure 27 Density estimates for the virgin spawning biomass from sensitivities using different CPUE series.**

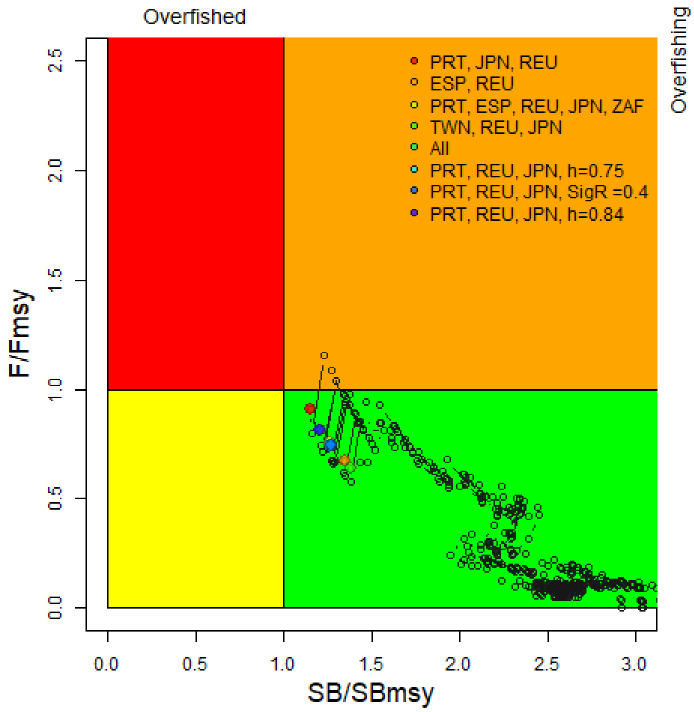



**Figure 28. Kobe plot showing the results the estimation of SB/SB$_{MSY}$ and F/F$_{MSY}$, for the terminal year of the model (2015).**

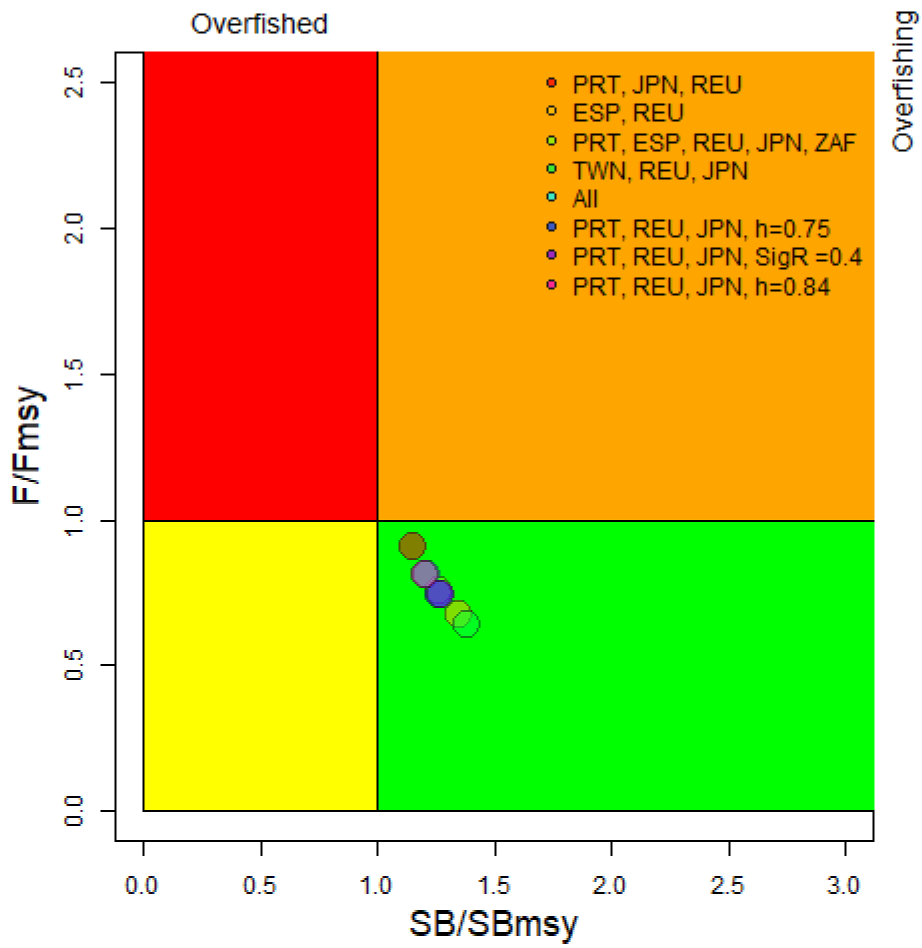



**Figure 29. Kobe plot showing the results the estimation of $B/B_{MSY}$ and $F/F_{MSY}$, for the terminal year of the model (2019) for the sensitivities using alternative groupings of CPUE series, steepness and sigmaR.**

# ANNEX 1

## Model Diagnostics

## Stock assessment of blue shark (*Prionace glauca)* in the Indian Ocean using SS3.

17th IOTC Working Party on Ecosystems and Bycatch
6 – 10 September 2021
Virtual Meeting

Joel Rice[2]

