IOTC-2017-WPEB13-23

APPROACHES TO THE RECONSTRUCTION OF CATCHES OF INDIAN OCEAN BLUE SHARK *(PRIONACE GLAUCA)*

IOTC Secretariat[1] (S. Martin and F. Fiorellato) and J. Rice[2]

ABSTRACT

Catch histories form an important component of stock assessments and so having a reliable and believable catch series is a key part in gauging the level of stock depletion. In data-limited situations, reported nominal catches are often not considered reliable and so reconstruction of catch histories plays an important role. The first Indian Ocean stock assessment of blue shark took place in 2015, however, due to the amount of uncertainty in the assessments, the conclusion regarding stock status remained as uncertain. The historic catch series was considered to be one of the key sources of uncertainty and so the Working Party requested that participants develop new approaches to reconstructing historic catches to be used as alternate series for assessment. This paper uses the available nominal catch data currently held in the IOTC database and explores the use of a disaggregation method followed by a ratio based method and a GAM statistical approach to reconstructing historic blue shark catches in the Indian Ocean.

The methods described in this paper attempt to account for two key sources of error in reported catches: (i) not reporting to species, and (ii) not reporting at all. A rule-based method to identify proxy fleets was used to disaggregate reports of 'sharks NEI' to address the limited reporting to species level, while ratio and GAM based models using target catches were used to predict the expected catches where there are zero reported catches. The ratio based method was based on the disaggregated catches while the GAM method was based on the IOTC nominal catches. The two resulting estimated catch series were very similar with catches increasing over the time period of the fishery, reaching approximately 50-60,000 t in recent years. However the GAM series produced higher estimated catches in early years and was still increasing at the end of the time period (2015) while the ratio estimates based on the disaggregated catches followed the disaggregated catch trend more closely and peaked in 2011. While a range of approaches have been explored, if a preferred catch series is to be used as an alternative series for the assessment, then it is recommended that the GAM estimated catch is used.

KEYWORDS: Catch reconstruction, catch estimation, catch history, data-limited stocks, nominal catch, blue shark, stock assessment.

[1] IOTC Secretariat, PO Box 1011, Victoria, Seychelles. Sarah Martin; email: sarah.martin@fao.org; Fabio Fiorellato: email: Fabio.fiorellato@fao.org.

[2]email: joelrice@uw.edu

1 Introduction

Catch histories form an important component of stock assessments and so having a reliable and believable catch series is a key part in developing a good estimate of the level of stock depletion. In data-limited situations, reported nominal catches are often not considered reliable and so reconstruction of catch histories plays an important role. This is particularly important for bycatch species where data are often sparse and of varying quality. Nominal catches of blue sharks in the Indian Ocean held by the IOTC[3] are considered to be highly uncertain, and are likely to be 'severe underestimates' of the actual catches taken as concluded by the Working Party on Ecosystem and Bycatch in 2015.

The first Indian Ocean stock assessment of blue shark took place in 2015, however, due to the amount of uncertainty in the assessments, the conclusion regarding stock status remained as uncertain[4]. The historic catch series was considered to be one of the key sources of uncertainty and so the Working Party requested that participants develop new approaches to reconstructing historic catches to be used as alternate series for assessment. There a number of approaches that may be used to produce catch history reconstructions. One method that has been used previously for Indian Ocean blue shark was based on information obtained from the shark fin trade, providing estimates used in the 2015 assessment[5] that were approximately four times higher than the IOTC nominal catches[6]. Another method has been developed which is based on expert knowledge of Indian Ocean fisheries to determine catch rates of sharks to target species and separating out the different shark species using a proportioning method[7]. Yet another approach that has been applied for southern bluefin tuna in the southern Ocean involved the use of random forests to predict CPUE of non-members based on the reported CPUE of members[8].

This paper uses the available nominal catch data currently held in the IOTC database and explores the use of a ratio based method and a GAM statistical approach to reconstructing historic blue shark catches in the Indian Ocean.

---

[3] IOTC Nominal catches: IOTC-2017-WPEB13-DATA03. www.iotc.org/meetings/13th-working-party-ecosystems-and-bycatch-wpeb13

[4] IOTC, 2015. Report of the 11th Session of the IOTC Working Party on Ecosystems and Bycatch. Olhão, Portugal, 7-11 September 2015.

[5] 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.

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

[7] Murua H., Santos, M.N., Chavance, P., Amande, J., Seret, B., Poisson, F., Ariz, J., Abascal, F.J., Bach, P., Coelho, R., Korta, M. 2013b. EU project for the Provision of Scientific Advice for the Purpose of the implementation of the EUPOA sharks: a brief overview of the results for Indian Ocean. 9th Working Party on Ecosystems and Bycatch, 12-16 September, La Reunion, French Overseas Territories. (IOTC Doc: IOTC–2013–WPEB09–19).

[8] Chambers, M. and Hoyle, S. 2015. Proposed approach to estimate non-member catch of SBT using ransom forests to model CPUE. CCSBT/CPUE2015/04

2 Methods

Data sources used: IOTC nominal catches

The best estimates of nominal catches of blue shark in the Indian Ocean are published annually on the IOTC website[9]. These are based on catches reported directly to IOTC both contracting and non-contracting parties fishing for tunas in the Indian Ocean and include best estimates in some cases where data are particularly poor or lacking altogether.

This data is available by flag state, species (including IOTC species and bycatch), fishing gear and area (east or west Indian Ocean) in live weight equivalent. The data set extends back to the 1950s when industrial longlining began in the Indian Ocean. The data are generally considered representative (though the level of accuracy varies by year) of the nominal catch of the main IOTC target species, however, the reporting of sharks over the time period has been somewhat more inconsistent.

The nominal catch dataset for blue shark and the main amendments to reported catches that have been made have been fully described (IOTC Secretariat, 2016[10]). The majority of nominal blue shark catches are taken by the Indonesian fleet (Figure 1) and catches are dominated by three major gear types: longline, gillnet and handline (Figure *2*). The Indonesian gillnet fleet is responsible for most of the historic catches of blue shark, followed by a transition to coastal longlines in the mid-1980s. In more recent years catches taken by the industrial longline fisheries have expanded, predominantly by the swordfish targeting longliners of EU-Spain and EU-Portugal, the deep-freezing longliners of Japan and Taiwan,China and the fresh longliners of Taiwan, China (Figure *3*).

A key issue with this dataset is the presence of the large *"Sharks various nei"* (SKH) category in the database which is assumed to include unidentified blue sharks. However, the extent to which these aggregates comprise blue sharks relative to other shark species is unknown. Another major issue is the apparent many incidences of 'missing' catch. For example two fleets fishing in the same vicinity catching the same target species using the same gear type but only one reports any catch of (blue) sharks. This is likely a reporting issue. A third key issue is inaccurate reporting, e.g., a fleet catches substantial quantities of blue shark and only reports a small fraction of this. The methods descried below aim to address these core problems with the dataset through a range of approaches explored to reconstruct historic blue shark catches.

Disaggregation of unidentified shark catches

The ***Nominal Catch Disaggregation*** process is the deterministic, non-linear process of breaking down all nominal catch records referring to either a gear or a species aggregate (or both) into records referring to single species and gears[11].

[9] IOTC Nominal catches: IOTC-2017-WPEB13-DATA03. www.iotc.org/meetings/13th-working-party-ecosystems-and-bycatch-wpeb13

[10] IOTC Secretariat, 2016. Blue Shark catches reported to the IOTC Secretariat and a review of current estimation procedures. IOTC-2016-WPEB12-INF04.

[11] IOTC Secretariat, 2016. Improving the core IOTC data management processes. IOTC-2016-WPDCS12-25_Rev1

This process is routinely used to provide the reference nominal catch data used as basis for the stock assessments performed during the WPTT, WPB, WPNT and WPTmT working parties. So far, it has only been applied to IOTC species and this is the first time that its application has been attempted outside this context.

The key concept behind the disaggregation process is that catch quantities for aggregated records should be assigned to the different combinations of species / gears that the aggregates are considered to comprise: therefore, in order to identify single target species and gear combinations and proportionally assign fractions of the original aggregated catches to said combinations, the disaggregation process applies a sequence of multiple *disaggregation procedures* to identify relevant *proxy records* within the original Nominal Catch dataset. Once proxy records have been identified, the proportion of catches by species and gears available for these records is eventually used to assign the original aggregated catches to its single components.

The disaggregation procedures identify proxy records by filtering the original dataset by *fleet*, *type of operation* (*Artisanal* / *Semi-Industrial* / *Industrial*), *region*, *area* and *timeframe*: for this reason, they rely on a specific configuration table that assigns – to each combination of fleet, gear and area for which at least one nominal catch record exists – a *region* of most-likely operation (see Table 1). A pseudo-code implementation of the disaggregation process is available in Table 1 of the Appendix.

Currently, the process implements eight different disaggregation procedures that are complemented by a ninth procedure (*manual disaggregation*) triggered when no proxy record can be identified by any of the other procedures. The definition of the proxying criteria adopted by the current disaggregation procedures is provided in Table 2.

In this exercise, the disaggregation process has been applied to all records reporting catches for SKH (*Sharks various nei)*, whose aggregation is currently defined as the combination of the following shark species (entries marked with * represent other species aggregates and are eventually further broken down by the disaggregation process into any of their components):

- *Hammerhead sharks nei**
    - Scalloped hammerhead
    - Smooth hammerhead
- *Mako sharks**
    - Longfin mako
    - Shortfin mako
- *Thresher sharks nei**
    - Thresher Shark
    - Bigeye thresher
    - Pelagic Thresher Shark
- *Blue shark*
- *Silky shark*
- *Oceanic whitetip shark*
- *Porbeagle*
- *Crocodile shark*
- *Great hammerhead shark*

- *Scalloped hammerhead*
- *Smooth hammerhead*

Given this definition of all component species for the SKH species aggregate, the disaggregation process is expected to produce – for each catch record originally referring to such aggregate – one or more records of catches for any of the species above. The actual result of all SKH catches disaggregation, limited to blue shark catches as produced by the disaggregation process, is eventually used to provide the *'Estimated – disaggregated'* quantities reported in Table 3. In practice, given the possible component species for the *unidentified shark species* aggregate, the disaggregation process may also be producing catch records for the other species in the list.

Table 4 provides a full breakdown of how all SKH catches have been disaggregated over the entire time series. Catches of shark species accounting for less than 10,000 t in total have been all reported under the '*OTH – Other sharks species*' category. Final disaggregation results do still include catches for '*SKH – Sharks various nei*' for all those years for which the disaggregation process could not identify any proxy records to break down the aggregated catches: these remaining aggregated catches could subsequently be manually broken down into their component species, however, for this exercise they have been left as originally reported.

Once the fraction of SKH catches by year assigned to blue sharks was determined, the original (i.e. nominal) blue shark catches were updated with the additional quantities resulting from the disaggregation, producing the time series reported in Table 3 under the *Estimated – disaggregated* column.

