Received: 16 August 2015

Catch Rates of Indo Pacific Sailfish (*Istiophorus platypterus*) as Calculated Based on IOTC Longline Dataset

Humber A. Andrade

Federal Rural University of Pernambuco (UFRPE)
Department of Fisheries and Aquaculture (DEPAq)
Laboratory of Applied Statistical Modeling (MOE)

Abstract

Estimations of relative abundance indices are cornerstones in most of the fisheries stock assessments. In tuna fisheries relative abundance indices are often calculated by standardizing the commercial catch-per-unit-effort (CPUE). Whenever the species of interest is bycatch the task may become difficult because the datasets are limited, incomplete or biased (e.g. underreports). However, in some cases like the Indo Pacific sailfish (*Istiophorus albicans*) to look at those limited databases may be the alternative. In this paper a simple model was used to standardize the CPUE of sailfish based on a limited database, which does not include fishing operational information (e.g. number of hooks between floats). In addition the data are aggregated by month and by square (5º latitude x 5º longitude). Time series of standardized CPUE based on the aggregated database were calculated for Korea and for Japan. Estimations for Korea in 1975-1987 timespan are probably useful for stock assessment. In that timespan the target species of Korean longline fishermen did not change much as indicated by the proportions of the tuna species in the catches. Estimations of standardized CPUE of Japan in 1994-2014 timespan as calculated based on the limited dataset were compared to others calculated using “richer” set by set database. Both estimations were very similar from 1994 to 2007. However, major differences had showed up in the end of the time series. As far as the calculations based on detailed database are less biased, the differences found stand as a warning about using standardized CPUE calculated based on aggregated dataset, at least for fleets like the Japanese one which has experienced major spatial and temporal changes concerning the target and the fishing operational characteristics.

Introduction

In the Indian Ocean most of Indo Pacific sailfish (*Istiophorus platypterus*) (SFA) have been caught by fishermen operating gillnets and handline gear (Anon, 2014). Data concerning SFA in the Indian Ocean are very limited. Only approximate estimations of catch are available for most of the fleets. More detailed catch and effort information is only available for some of the longline fleets. However, longline fishermen aim at tuna or swordfish, while SFA is an eventual bycatch. Often the available data concerning bycatch species are biased or incomplete due. There are many causes of the low quality and quantity of bycatch datasets, like misidentification and underreports for example. The later is probably the major issue for the sailfish, because of its morphological characteristics which make it easier to identify than other billfish.

Data-poor or data-limited terms have been applied broadly to several different cases, but in general they have been used to classify the situations when the best scientific information is not good for determining management benchmarks and the status of the stocks with respect to those reference points (Richards and Maguirre, 1998; Pilling et al, 2008). To look at all available data is a reasonable advice in most of the cases, but mandatory in data-poor cases. All information, even incomplete, may help on taking management decisions concerning data-poor stocks. That was the motivation to look at longline databases of the Indian Ocean Tuna Commission (IOTC) aiming at calculation of catch per unit effort time series, which may be useful to assess the status of the SFA.

Commercial catch rates (e.g. CPUE) of tuna fisheries are often used to estimate relative abundance indices, which are required to run several stock assessment models. Ideally all factors that affect the CPUE should be taken into account to "standardize" the CPUE in order to estimate relative abundance indices. Therefore, detailed data concerning the fishing operation (e.g. location, number of hooks per basket) are of most importance. However, catch and effort tuna longline data as reported to international agencies (e.g. IOTC) are undetailed. Typically the catch and the effort (e.g. number of hooks) data are aggregated (summed up) by month and by 5ºx5º (latitude x longitude) square. Further information on fishing set characteristics, like number of hooks per basket and length of branch lines, are also not available. Nevertheless, such detailed information is collected in the sampling programs of some of the contracting parties.

Whenever possible, calculations of standardized catch rates are to be carried out using detailed dataset. However, in some situations the gross data as reported to the agencies are the only available data, and more simple models are the alternative to standardize CPUE. In principle, calculations using simple models lacking variables are more biased than calculations using models that includes that explains the variability of CPUE. Hence, comparisons of standardized CPUEs as calculated in "rich" situations (more data is available) with those calculated in "poor" situations, may be useful to assess the bias of the poor calculations and the risk of using them later in stock assessments. Therefore, poor standardized SFA CPUE were calculated with two purposes: to provide the very first estimation of relative abundance indices based on SFA caught by some contracting parties for which detailed data are not available, and to provide estimations to be compared to rich standardized CPUE in order to assess the relative bias of poor standardized CPUE.