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

# 9 Annex 1. Diagnostic Tests for the Diagnostic Case SS3 Model

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

## 9.1 Goodness of fit

*Residual and Hierarchal cluster analysis*

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

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

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

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

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

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

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

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

*Runs test and joint residual plots.*

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

## 9.2 Model consistency

*R0 Profile*

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

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

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

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

### 9.3 Prediction Skill

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

## 10 Figures

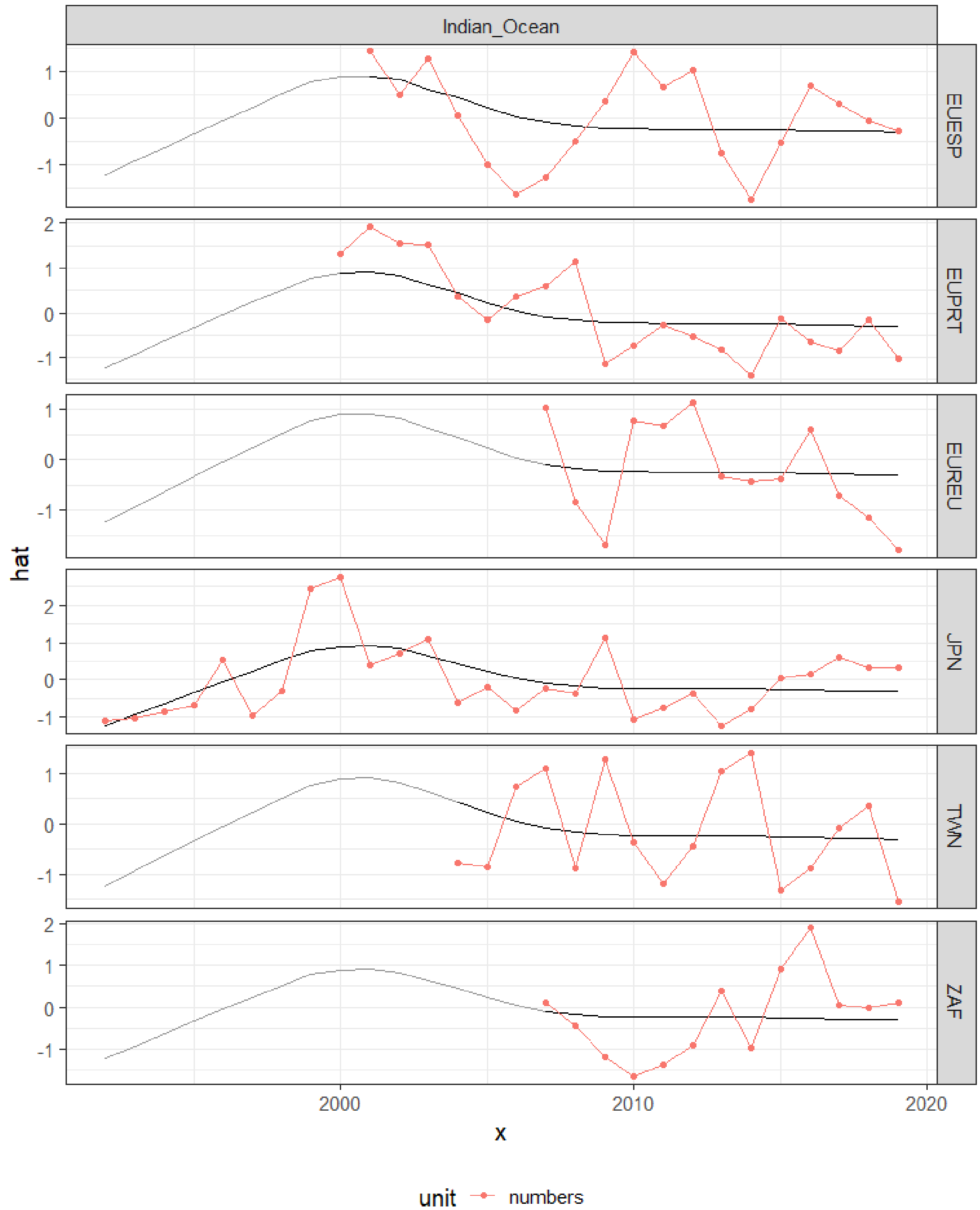



**Figure A1.** Indian Ocan time series of agreed CPUE indices; Points are the standardized values, continuous black lines are a lowess smoother showing the average trend by area (i.e. fitted to year for each area with series as a factor). X-axis is time, Y-axis are the scaled indices.

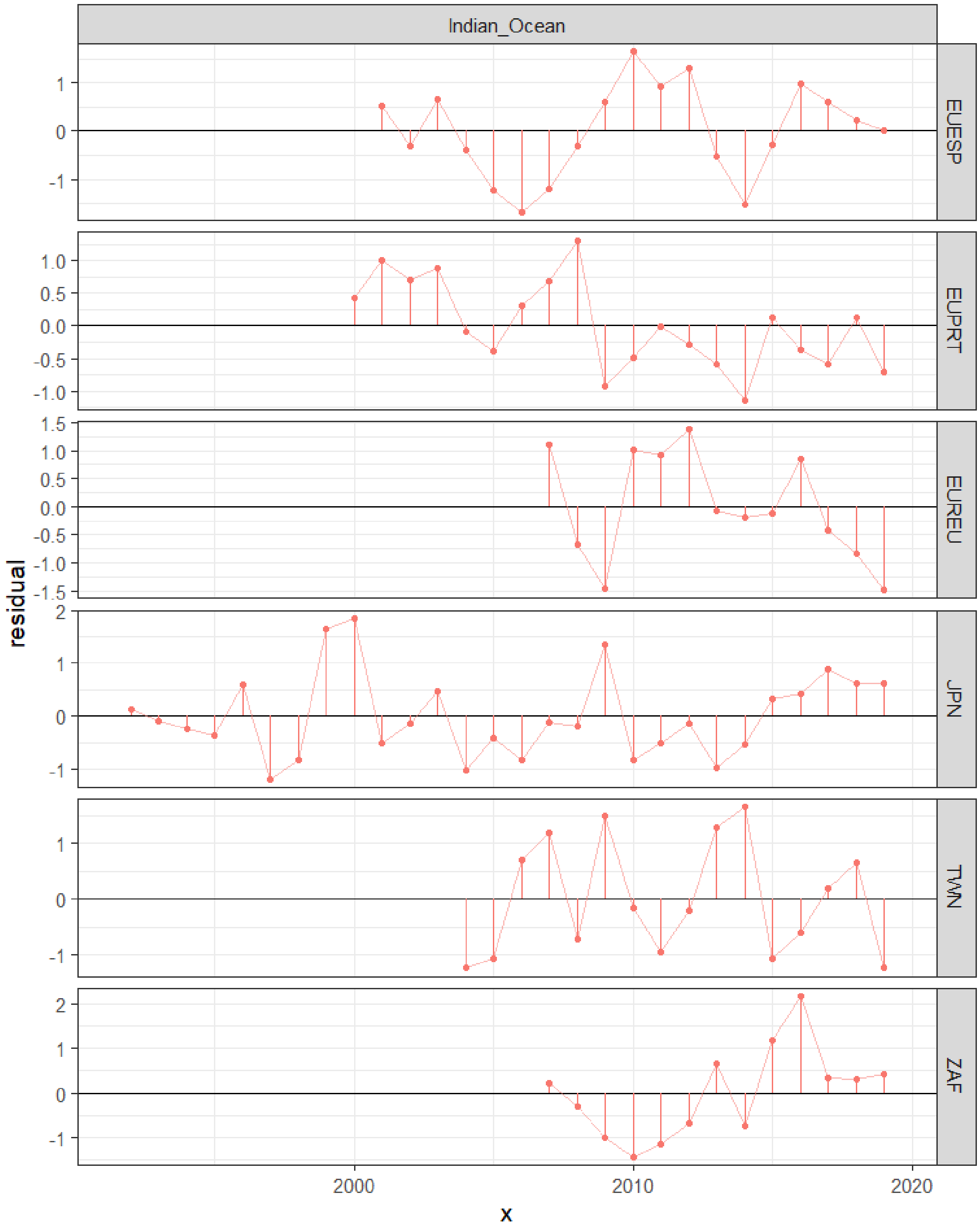



**Figure A2.** Time series of residuals from the smooth fit to CPUE indices. X-axis is time, Y-axis are the scaled indices.