Ratio method to estimate unreported blue shark catches

A second method based on the ratio of blue shark to target species was used in an attempt to estimate the unreported component of blue shark catches. Target species were defined as yellowfin tuna, bigeye tuna, skipjack tuna, albacore and swordfish. Nominal catches of these species are considered to be relatively accurate.

Starting from the blue shark nominal catches plus the blue shark component of the disaggregated catches, records were separated out into four components where fleets were reporting:

1) Positive catches of target species and positive catches of blue shark where the target species catch is greater than the blue shark catch (used to calculate catch rate)
2) Positive catches of target species but zero blue shark catches (assumed to be non-reporting so were not included in the catch rate calculation)
3) Positive catches of blue shark but zero target species catches or positive catches of target species and positive catches of blue shark where the blue shark catch is greater than the target species catch (it is assumed here that blue sharks are actually the target species in this case and so the reporting is likely to be accurate, hence these records were excluded from the catch rate calculation)
4) Zero catches of both target species and blue sharks reported (these records were not used)

Blue shark catch rates were calculated, defined as the ratio of blue shark to the total target species catch where positive catches of target species and blue shark were caught and where the target species catches were greater than the blue shark catches. These catch rates were calculated by fleet, year and gear type (the finest scale gear classifications stored in the IOTC database). Catch rates were averaged across all fleets reporting blue shark catches for each gear-year combination (Figure *4*). Fleets reporting zero catches of blue sharks for a year-gear combination where other fleets were reporting positive blue shark catches were assumed to be false zeros and so were not used in calculating the average, while records where catches of blue shark were greater than the target species catches were also not used as in these cases, the blue shark was assumed to be the target species and should be more accurately reported. Unclassified gear types were removed to avoid meaningless predictions from unrelated gear types.

These ratios were then used to estimate the unreported blue shark catch component (defined as fleets reporting zero catches of blue sharks for a year-gear combination where other fleets were reporting positive blue shark catches). Fleets reporting zero blue shark catches were allocated catches by multiplying the average catch rate by the target catch for the fleet.

As a second step in the process, a moving average was used to smooth the catch rates by gear type over time. This was explored for a number of different years (3,5,8) but had little impact on the final predictions and was not used in the final estimation.

GAM approach to estimate unreported blue shark catches

A second method was used to attempt to estimate blue shark catches based on the nominal catches in the IOTC database. A statistical modelling approach based on generalized additive models (GAMs) was used to predict unreported catches. The model was set up incorporating a number of explanatory variables thought to be influential in determining whether a fleet catches blue sharks. The model was parameterised based on the records where reported blue shark and the selected covariates were available and the model was run on the remaining dataset where zero blue shark catches were reported, and where sufficient levels of the covariates were available for prediction. Records with levels outside the model, and so for which prediction was not possible, were dropped.

The log transformed nominal blue shark catches were used as the response variable. A filter was applied to remove extremely high catch rates by selecting only those records where catches of blue shark were less than 80% of the total catches of non-shark species. This was performed to remove those high values where the fishery is likely to be targeting blue sharks and therefore more likely to be accurately reporting those sharks. Outliers were not well predicted by the model so the dataset on which to predict the unreported blue shark catches was also filtered to remove extreme values (records where target catches >80,000 t) which had a disproportionately large effect on the results. This resulted in the removal of 77 outliers which was 1.06% of the data set.

The explanatory variables year, target species catch, gear, area (E/W) and fishing ground (coastal, pelagic or all). Different classifications of non-blue shark species were also explored including separate covariates for temperate tuna species, tropical tunas, other shark species and all other species, added using splines. To avoid over-parameterisation, models were run sequentially starting from the simplest model and incorporating covariates and interactions, where they made sense theoretically (e.g. area-gear interactions) in an iterative manner. Models were evaluated based on AIC values.

For this method, including too many factor levels can result in predictions which are made from non-significant levels or factor levels that are not present in the information for the non-reporting fleets, making predictions from these levels difficult. Therefore, to reduce the number of factor levels in the model, regression trees were used to partition the predictor space for the variables *gear* and *area* into groupings that have the most homogeneous response to the prediction variable. For this exercise the regression tree fit the mean response in each group assuming normally distributed errors based on the formula log(BSH_catch) ~ variable, where variable is either area or gear. These models were considered minimally sufficient to reduce the number of levels of each covariate. The covariate *gear* originally had 13 levels and was reduced to 5 levels and *area* originally had 17 levels and was reduced to 5 (Table 5) (Figure *5*).

3 Analysis of results

Disaggregation of aggregate shark catches

Figure *6* shows the results of the disaggregation of unidentified shark catches; catches identified as blue sharks and the addition of these to the nominal blue shark catches. The third column in Table 2 provides the final estimated catch figures derived from the disaggregation process (disaggregated catches allocated to blue sharks + nominal blue shark catches).

Table 5 shows the distribution of SKH catches among the component species resulting from the Nominal Catch Disaggregation process. An interesting result of this exercise is that the majority of SKH catches over the entire time series is assigned to Silky shark (2,049,006 t), with Blue shark coming third (247,968 t) and receiving a fraction of the SKH catches roughly at the same level of Smooth hammerhead (259,275 t).

This result is not particularly surprising as the majority of SKH catches in the IOTC Nominal Catch database are recorded (as reported) under *Artisanal* or *Semi-Industrial* gears in the Western Indian Ocean (Figure *7* and Figure *8*). Based on the way in which the Nominal Catch Disaggregation process is defined, aggregate catches are broken down by species using proxy records referring to gears of the same type and in the same Indian Ocean area. The fact that the majority of nominal blue shark catches are recorded in the Eastern Indian Ocean (Figure *9*) and under *industrial* gears will *de facto* prevent the Nominal Catch disaggregation process from assigning those *non-industrial* SKH catches from the Western Indian Ocean to blue shark.

Furthermore, the Nominal Catch Disaggregation process is also very sensitive to the *region* of operation of any given fleet (Table 1) and this approach will, for certain proxy fleet / gear combinations known for fishing mainly one specific component of the species aggregates in a given region, result in a strong predominance of that component species in the disaggregated results.

Figure *10* presents the breakdown of the total disaggregated estimated blue shark catches by gear and by fleet. The allocation of unidentified sharks to the blue shark series increases the catches in early years, primarily based on the Australian gillnet fleet in the 1970s and early 1990s and catches by Bangladesh in the late 2000s which have been assigned to blue sharks.

Estimation of unreported blue shark catches based on target species ratios

The fourth column in Table 2 provides the final results of the ratio estimate (based on the disaggregated blue shark catches as a starting point). The estimated unreported catch component is shown in Figure *11* by aggregate gear group. The estimated unreported catches peak around 2008 and 2015 with a reduction in estimates for the years 2011-2012. This overall trend is unsurprising as it is similar to that of the target catches where numbers declined around the late 2000s due to the impact of piracy in offshore waters on catches of pelagic species (Figure *12*). Unreported catch estimates are only available for those gear types that have been reporting catches of blue shark over time (gillnets, longlines and other lines). The estimates are dominated by the longline catches in early years, followed by other lines and gillnets in very recent years. Figure *13* provides a more detailed breakdown of the estimates by gear type, highlighting the particularly high estimated catches by handlines and coastal longlines. Estimated gillnet catches are very low until 2010, due to the low catch rates reported by gillnet fleets from around 1985 to 2010. Subsequent reported catch rates are much higher for the gillnet fleets, reaching around 40% relative to catches of target species (Pakistan and Yemen) (Figure *14*).

The final estimates from the ratio method are presented in Figure *15*. The overall estimated quantities are higher in recent years. The peak in catches in 2011 that is present in the disaggregated catch series is smoothed out due to the decline in estimated unreported catches at this time and total catch estimates instead remain relatively stable from around 2008- 2015.

Estimation of unreported blue shark catches based on GAMs

A range of explanatory variables were explored through the GAM models: Year, Gear, Area, Fishing Ground, Target Catch (YFT+BET+SKJ+ALB+SWO), Tropical tunas (YFT+BET+SKJ), Temperates (ALB and SWO), Other (not target or shark), Other sharks and BSH catch. Target catch is the sum of Tropical tuna and temperate catch. Given that the aim of the method was to predict the catches of countries that had not reported BSH catches, country was not used as an explanatory variable. The model was set up using only those records where blue shark was reported and the resultant coefficients were estimated. These were then used to estimate the unreported catch component by predicting the missing values based on the records where blue shark was not reported.

Stepwise model development resulted in the range of models shown in Figure *16*. Multiple other models were also fit to the data, however the resulting estimates of catch were often highly variable (with inter-annual fluctuations in the order of 10-20 thousand t), or estimated extremely high catch in the early part of the model when the exploitation was thought to be lightest. The following model was selected as the best based on AIC ranking:

gam(formula = log(BSH_catch) ~ as.factor(Year) + s(TAR_catch) + Gear * Area + Fgrounds)

The residual diagnostics are shown in Figure *17* and Figure *18*. The results of the GAM modelling provide final estimates that are very similar to the ratio based estimates, however there are greater estimated catches in the early years resulting in a slightly flatter overall trend (Figure *15*). Estimated catches in the early years are primarily attributed to the Japanese longline with a small amount estimated for the Taiwanese longline and the gillnet fleets. In later years, the relative distribution of catches across fleets remains fairly consistent, however, the scale is greater and total catches are estimated to reach approximately 60,000 t (Figure *19*).

4 Discussion

The methods described in this paper attempt to account for two key sources of error in reported catches: (i) not reporting to species, and (ii) not reporting at all. The procedure used to disaggregate reports of 'sharks NEI' has been used to address the limited reporting to species level, while ratio and GAM based models using target catches can be used to predict the expected catches where there are zero reported catches. The accuracy of all of these methods is entirely dependent on the quality of the original data on which they are based.

The disaggregation approach is the same procedure as that applied to each of the main IOTC species ahead of each assessment, so the approach has been approved by the IOTC Scientific Committee and has been established as the best practice method to use. Nevertheless, the results have not been fully evaluated with respect to situations of poorer data quality and may require more manual oversight to ensure appropriate proxy fleets are assigned where data are particularly sparse.

The ratio and GAM based methods both provide different approaches to the estimation of the 'missing' blue shark catches. The ratio based method used the nominal catches plus the results of the disaggregation process as the starting point for estimations, while the GAM was based on the reported nominal catches.
A key assumption of both of these methods is that all zero reported catches, where there are reported catches of target species present, are false. This might present an overestimation bias in the results by estimating catches where there were actually zero catches. Nevertheless, the data used were based on aggregated annual values and so, given this time period of aggregation, the assumption that reported zero catches are false seems reasonable. These methods also make the assumption that target catches are reported accurately. If target catches are in fact also under-reported, then this may result in an underestimation bias in the results. Nevertheless, as only the five species for which data are deemed to be of reliable quality are used, this should also be a reasonable assumption.