Data and Analyses

Database analyzed concerns catch and effort of longline fleets that operated in the Indian Ocean as reported to the Indian Ocean Tuna Commission (IOTC) (IOTC, 2015). Catches of the species were reported in number of fish or weight. Number of hooks is the unit of effort. There are also information concerning the year, month and location (5º latitude x 5º longitude) where the longline was deployed in the water. Number of missing data and of catches equal to zero were calculated by year and month for each fleet in order to uncover underreported and false zeros. Exploratory analyses of the positive catches were also warranted.

In order to try to estimate a relative abundance indice, catch rates were standardized using a generalized linear model (Dobson, 2002). Because there were not catches equal to zero in the reports, only positive data entries. Further discussions on the motivations to analyze only positive data are below in the discussion. However, to analyze only positive data is an alternative when dealing with bycatch (e.g. Baum et al., 2003; Baum and Blanchard, 2010). Further comparison of the results of this approach with those gathered in a "rich" situation (e.g. there are zeros and they are assumed to be true zeros), may be useful to assess the usefulness of the alternative to assess only positive data zero reports concern.

In this paper gamma distribution, which is for positive data was used to model catch per unit effort (fish/hook). However, truncated distributions for counting data and the use of effort as offset are alternatives to be investigated in the future. After initial exploratory analyzes logarithm link function performed better. The model has converged and the estimations with log link function resulted in lower values of Akaike Information Criterion (AIC – Akaike, 1974).

Explanatory variables straightforward available for analyzes are year, month and approximate location of the catch. However, there are not information about the strategy of fishermen concerning the longline deployment in the water (e.g. day or night), the operation procedure, and the gear characteristics (e.g. number of hooks between floats) in the IOTC datasets. That information would be useful to figure out which species was the fishermen target. An alternative is to use the proportions of the different species in the total catch as proxies of the fishermen target. When the species analyzed (the response variable) is also one of the target species the use of proportions as explanatory variables may be a misleading approach because of the circular logic. However, sailfish was always a bycatch for the longline fleets, hence I have assumed the proportions of yellowfin, bigeye and albacore in the catch were potentially useful proxies of fishermen's targets.

Main effects and first order interactions of variables (year, month, area and proportions of yellowfin, bigeye and albacore as target proxies) were considered in the model. In order to select the important exploratory variables, I have started with a full model and I have used backward procedure with AIC as criterion. However, I have also took into account hypothesis tests based on the reduction of deviance. If the inclusion of the explanatory variable or interaction did result in significant reduction of the deviance, it was discarded. Main effect terms were discarded only all the interactions including it were already discarded. Detailed information on model selection can be found in McCullagh and Nelder (1989).

In the generalized linear models the explanatory variables may be considered as fixed effects or as random effects. If all the explanatory variables are fixed effects the usual model acronym is GLM. If both fixed and random effects are included the model is denominated as Generalized Linear Mixed Models (GLMM) (Pinheiro and Bates, 2000). The inclusion of year in interactions as fixed effect imposes difficulties to calculate the standardized CPUE, which should reflect the separated effect of year. The solution usually is the use of a weighted average calculation (e.g. "least-square means" or "population marginal means" – Searle et al., 1980) in which weights have to be assigned to the levels of factor interacting with year (Maunder and Punt, 2004). In opposition if interactions with factors and year are included in the models as random effects, the separated effect of year can be easily calculated. In this paper I have tried to fit both models, GLM and GLMM, in an attempt to calculate the year effect and the standardized catch rates. In this paper the equal weights when calculating standardized catch rates using "least square means".

Results

There are 105,169 entries in the database, but catch in number is missing in 95,731 (91.02 %) entries, while catch in weight is missing in 101,870 (96.9 %). There were not reports of catches in number or weight equal to zero. The minimum catch in weight was 0.285 kg, while the minimum catch in number was 1. The variable catch in weight lacks more data than catch in number. Hereafter the analyses were based upon the later variable. Only Australia (n=31), La Réunion (n=230), Japan (n=4716), Korea (n=3849), Mauritius (n=6) and Seychelles (n=612) have reported catches of sailfish in number.

Effort, proportion of positive catches, number of fish caught and catch-per-unit-effort (CPUE) are shown in Figure 1. Time series of Australia and La Réunion are short. Seychelles database is not that short, but it is inconsistent in the sense the catch rates increases from close to zero to a very high value and drop back to approximately zero in a short time span. It is important also to highlight that Seychelles fleet usually operates in an small part of the Indian Ocean. Time series of Korea seem useful from mid 1970's until the beginning of 1990's, in the sense there are both, information about effort and about catch in that time span. Japan database is the longest one, and it conveys information about catches from 1993 to 2014, in spite of the peak in 2000's.