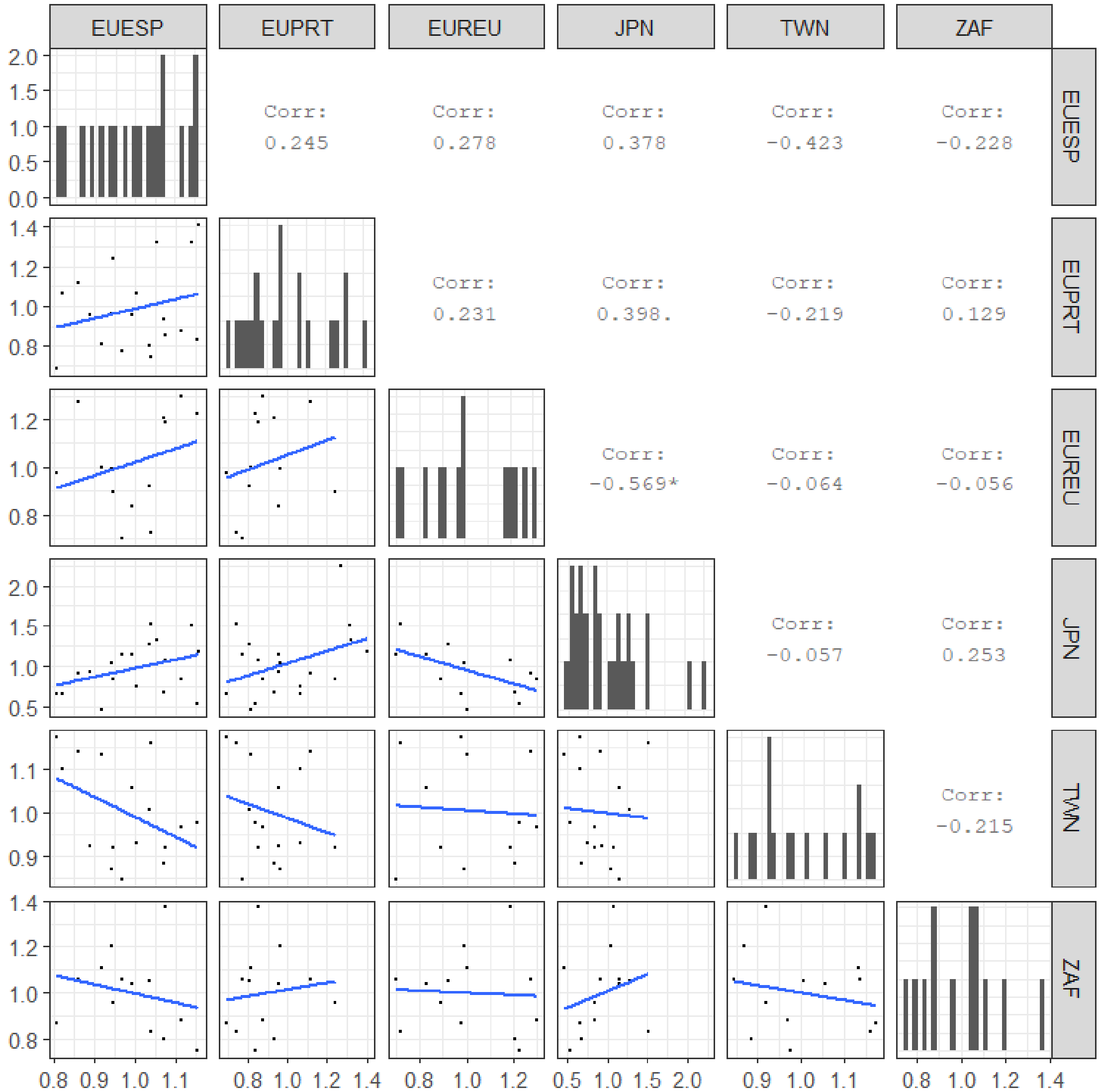



**Figure A3.** Pairwise scatter plots for CPUE indices. X- and Y-axis are scaled indices.

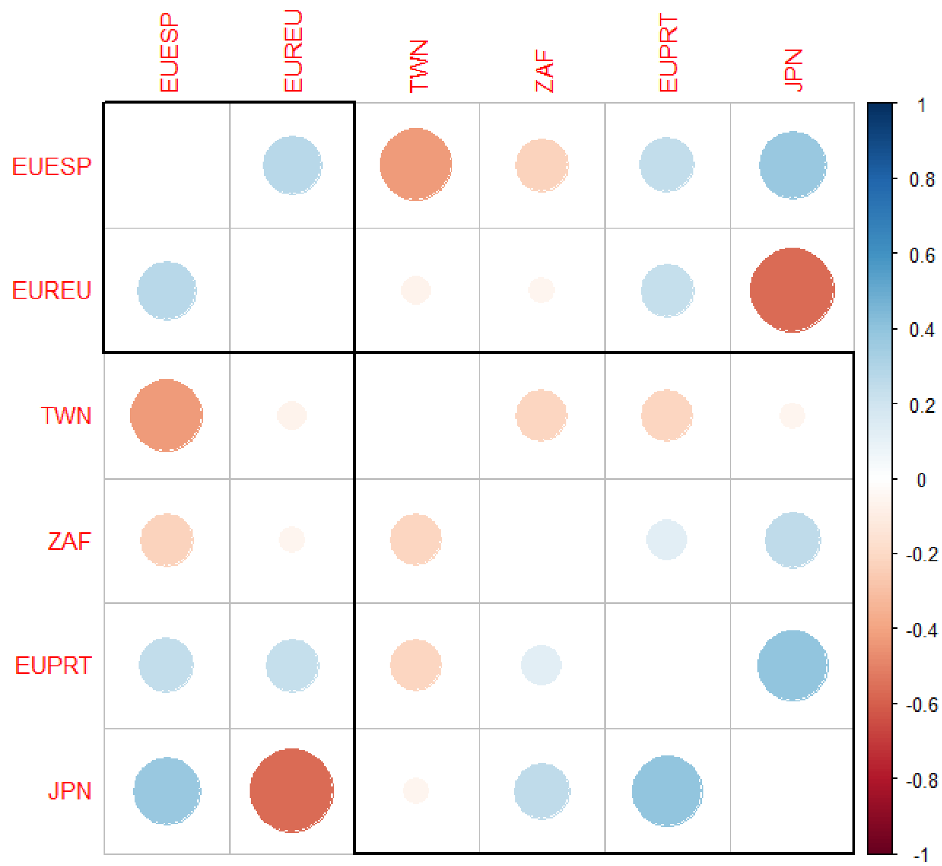



**Figure A4.** Correlation matrix for CPUE indices; blue indicates positive and red negative correlations, the order of the indices and the rectangular boxes are chosen based on a hierarchical cluster analysis using a set of dissimilarities.

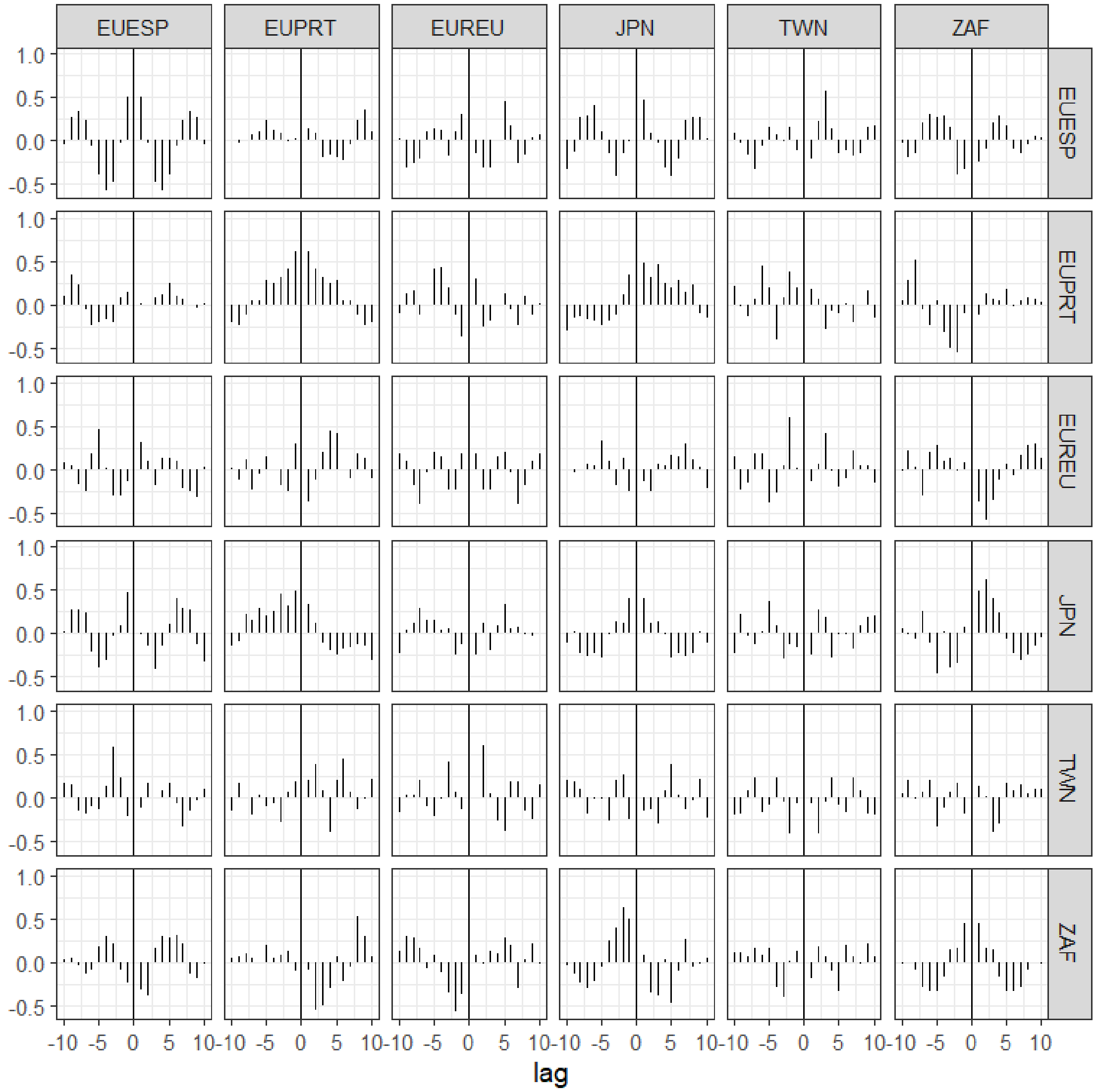



**Figure A5** Cross-correlations between CPUE indices to identify lagged correlations (e.g., due to year-class effects). X-axis is lag number, and y-axis is cross-correlation.