A further assumption these methods make is that those fleets that are reporting positive blue shark catches are doing so accurately. Due to issues with the reporting of processed weight rather than round weights and retained catches rather than total catches, this may also lead to an underestimation bias in the results. Estimated catches will be greatest for gear types for which there are a large number of zero reporters (with substantial target catches) and a high average catch rate by the reporting fleets. If there are few zero reporters but many under-reporters, this will result in under-inflated catch rates and underestimates for the final catches. A filtering approach was used here to remove fleets which were deemed to be targeting sharks to avoid over-inflated catch rates, however, establishing lower thresholds was more problematic with the data available.

The GAM method uses a statistical approach to fill in the gaps where data are lacking and so provides advantages over the ratio method where simple average catch rates are used. The GAM method also uses a greater number of predictor variables to account for items such as spatial differences in catch rates where the sparse and patchy nature of the data means that this is not appropriate for the ratio method.

Any type of catch reconstruction that is attempted will include some level of error, so in practice it is common to include multiple alternative catch time series in assessments for data limited stocks such as these and to explore the outcomes based on the different sensitivity runs. This paper outlines the methods and results for two new alternative catch series that may be used in the assessment model; a series based on disaggregated catches followed by a ratio approach to estimation and a GAM estimation method. If a preferred catch series is to be used as an alternative series for the assessment, then it is recommended that the GAM estimated catch is used.

5 Tables

Table 1. A sample of the fleet / gear / area / region / type of operation configuration mappings used by the Nominal Catch disaggregation

Country	Rep. country	Gear	Area	Region	Type of operation
…	…	…	…	…	…
ESP	ESP	LLEX	IREASIO	EASIO	IND
ESP	ESP	LLEX	IRWESIO	WESIO	IND
ESP	ESP	PS	IREASIO	EASIO	IND
ESP	ESP	ELL	IREASIO	SWEIO	IND
ESP	ESP	LL	IREASIO	SWEIO	IND
ESP	ESP	ELL	IRWESIO	SWEIO	IND
ESP	ESP	LL	IRWESIO	SWEIO	IND
ESP	ESP	BB	IRWESIO	WESIO	IND
ESP	ESP	PS	IRWESIO	WESIO	IND
ESP	ESP	SUPP	IRWESIO	WESIO	IND
FRA	FRA	HAND	IRWESIO	MOZCH	ART
FRA	FRA	TROL	IRWESIO	MOZCH	ART
FRA	**FRA**	**ELL**	**IRWESIO**	**SWEIO**	**IND**
FRA	FRA	PS	IREASIO	EASIO	IND
FRA	FRA	PS	IRWESIO	WESIO	IND
FRA	REU	LLCO	IRWESIO	SWEIO	ART
FRA	REU	HAND	IRWESIO	SWEIO	ART
FRA	REU	HATR	IRWESIO	SWEIO	ART
FRA	REU	TROL	IRWESIO	SWEIO	ART
FRA	REU	ELL	IRWESIO	SWEIO	IND
FRAT	FRA	PS	IREASIO	EASIO	IND
FRAT	FRA	HAND	IRWESIO	MOZCH	ART

FRAT	FRA	HATR	IRWESIO	MOZCH	ART
FRAT	FRA	TROL	IRWESIO	MOZCH	ART
FRAT	FRA	ELL	IRWESIO	SWEIO	IND
FRAT	FRA	PS	IRWESIO	WESIO	IND
…	…	…	…	…	…

Table 2. Criteria for identifying proxy fleets defined through the eight procedures used in the Nominal Catch disaggregation process

Procedure #	Fleet	Type of operation	Region	Area	Years
1	Same	Same	Same	Same	Same
2	Same	Same	Same	Same	+ / - 5 years
3	*Any*	Same	Same	Same	Same
4	Same	Same	Same	Same	+ / - 10 years
5	Same	Same	*Any*	Same	Same
6	*Any*	Same	*Any*	Same	Same
7	*Any*	Same	*Any*	Same	*Any*
8	*Any*	Same	*Any*	*Any*	*Any*

Food and Agriculture Organization of the United Nations

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

Table 3. IOTC nominal catches and catch estimates

Year	IOTC Nominal	Estimated - disaggregated	Estimated - Ratios	Estimated - GAM
1950	47	126	210	142
1951	269	395	474	806
1952	293	388	478	1,756
1953	297	362	448	2,241
1954	367	426	514	3,667
1955	367	422	512	4,908
1956	389	456	533	3,349
1957	372	456	541	4,031
1958	371	457	544	4,633
1959	372	464	555	4,634
1960	367	484	568	4,775
1961	394	532	625	4,923
1962	488	695	774	4,154
1963	497	774	859	5,084
1964	2,679	2,949	3,313	6,643
1965	1,859	2,120	2,371	4,563
1966	2,048	2,351	2,617	5,627
1967	2,906	3,244	3,834	8,970
1968	2,217	2,592	3,750	7,994
1969	2,452	2,863	5,242	7,929
1970	1,470	8,399	9,995	4,968
1971	1,506	8,143	8,980	5,246
1972	1,536	8,519	9,497	5,259
1973	1,158	3,740	4,424	3,300
1974	1,531	5,449	6,613	4,689
1975	1,851	4,120	5,053	4,973
1976	1,654	5,453	8,245	4,722
1977	1,888	8,275	9,735	5,264
1978	2,122	9,788	10,871	6,290
1979	1,936	8,904	10,109	7,828
1980	2,080	10,475	12,153	7,422
1981	2,464	3,434	5,071	10,052
1982	2,919	3,895	5,193	9,663
1983	2,981	4,121	5,701	10,632
1984	3,111	3,815	6,343	11,045
1985	2,892	3,705	6,663	7,111
1986	2,973	3,885	7,849	10,452
1987	2,911	3,859	8,604	9,740
1988	3,363	4,533	10,497	11,387
1989	3,768	11,963	17,514	14,461

1990	3,013	9,868	14,646	9,950
1991	3,733	11,882	16,316	11,671
1992	3,567	15,137	20,481	10,910
1993	5,169	18,970	24,053	13,254
1994	6,499	15,748	23,646	17,120
1995	6,841	12,599	18,009	18,086
1996	7,421	11,166	17,164	18,831
1997	8,847	12,326	19,575	26,937
1998	8,876	13,196	19,971	17,167
1999	12,123	16,762	23,341	22,099
2000	12,404	16,634	26,084	29,921
2001	10,484	15,246	21,158	22,442
2002	11,854	19,612	26,350	26,748
2003	15,354	21,612	30,835	34,442
2004	21,399	25,284	38,192	42,024
2005	24,393	27,264	36,735	44,194
2006	21,447	24,710	35,282	40,243
2007	23,293	25,170	36,527	39,064
2008	24,145	31,518	51,154	44,061
2009	26,563	33,807	46,214	48,336
2010	27,414	36,645	51,140	46,639
2011	28,033	46,974	56,587	49,034
2012	28,159	35,109	44,140	52,931
2013	32,302	39,091	51,675	60,400
2014	29,124	30,472	42,300	57,867
2015	29,916	31,671	46,473	54,735

Table 4. Yearly breakdown of all original SKH ('Sharks various nei') catches into their component species (reported quantities are in mt)

	Species								
	OTH	POR	SKH	SMA	OCS	BSH	SPZ	FAL	
Year	Other shark species	Porbeagle	Sharks various nei	Shortfin mako	Oceanic whitetip shark	Blue shark	Smooth hammerhead	Silky shark	Total
1950	25	15		233	392	79	578	6,353	7,676
1951	41	15		247	462	126	589	7,513	8,993
1952	31	17		278	468	96	687	7,875	9,451
1953	20	19		295	457	65	751	7,604	9,210
1954	19	17		267	413	59	679	7,032	8,487
1955	17	21		314	470	55	807	7,881	9,564

1956	21	17		265	419	67	669	6,891	8,349
1957	26	31		471	709	84	1,211	10,701	13,234
1958	28	17		266	442	86	662	7,222	8,723
1959	30	17		277	464	93	687	7,665	9,232
1960	38	20		323	551	117	797	8,857	10,703
1961	45	17		281	521	139	674	8,589	10,265
1962	68	21		359	698	208	841	11,267	13,462
1963	91	25		432	869	277	997	13,869	16,559
1964	88	37		605	1,087	270	1,462	16,729	20,279
1965	85	43	4	689	1,188	261	1,694	18,038	22,003
1966	97	70		1,087	1,753	302	2,731	25,673	31,713
1967	108	80		1,242	1,994	338	3,125	29,259	36,145
1968	121	80		1,248	2,044	375	3,122	30,129	37,118
1969	133	77		1,215	2,040	411	3,013	30,302	37,191
1970	105	82		1,270	2,021	6,929	3,205	30,033	43,645
1971	77	120		1,482	2,025	6,637	3,575	30,436	44,352
1972	120	158		1,980	2,814	6,984	4,894	42,154	59,104
1973	147	184		2,350	3,285	2,582	5,505	47,980	62,032
1974	126	129		1,399	1,869	3,917	2,884	30,193	40,517
1975	103	124		1,302	1,839	2,269	3,260	29,800	38,695
1976	121	119		1,264	1,914	3,800	3,337	32,148	42,703
1977	86	209		2,091	2,432	6,387	4,562	39,064	54,830
1978	109	245		2,150	2,399	7,665	4,661	40,046	57,276
1979	109	212		2,053	2,500	6,968	4,698	40,648	57,187
1980	112	166		1,524	1,945	8,395	3,577	33,804	49,523
1981	128	142		1,342	1,699	970	2,667	28,906	35,854
1982	85	142		1,253	1,310	976	2,123	23,547	29,436
1983	123	136		1,054	1,103	1,140	1,493	20,936	25,986
1984	79	100		844	994	703	1,642	18,110	22,472
1985	86	128		1,150	1,337	812	2,274	24,804	30,590
1986	81	147	238	1,254	1,303	912	2,265	26,373	32,573
1987	84	139	10,427	1,230	1,301	948	2,138	27,091	43,358
1988	88	161	638	1,421	1,497	1,170	2,546	28,577	36,097
1989	91	152	7,403	1,277	1,447	8,195	2,475	28,104	49,143
1990	95	165	1,150	1,591	1,837	6,854	3,238	32,834	47,764
1991	98	211	8,740	1,974	2,026	8,149	3,547	37,785	62,531
1992	102	415	2,608	3,209	2,139	11,570	3,765	47,993	71,801
1993	122	598		4,395	2,166	13,801	3,677	51,004	75,763
1994	121	199		1,241	1,280	9,249	1,913	41,880	55,882
1995	107	388		3,502	2,278	5,758	4,115	52,292	68,441
1996	85	265		3,416	2,256	3,746	4,405	52,995	67,167
1997	52	397	25	2,779	2,758	3,480	9,954	58,267	77,713

1998	272	467	27	3,350	2,982	4,320	8,967	61,222	81,607
1999	379	291	636	2,343	2,720	4,640	8,356	54,512	73,876
2000	120	256	1,181	2,089	2,576	4,231	7,224	57,115	74,792
2001	94	252	918	1,976	2,373	4,762	7,325	56,661	74,360
2002	69	241	1,059	3,976	2,309	7,758	7,516	55,808	78,737
2003	135	307	867	1,766	2,163	6,258	8,954	59,361	79,812
2004	50	407	848	1,911	2,167	3,885	8,477	59,544	77,288
2005	258	250	789	1,413	1,875	2,871	7,710	46,895	62,062
2006	48	251	941	1,390	1,727	3,263	7,507	44,558	59,686
2007	44	234	896	1,054	1,499	1,877	7,101	40,717	53,423
2008	43	264	674	956	1,269	7,373	6,394	30,630	47,603
2009	29	372	1,089	929	1,171	7,244	6,334	32,430	49,598
2010	36	254	1,107	945	1,209	9,230	6,606	30,694	50,082
2011	31	320	1,143	611	775	18,941	7,055	25,422	54,298
2012	27	513	1,253	1,144	1,700	6,951	6,156	27,374	45,118
2013	35	310	1,323	1,328	1,619	6,788	7,693	31,185	50,281
2014	72	2,258	1,203	958	1,446	1,349	6,685	27,596	41,566
2015	385	0	929	1,455	3,927	1,754	9,045	40,030	57,525
Total	6,100	13,538	48,115	91,781	106,723	247,968	259,275	2,049,006	2,822,506
	OTH	POR	SKH	SMA	OCS	BSH	SPZ	FAL	

Table 5. Explanatory tables for regression trees on area(top) and gear. The columns include *var*, the variable used at the split (or <leaf> for a terminal node), *n*, the number of cases reaching that node, *dev* the deviance of the node, *yval*, the fitted value at the node. This table should be viewed in conjunction with Figure *5*.

Explanatory table for regression tree on area

var	n	dev	yval	splits.cutleft	splits.cutright	Terminal Node	Node Label
Area	1549	11733.5	3.7	:dghijkln	:abcefmopq		Root
Area	282	1181.0	0.6	:hjk	:dgiln		HAND,SLL,SPOR,TROL
<leaf>	25	77.0	-1.7			*	TROL
<leaf>	257	949.2	0.9			*	HAND,SLL,SPOR
Area	1267	7364.4	4.4	:abemopq	:cf		ELL,FLL,GILL,GIOF,GL,LG,LL,LLCO,LLEX
Area	1161	6300.8	4.1	:bep	:amoq		ELL,FLL,GILL,GIOF,GL,LG,LL,LLEX
<leaf>	629	3633.5	3.8			*	ELL,FLL,GILL,GIOF,LL,LLEX
<leaf>	532	2483.9	4.5			*	GL,LG
<leaf>	106	188.1	7.1			*	LLCO

Explanatory Table for regression tree on gear

var	n	dev	yval	splits.cutleft	splits.cutright	Terminal Node	Node Label
Gear	1549	11733.5	3.7	:fklm	:abcdeghij		Root
Gear	238	1357.3	0.9	:m	:fkl		EAFRI,IRAN,MALDI,MOZCH,PERSG,REDSE,SAUAR,SEYCH
<leaf>	49	338.7	-0.6			*	MALDI,PERSG,REDSE
<leaf>	189	878.9	1.2			*	EAFRI,IRAN,MOZCH,SAUAR,SEYCH
Gear	1311	8120.0	4.2	:abcdeghj	:i		ANDAS,ARABS,BAYBE,EASIO,INDON,SEAIO,SRILA,SWEIO,WESIO
Gear	1277	7473.7	4.1	:abcdhj	:eg		ANDAS,ARABS,EASIO,SEAIO,SRILA,SWEIO,WESIO
<leaf>	1221	7288.9	4.0			*	ARABS,EASIO,SWEIO
<leaf>	56	63.1	5.5			*	ANDAS,SEAIO,SRILA,WESIO
<leaf>	34	86.1	8.2			*	BAYBE,INDON

Table 6. Deviance table for GAM estimation model of blue shark catch

Model	Model Component	Resid.DF	Resid.Dev	DF	Deviance
1	log(BSH_catch)~year	542.00	2270.61		
2	+Target Catch	533.09	1487.77	8.90602	782.83
3	+Gear	521.04	910.57	12.05046	577.21
4	+Area	515.04	684.95	6.00802	225.61
5	+Fishing Ground	514.04	684.80	0.99957	0.15
6	+ Gear:Area	509.03	671.56	5.00312	13.23

6 Figures

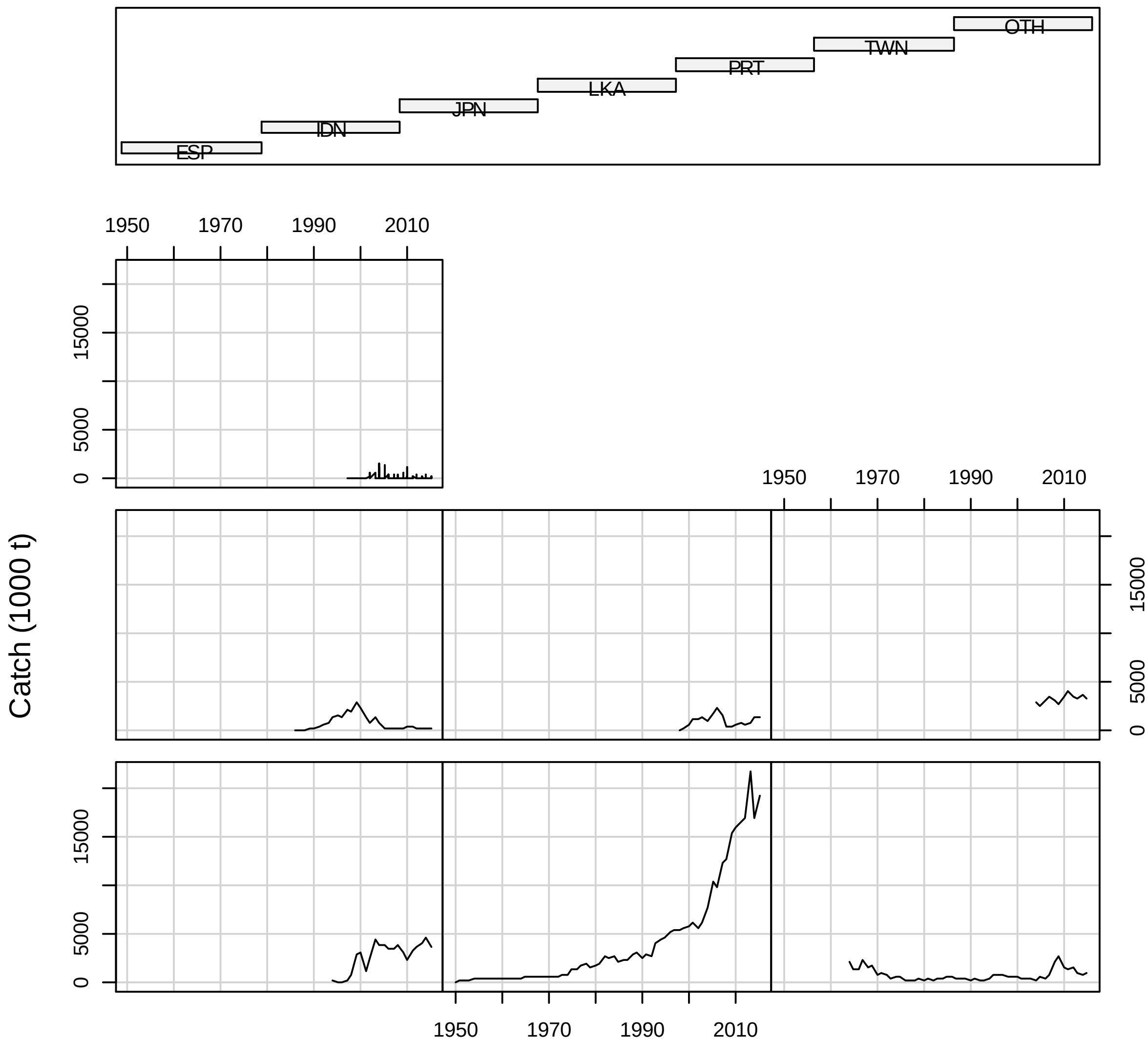


Figure 1. Catches of blue sharks in the IOTC area of competence by CPC (Nominal catches, IOTC database, 2017)

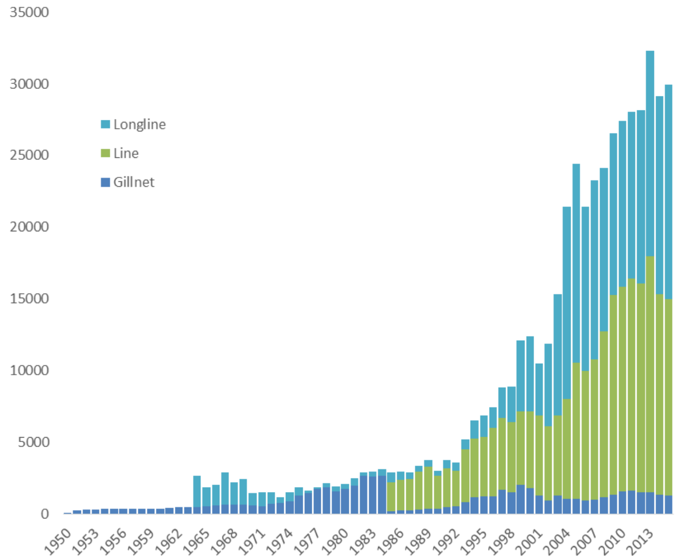


Figure 2. Nominal catches (t) of Indian Ocean blue sharks by gear (IOTC database, 2017)