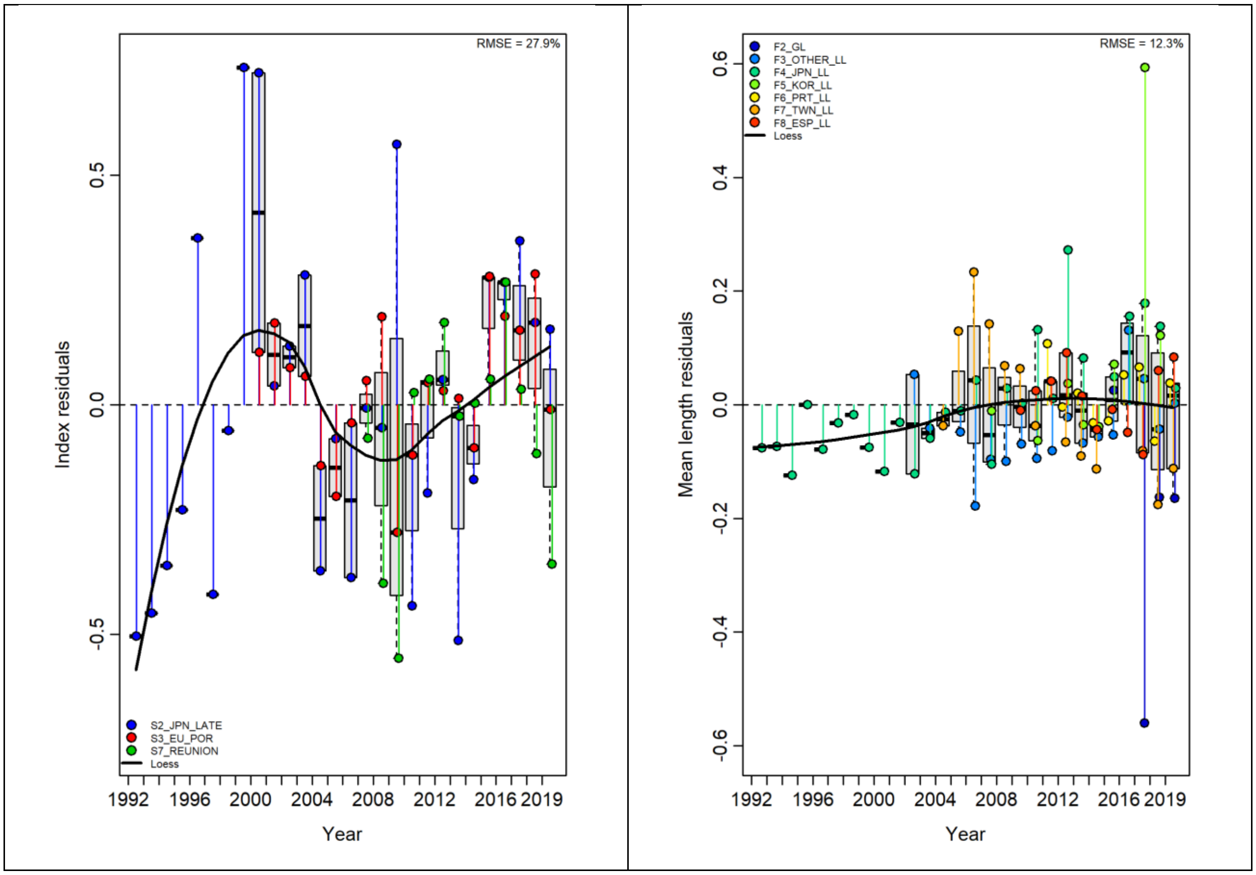



**Figure A6**. Joint residual plots for CPUE fits (left) from the Japanese (blue, S2), EU Portugal (Red, S3), and EU Reunion (Green S7) from the Indian Ocean blue shark Assessment,. And (c) annual mean length estimates for multiple fishing fleets. Vertical lines with points show the residuals (in colors by index), and solid black lines show loess smoother through all residuals. Boxplots indicate the median and quantiles in cases where residuals from the multiple indices are available for any given year. Root-mean squared errors (RMSE) are included in the upper right-hand corner of each plot.

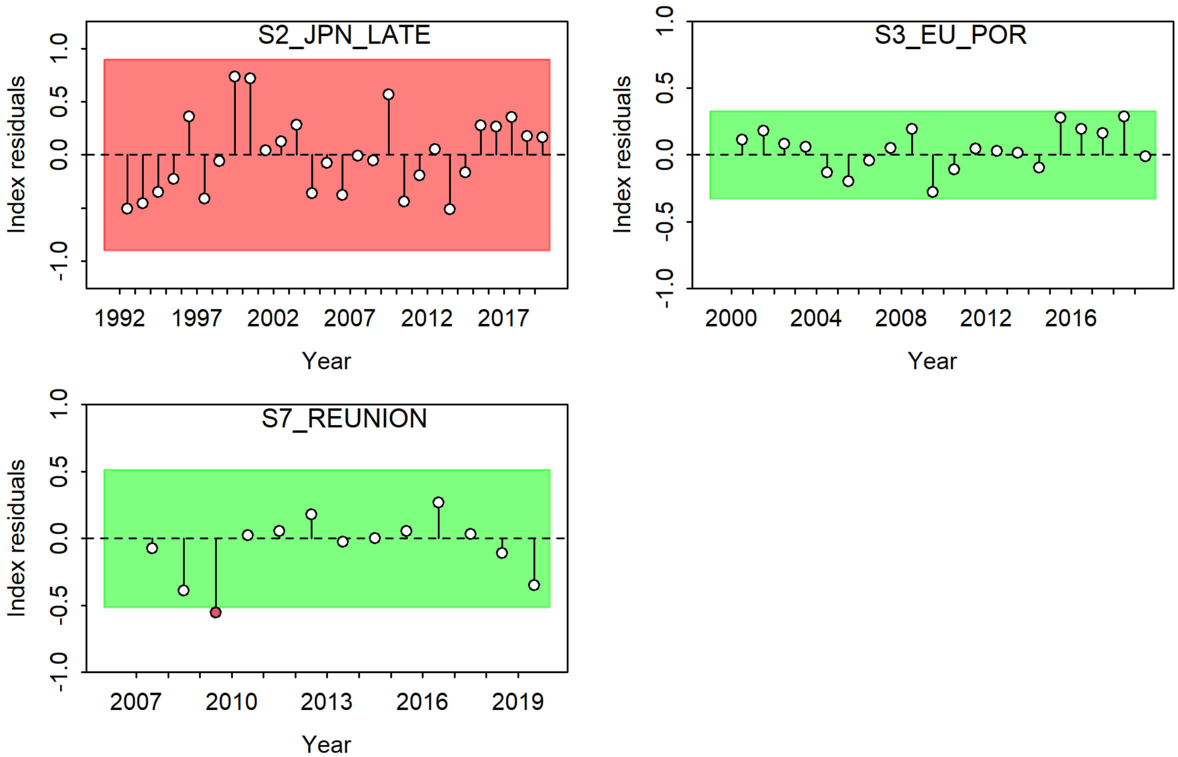



**Figure A7**. Runs tests results illustrated for three catch-per-unit-effort (CPUE) fits (top left pane Japanese LL, top right EU Portugal, bottom (Reunion) from the Indian Ocean SS3 blue shark (BSH). Green shading indicates no evidence ($p \geq 0.05$) and red shading evidence ($p < 0.05$) to reject the hypothesis of a randomly distributed time-series of residuals, respectively. The shaded (green/red) area spans three residual standard deviations to either side from zero, and the red points outside of the shading violate the 'three-sigma limit' for that series.

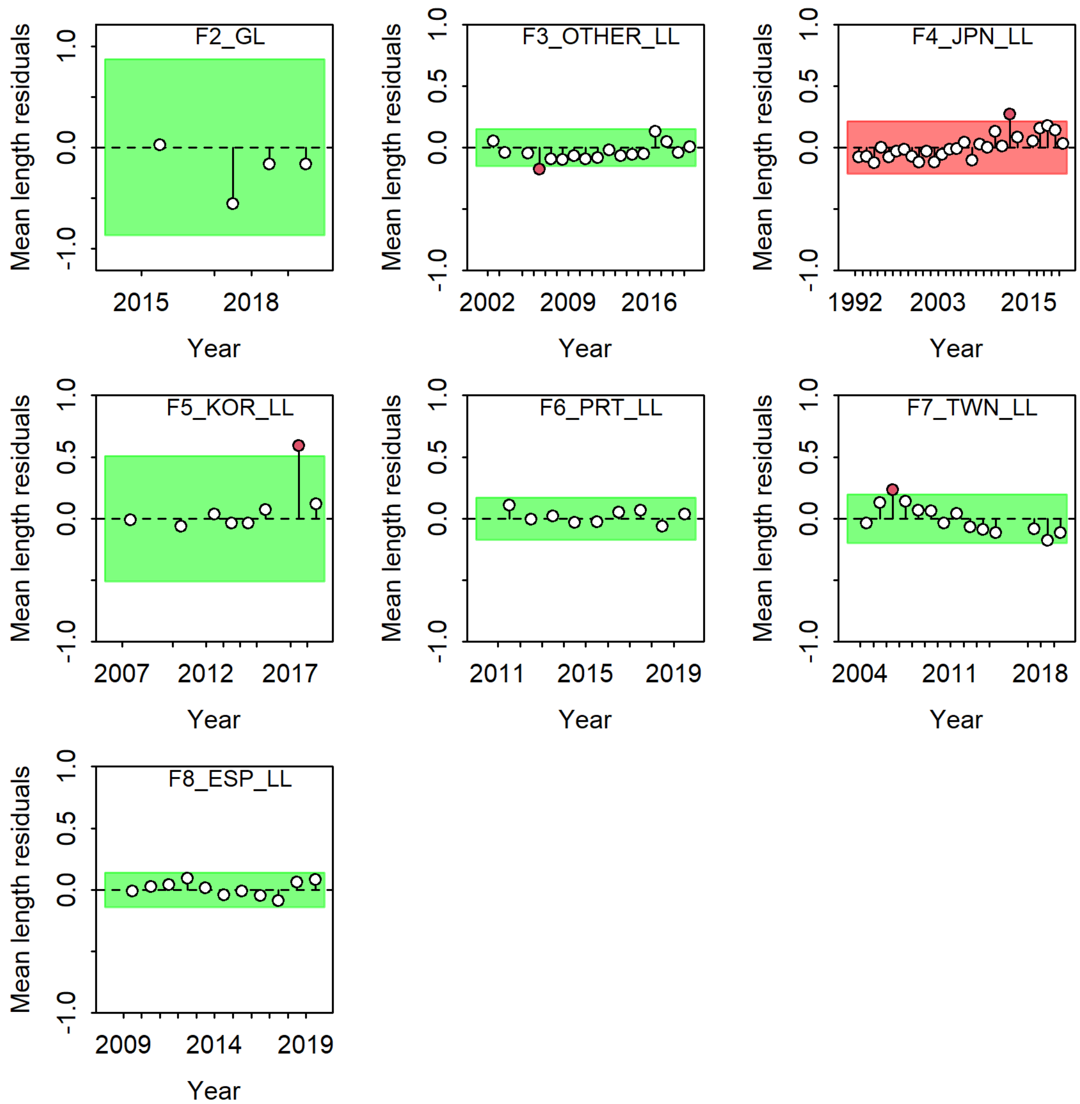



**Figure A8**. Runs tests results illustrated for the length composition fits (with names and fleet numbers on top) from the Indian Ocean SS3 blue shark (BSH). Green shading indicates no evidence ($p \geq 0.05$) and red shading evidence ($p < 0.05$) to reject the hypothesis of a randomly distributed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