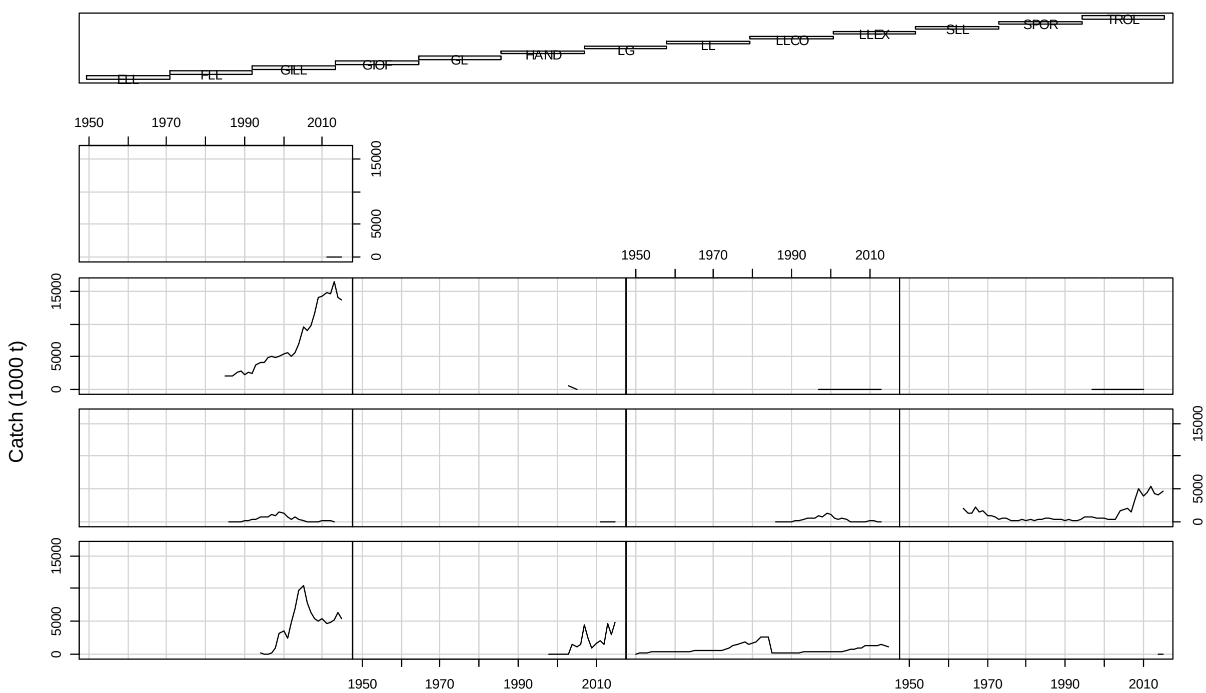


Figure 3. Nominal catches of Indian Ocean blue sharks by gear (IOTC database, 2017)

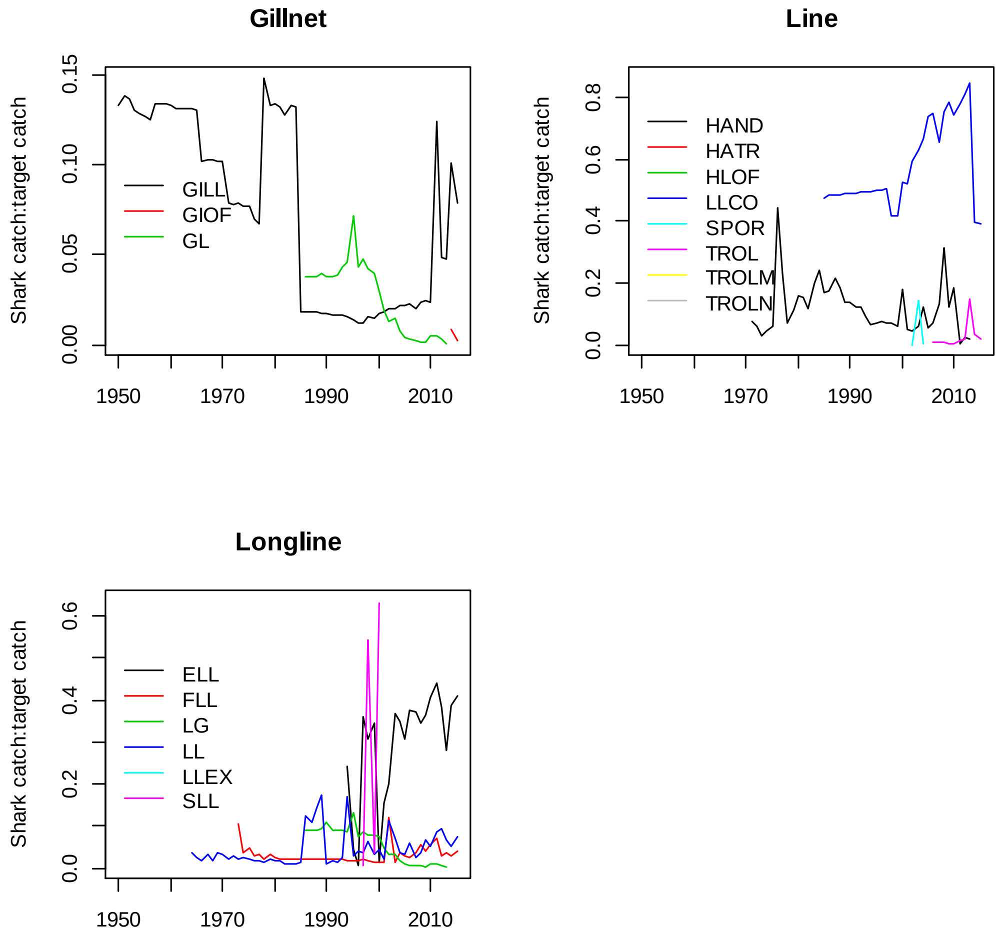


Figure 4. Ratio of blue shark catch to target catch by gear over time

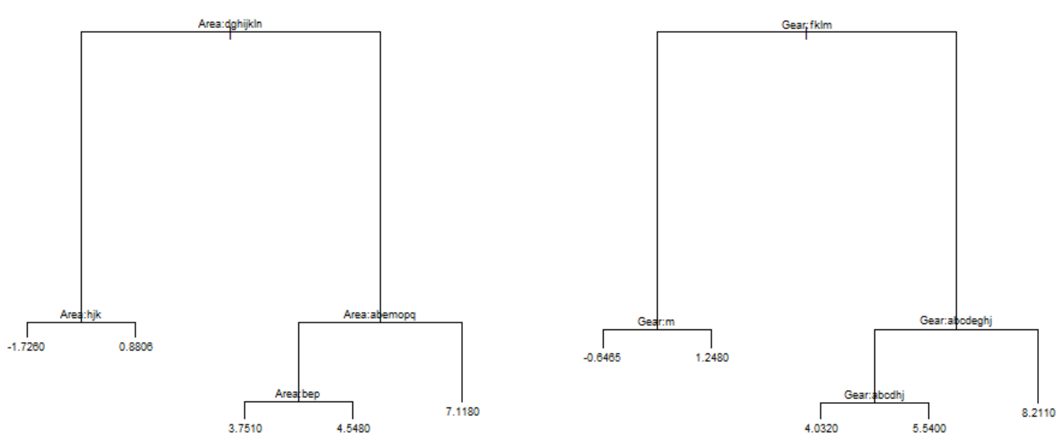


Figure 5. Regression tress for the gear and area groupings. Labels are explained in Table 6.