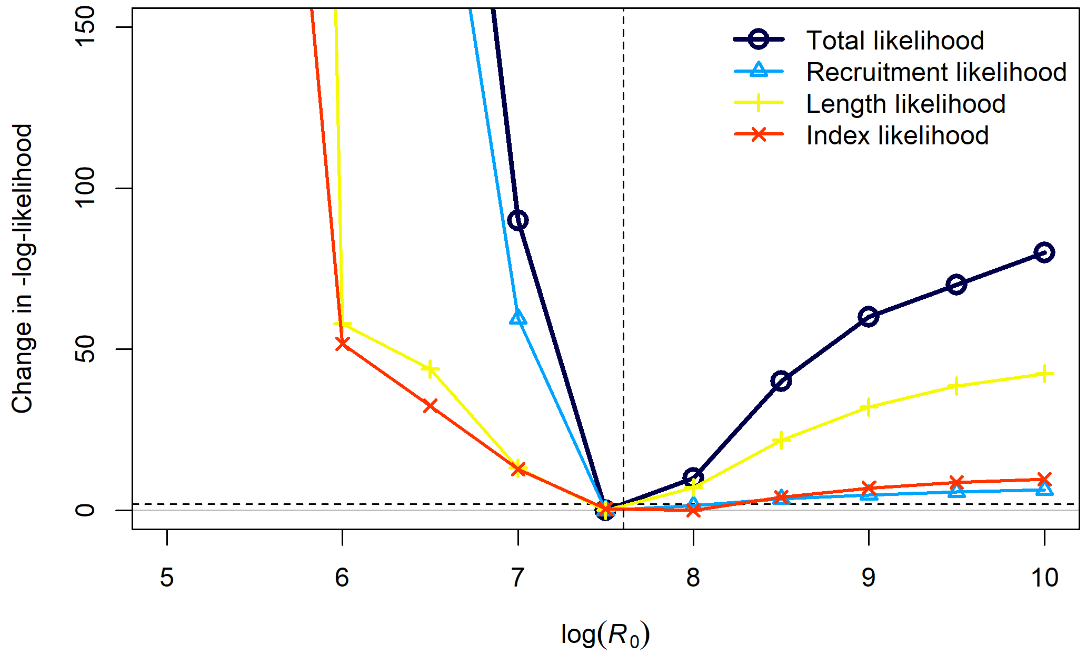



**Figure A9**. Total likelihood and component profiles (recruitment, length and index (CPUE) components).

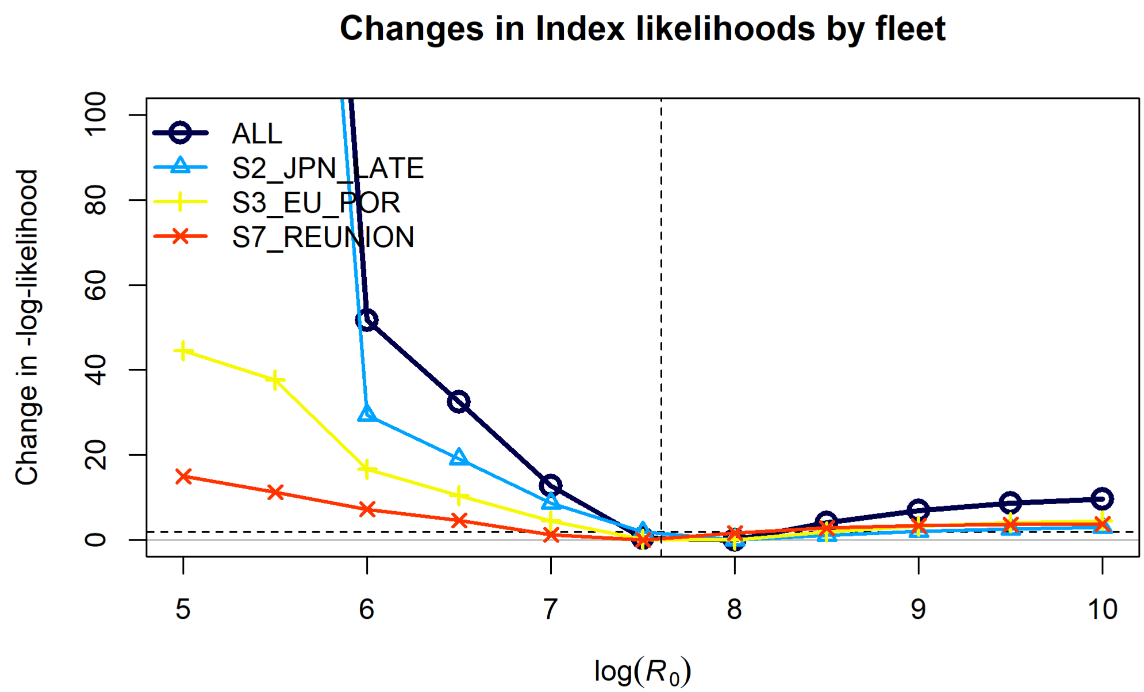



**Figure A10**. CPUE likelihoods for the diagnostic model.

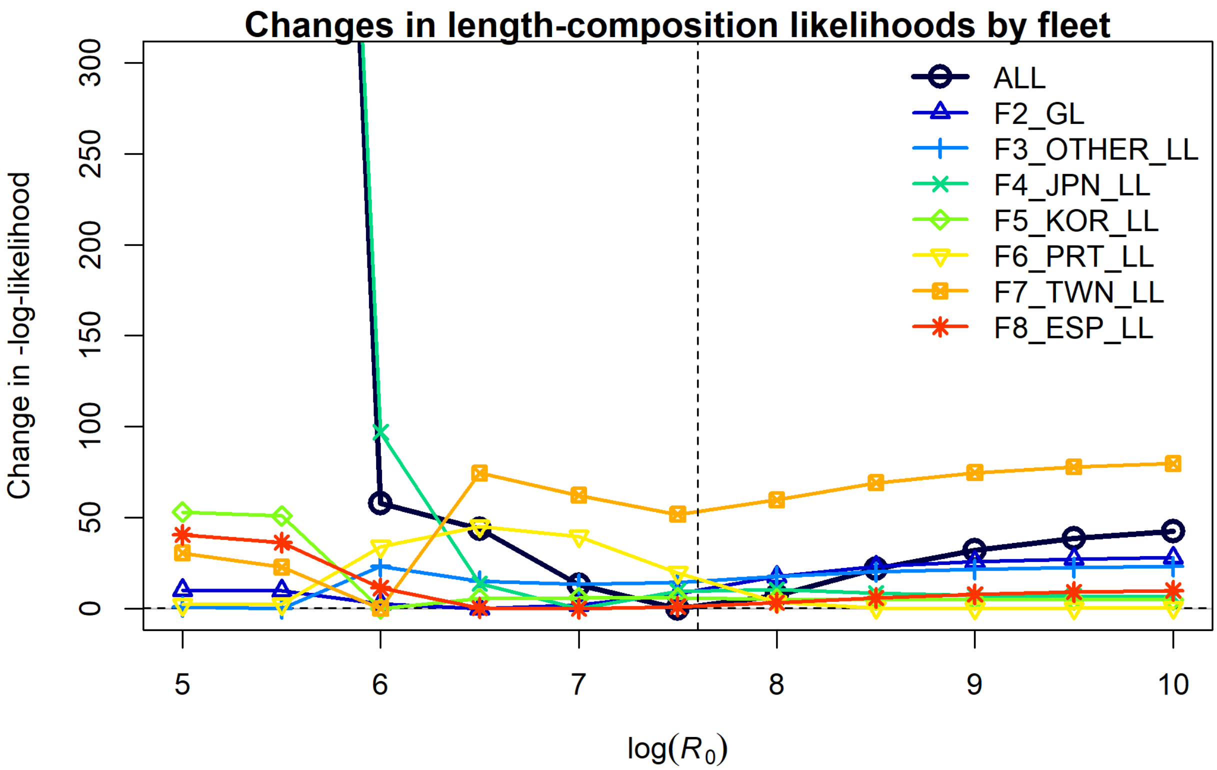



**Figure A11**.R0 profiles likelihoods for the fit to the length data.

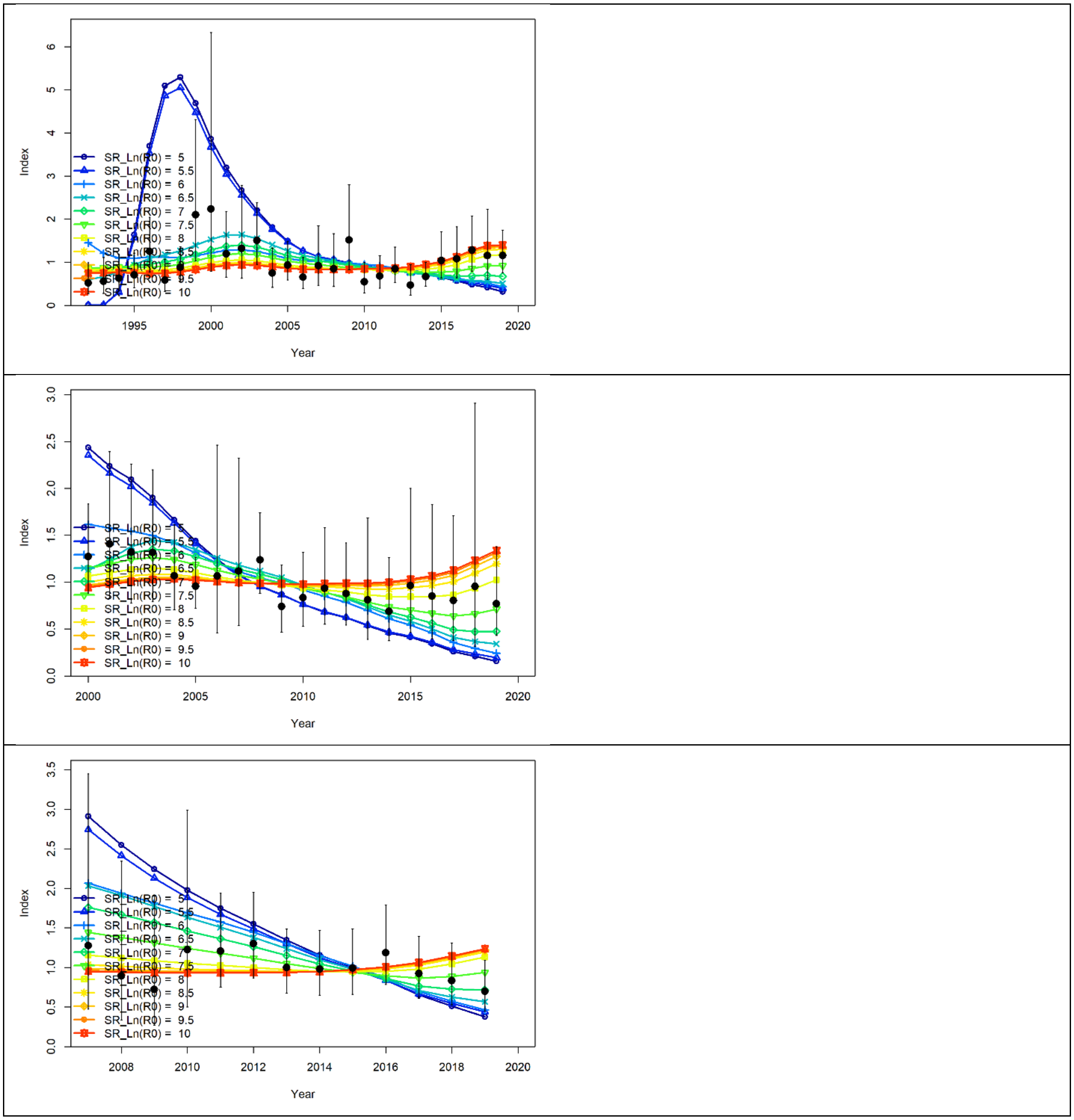



**Figure A12**. Fits to the component CPUE series (black dots and lines) for various profile likelihood values colored lines. Only the CPUE series that were fit in the diagnostic case are presented. The top panel is the fit to the Japanese CPUE, middle panel is the fit to the Portuguese CPUE and the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