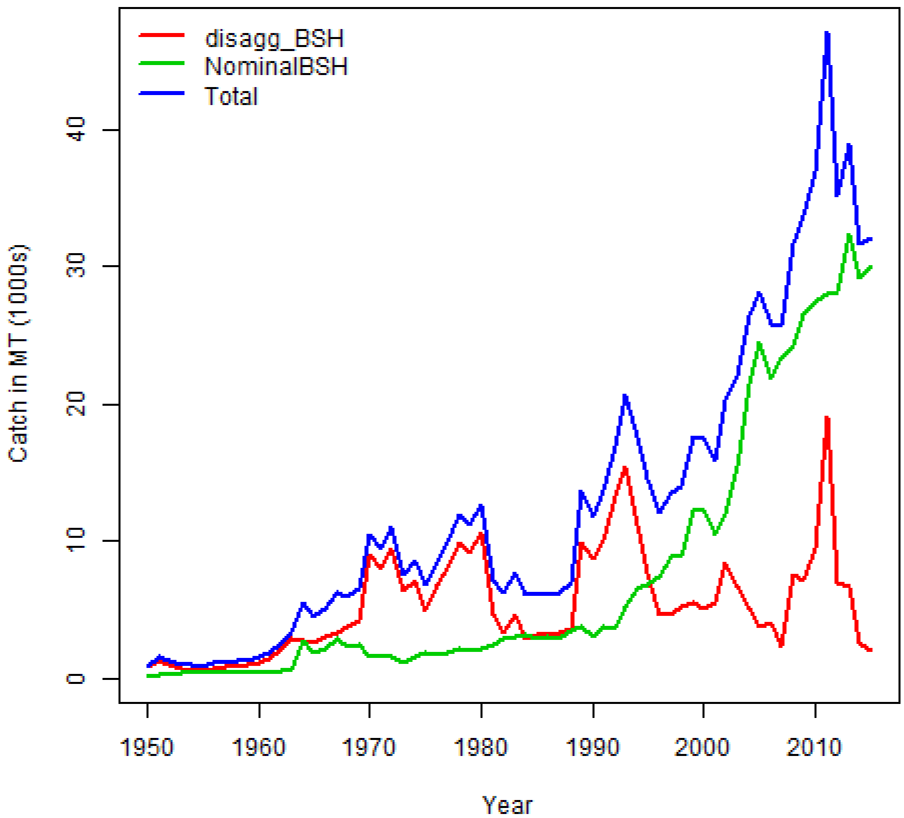


Figure 6. IOTC nominal catches, disaggregated catches and combined catches

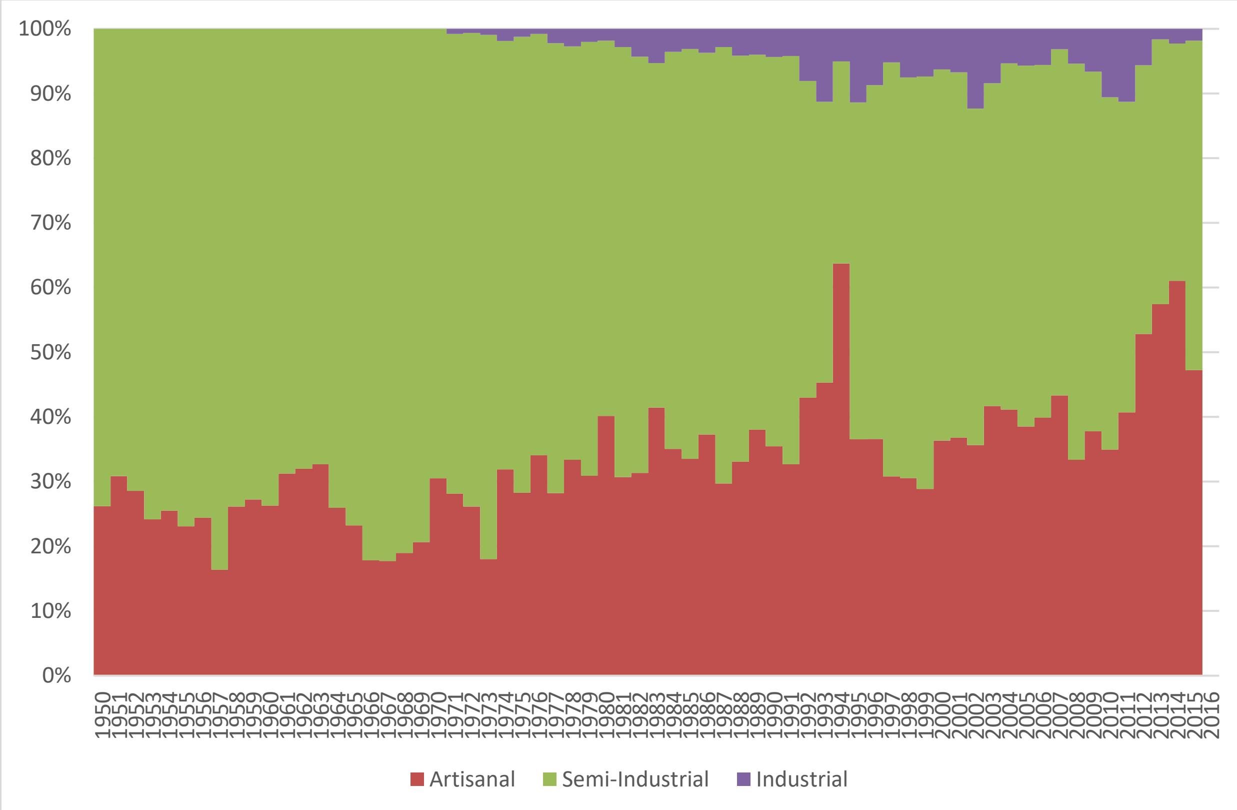


Figure 7. Percentage of Sharks various nei (SKH) catches by operation type

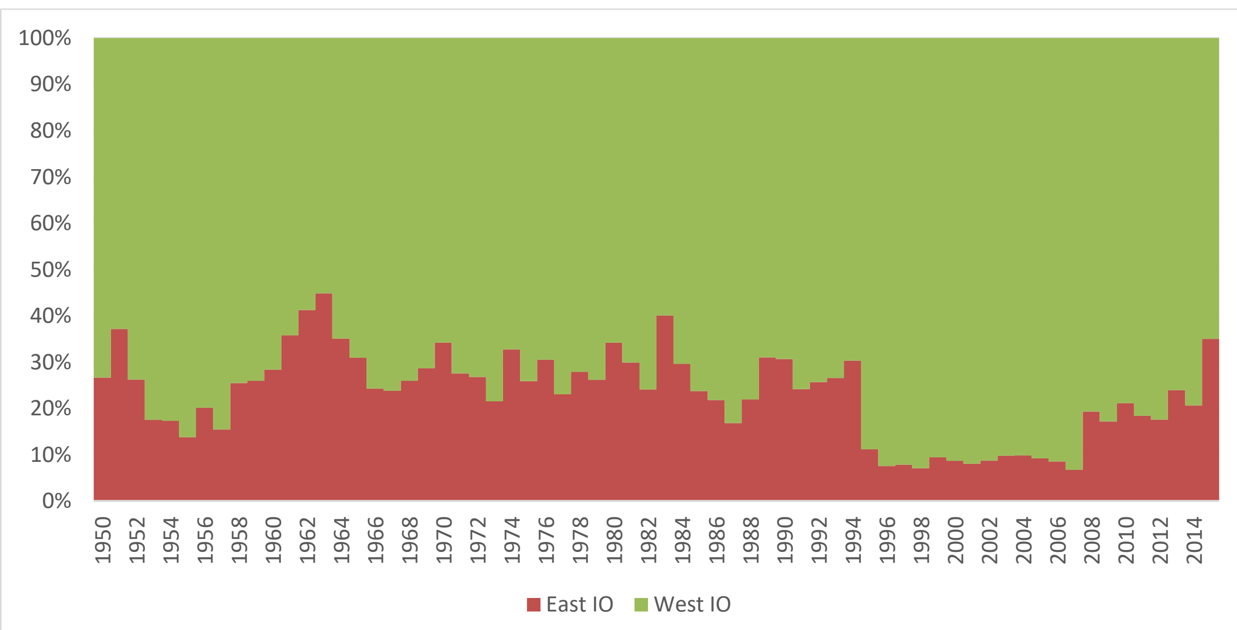


Figure 8. Percentage of Sharks various nei (SKH) catches by Indian Ocean area

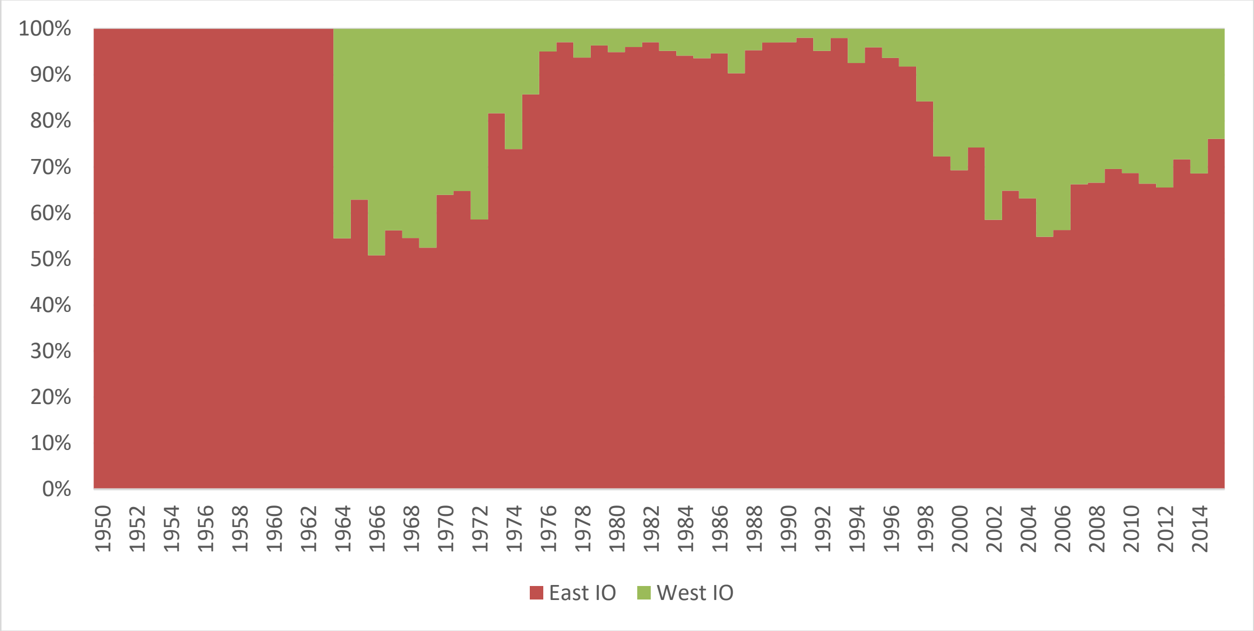
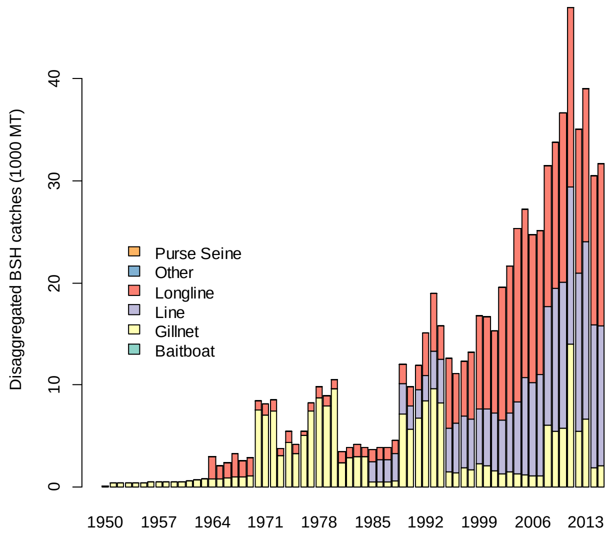


Figure 9. Percentage of blue shark (BSH) catches by Indian Ocean area



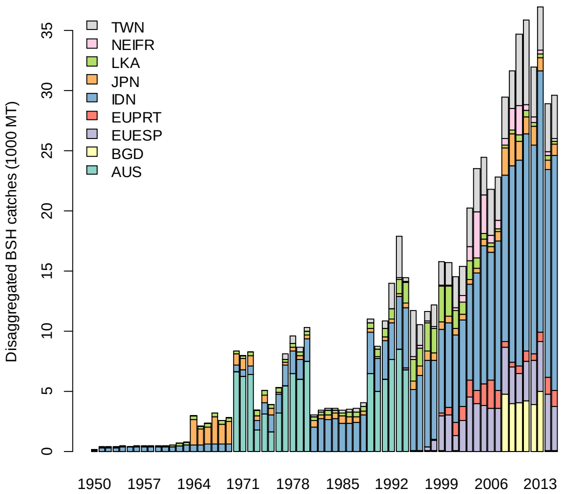


Figure 10. Disaggregated catches by gear type and fleet

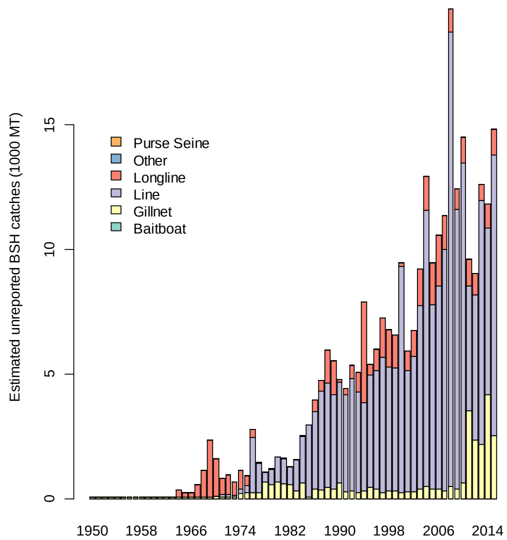


Figure 11. Estimated unreported blue shark catches by gear type

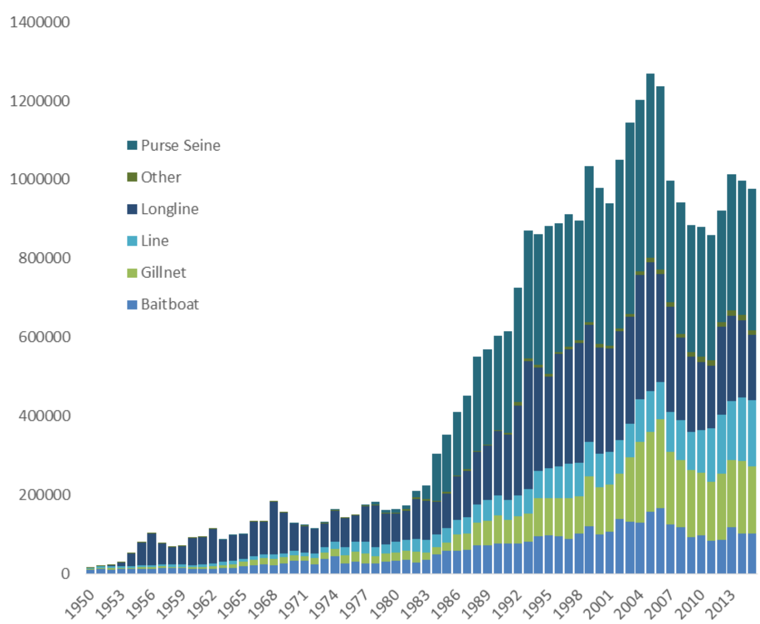


Figure 12. Nominal target catch (t) (YFT,SKJ,BET,ALB,SWO) in the Indian Ocean by gear