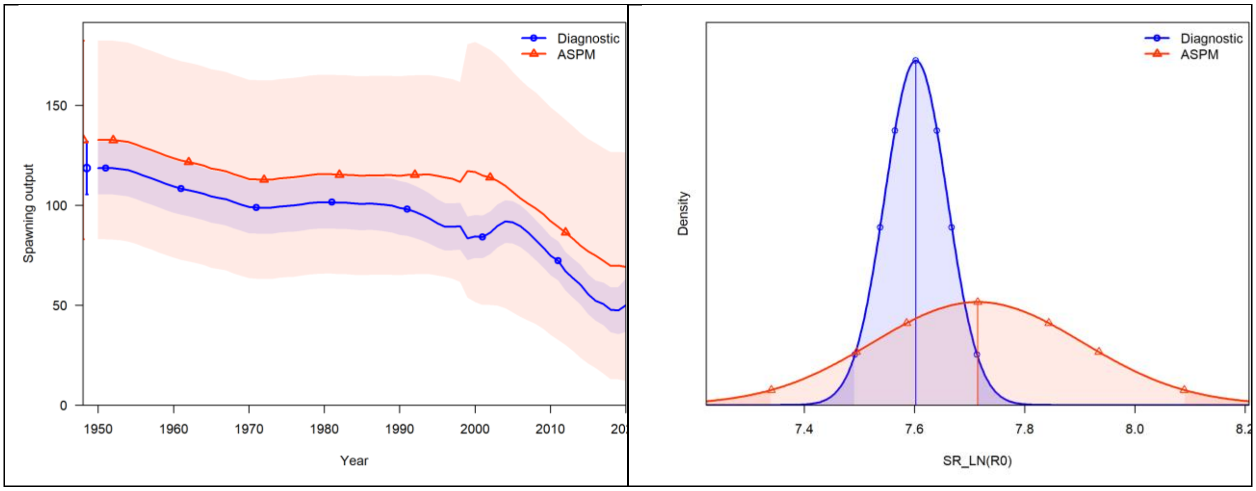



**Figure A13.** Comparison of spawning biomass trajectories for the ASPM and the diagnostic case of the assessment model carried out in stock synthesis (left panel), and the estimation of the densities for the stock recruitment parameter *LN(R0)* on the right hand side.

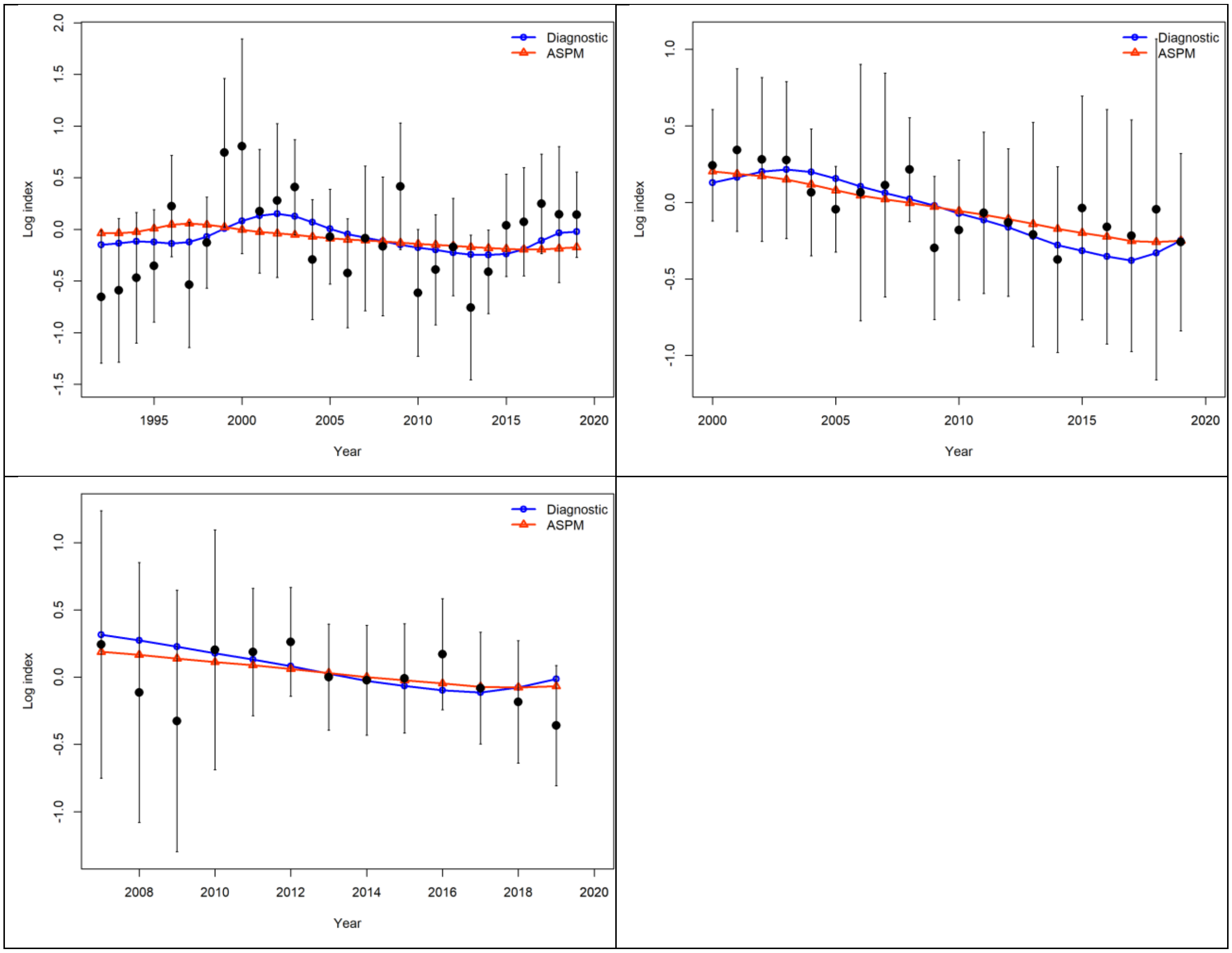



**Figure A14.** Fits to the indices for the ASPM and Diagnostic case on the log scale. The panels indicate the fits to the Japanese (top let), Portuguese (top right) and Reunion (bottom) CPUE series.

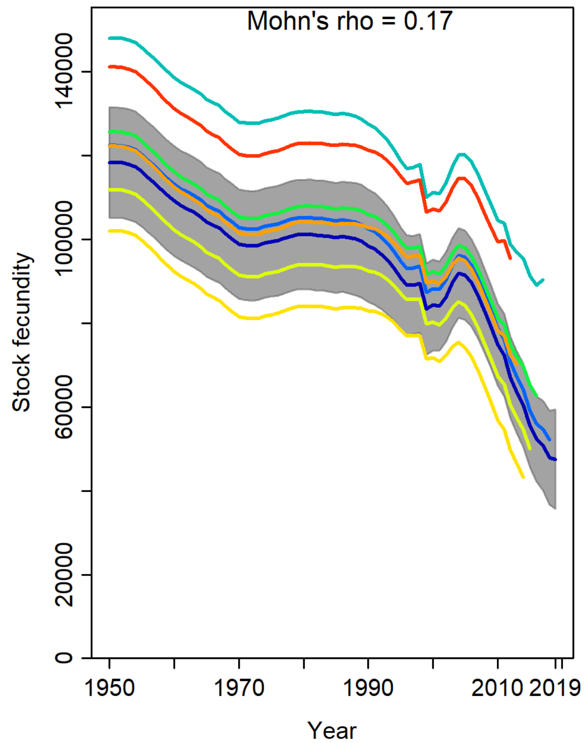



**Figure 15**. Retrospective analysis of spawning stock biomass (SSB) and fishing mortality estimates for Indian Ocean blue shark conducted by re-fitting the reference model (Ref) after seven years, one year at a time sequentially. Mohn’s rho statistic are denoted on top of the panels. Grey shaded areas are the 95 % confidence intervals from the reference model in cases where the analysis was run with Hessian.

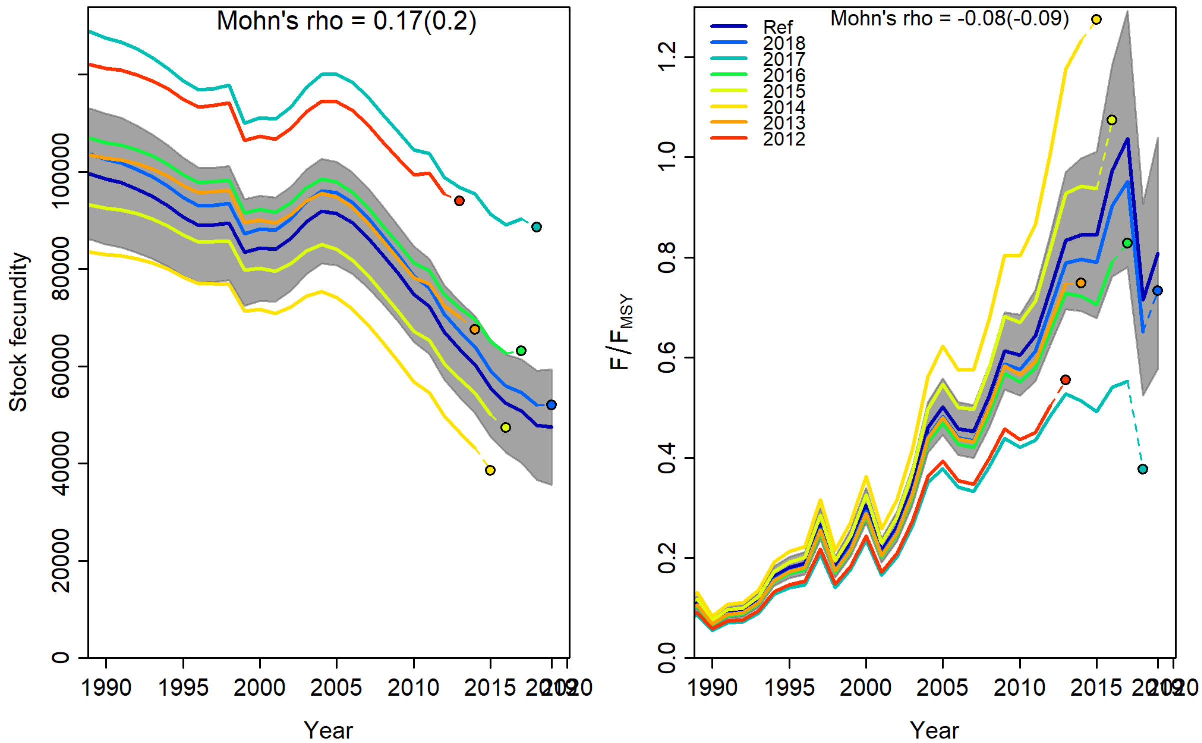



**Figure 16**. Retrospective results shown for the most recent years only. Mohn's rho statistic and the corresponding 'hindcast rho' values (in brackets) are printed at the top of the panels. One-year-ahead projections denoted by color-coded dashed lines with terminal points are shown for each model.

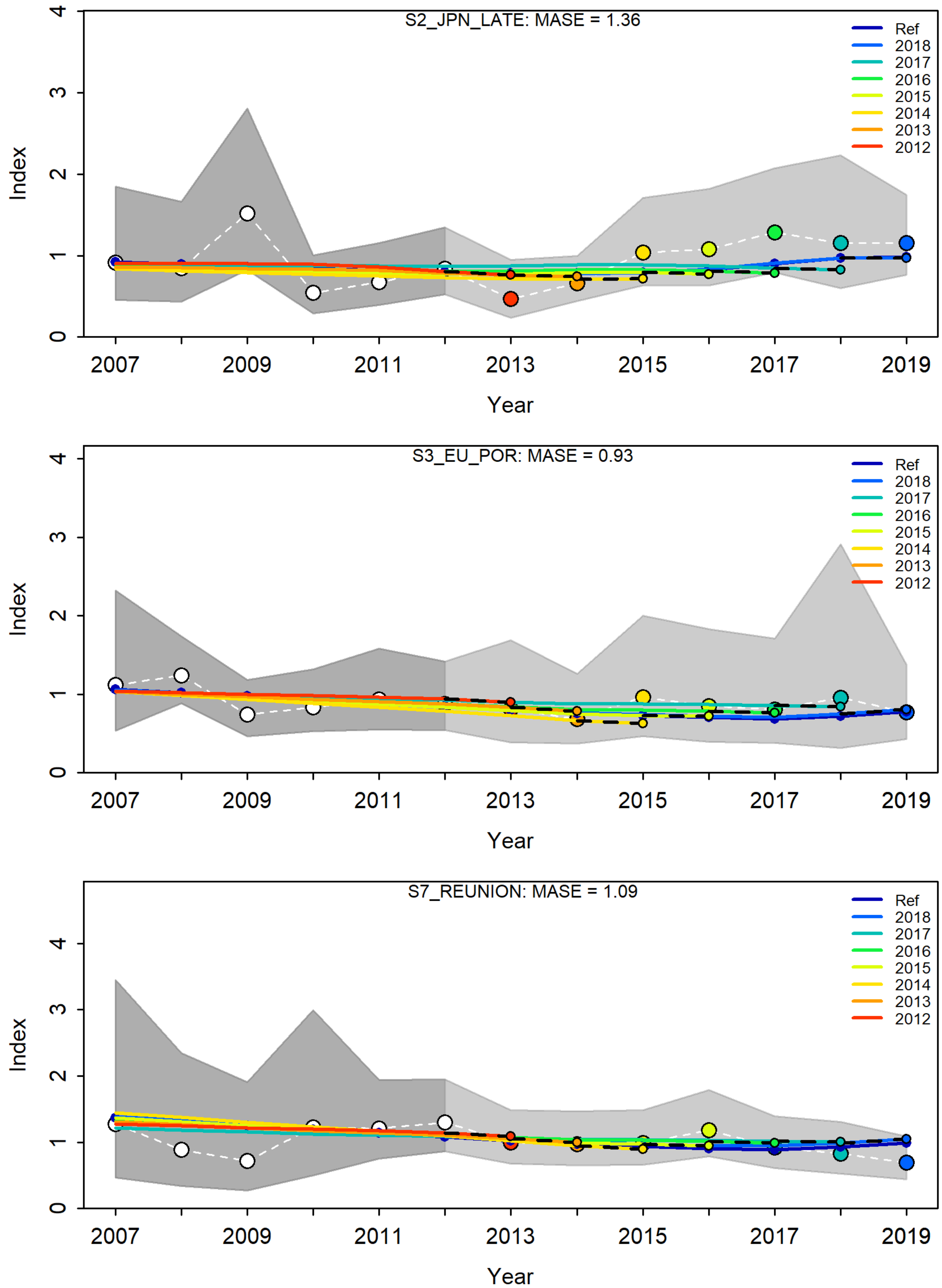



**Figure A17**. Hindcasting cross-validation (HCxval) results from CPUE fits, showing observed (large points connected with dashed line), fitted (solid lines) and one-year ahead forecast values (small terminal points). HCxval was performed using one reference model (Ref) and five hindcast model runs (solid lines) relative to the expected CPUE. The observations used for cross-validation are highlighted as color-coded solid circles with associated 95 % confidence intervals. The model reference year refers to the endpoints of each one-year-ahead forecast and the corresponding observation (i.e., year of retrospective + 1). The mean absolute scaled error (MASE) score associated with each CPUE.

## 11 References

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

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

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

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

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

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

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

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

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

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

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

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