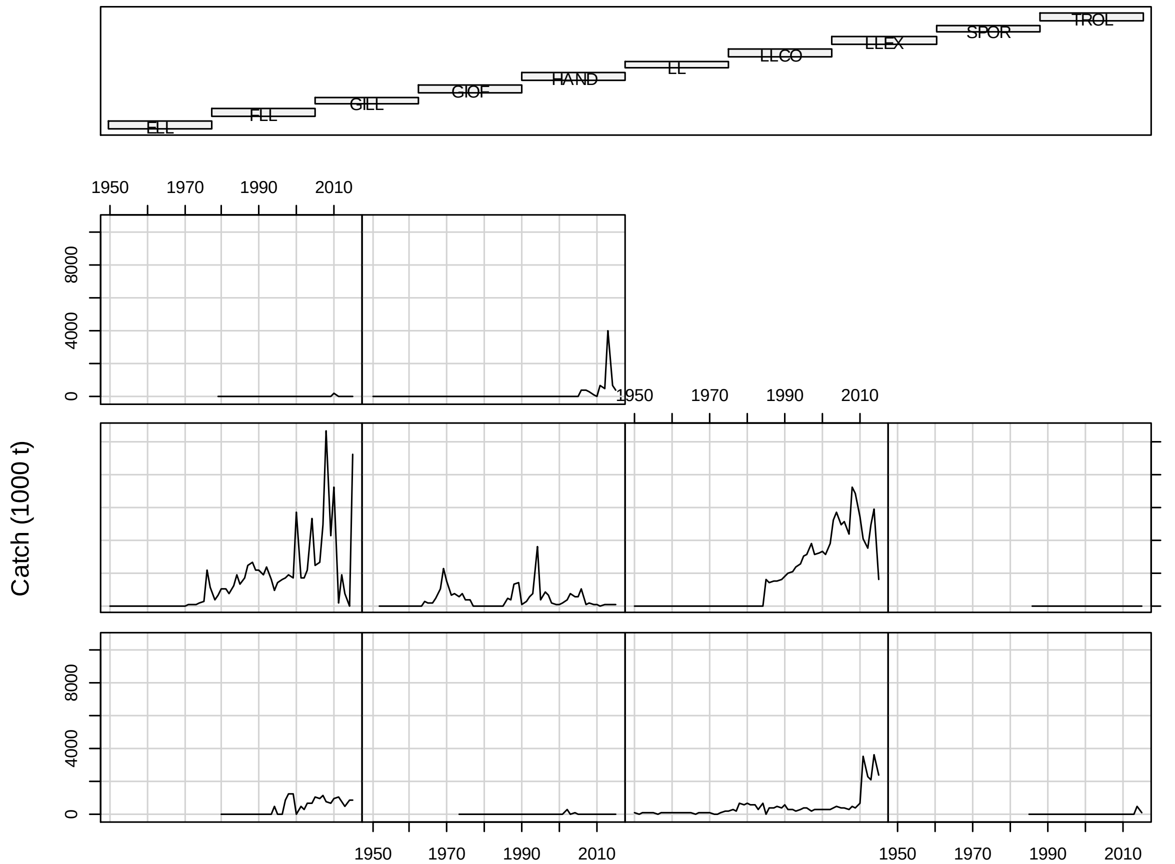
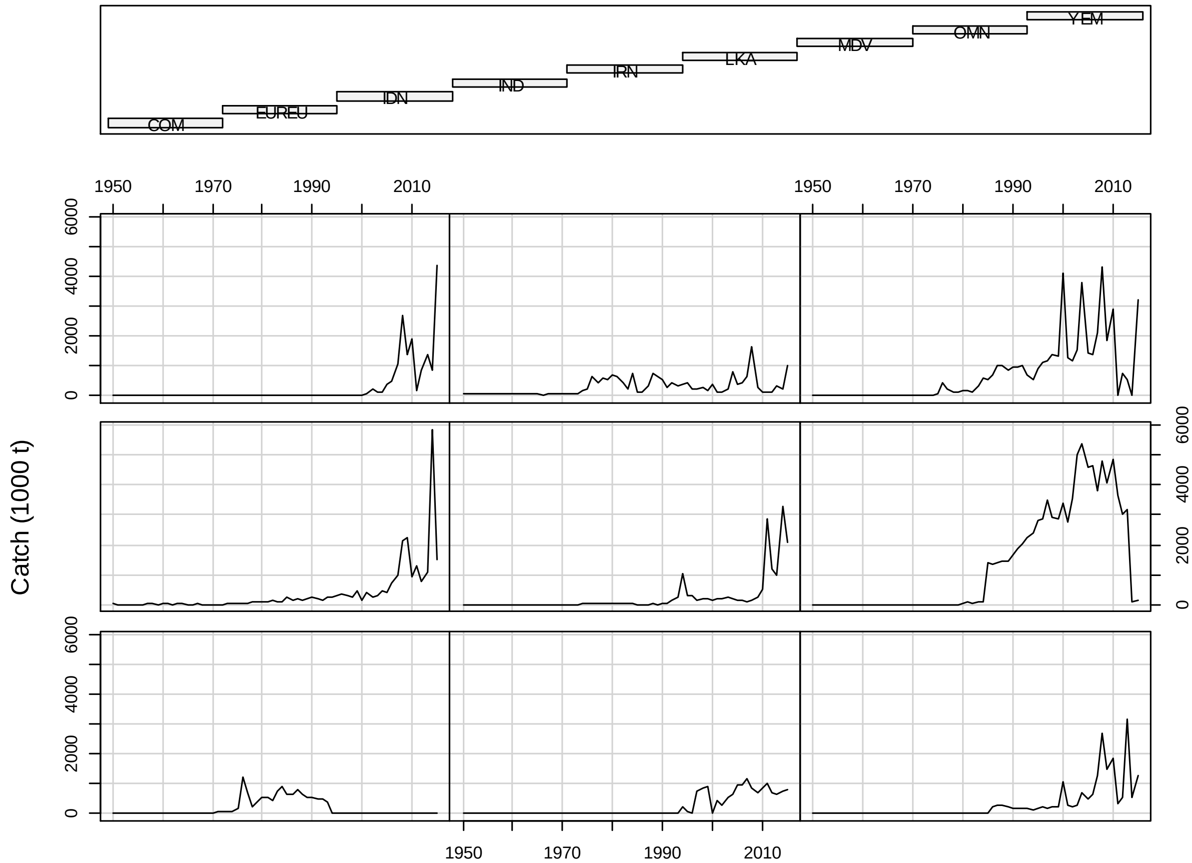


Figure 13. Estimated unreported blue shark catches by gear type



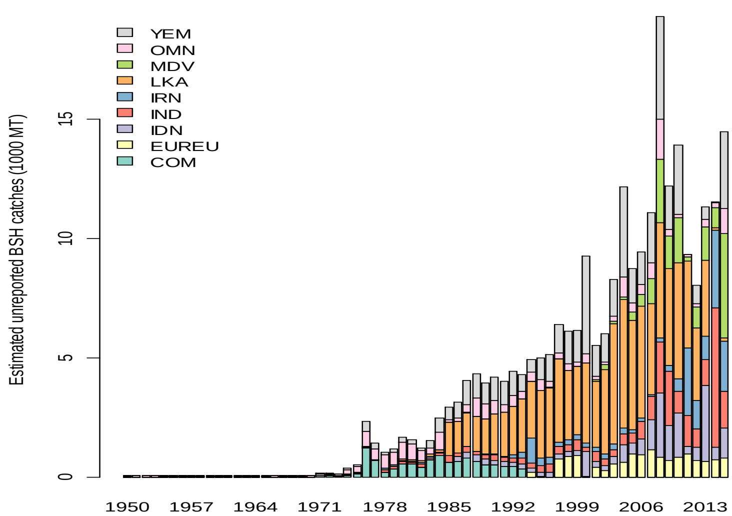


Figure 14 a&b. Estimated unreported blue shark catches by fleet

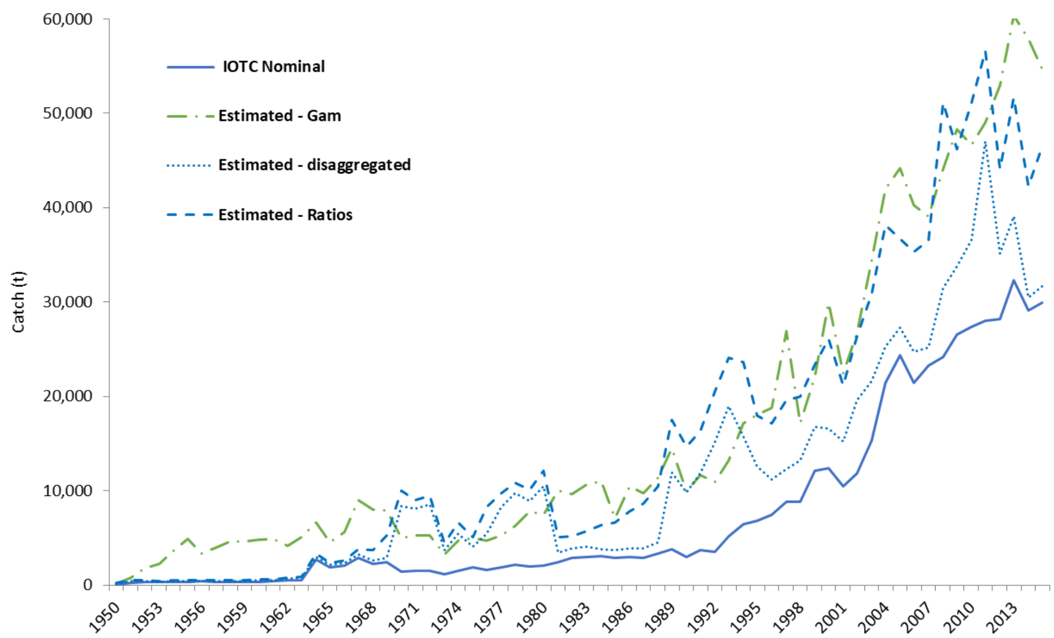


Figure 15. Reported, disaggregated, ratio based and GAM estimates of Indian Ocean blue shark catches

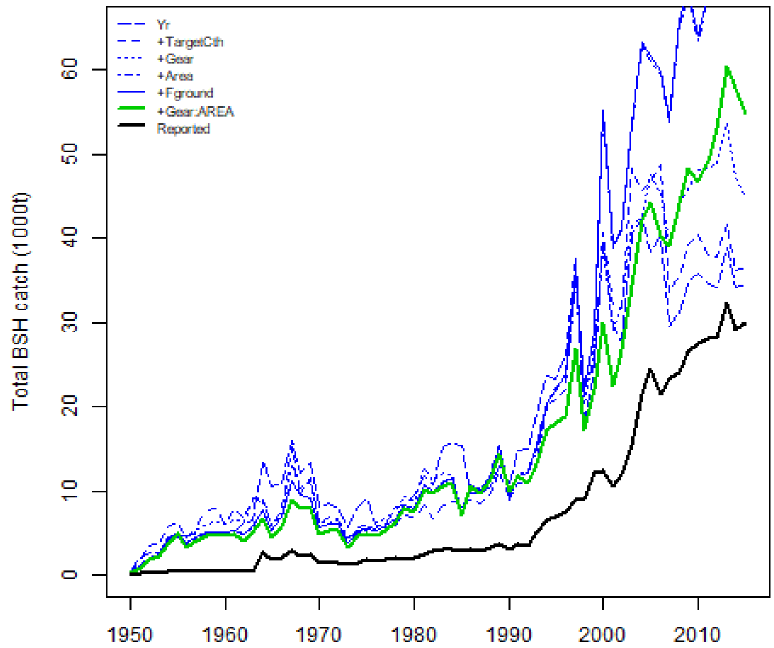


Figure 16. Stepwise results of predicted catch via GAM on the nominal catch data set (selected model = green line).

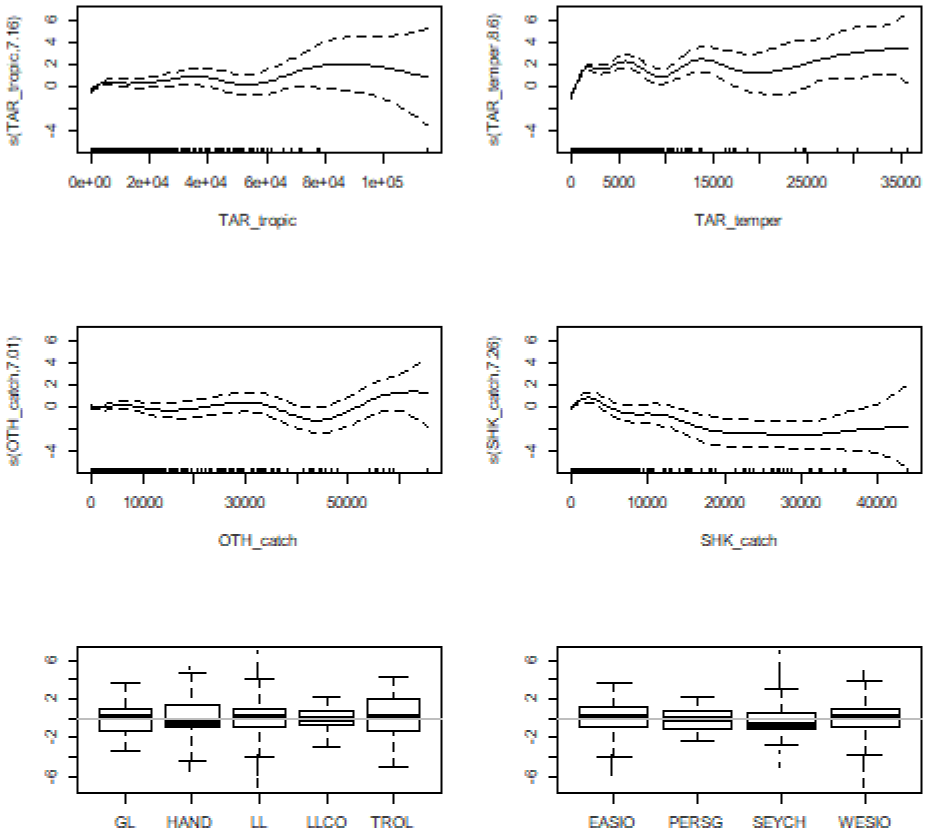


Figure 17. Residual plots of final GAM model

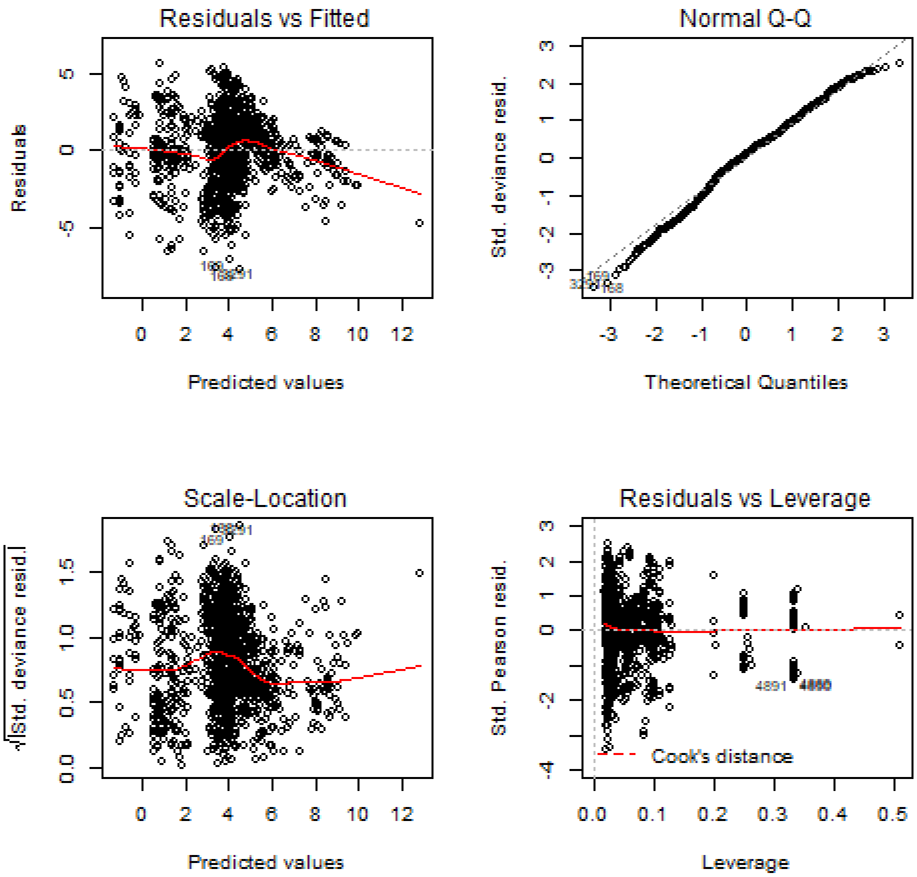


Figure 18. Additional model diagnostics for the GAM model

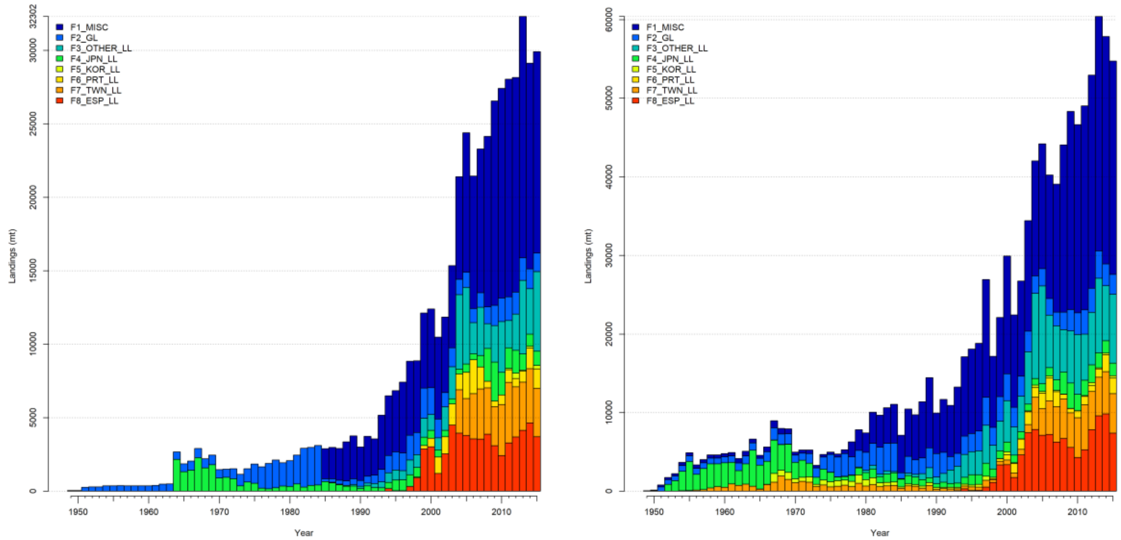


Figure 19. Nominal catch by fleet (left panel) and estimated catch by fleet based on the GAM model (right panel). Note the difference in scale of the y axis.

7 Appendix

Table 1. Pseudo-code implementation of the Nominal Catch Disaggregation process

```
var DisaggregationProcedures := { P1, P2, ..., P8 };

var NC_Original := { NC1, NC2, ..., NCn };

var NC_NonAggregated := { };
var NC_Aggregated := { };

var NC_Proxies := { };

var NC_Disaggregated := { };

for each NC in NC_Original:
      if !NC.isAggregated
            NC_Disaggregated.add(NC);
            NC_NonAggregated.add(NC);
      end if;
end for;

for each NC in NC_Original:
      if NC.isAggregated
            var Disaggregated := false;

            loop: for each Proc in DisaggregationProcedures:
                  NC_Proxies := Proc.apply(NC, NC_NonAggregated);
                  if NC_Proxies != { }
                        NC_Disaggregated.addAll(NC.breakdown(NC_Proxies));
                        Disaggregated := true;
                        break loop;
                  end if;
            end for;

            //Manual breakdown is required if none of the procedures
            //identifies any proxy record to use for the disaggregation
            //of current record
            if(Disaggregated == false)
                  NC_Disaggregated.addAll(NC.manualBreakdown);
      end if;
end for;

return NC_Disaggregated;
```

Where:

- P_1, …, P_8 are the eight currently available Disaggregation Procedures;
- NC_1, …, NC_n is the input Nominal Catch dataset;

- *`procedure`*`.`**`apply`**`(`*`<NC record>`*`, `*`<non aggregated NC records>`*`)` returns the proxy records (according to the current disaggregation *`procedure`*) for the aggregated record *<NC record>* being processed, as these are identified within the full set of *<non aggregated NC records>*;
- *`record`*`.`**`breakdown`**`(`*`<NC proxies>`*`)` breaks down the original, aggregated *record* into multiple disaggregated records, whose catch quantities (and species / gears) are proportionally assigned based on the identified *<NC proxies>*;
- *`record`*`.`**`manualBreakdown`** prompts users for their own, manual breakdown of the original aggregated *record*, as none of the disaggregation procedure was able to identify any valid proxy record for it;