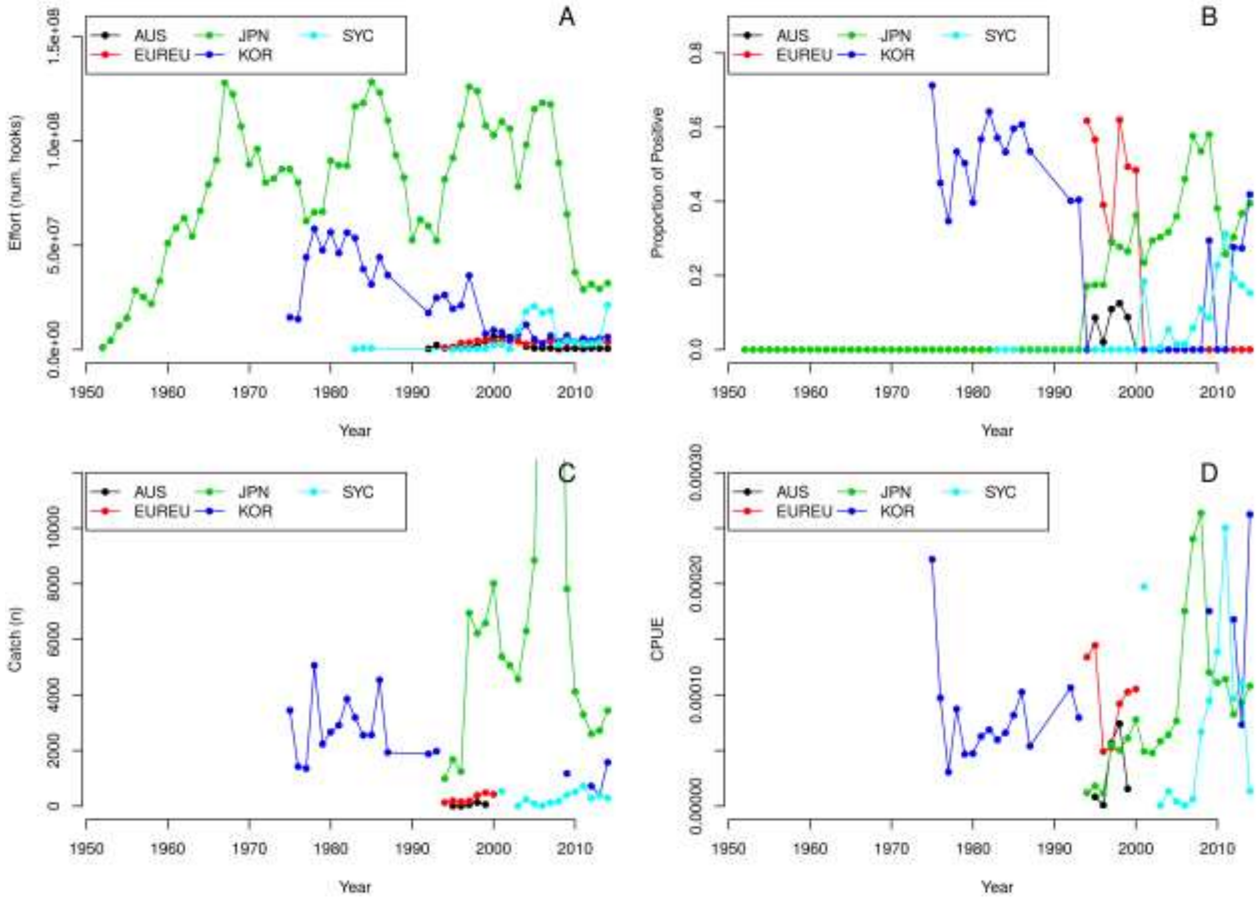

Figure 1 – Effort, proportion of positive catches, number of fish caught and catch-per-unit-effort (CPUE) of Indo Pacific sailfish. Australia (AUS); La Réunion (EUREU); Japan (JPN); Korea (KOR); Seychelles (SYC).

Follow below a brief description of Japan and Korea time series, which seem to be the more informative database. Effort time series of Japan is the longest (Figure 1 A). It covers from 1950's to 2014. However catches were reported since mid 1990's (Figures 1 B and C). The proportion of positive reports in the japan dataset increased from 0.2 to approximately 0.3 from 1994 to 1997, and it was still close to 0.3 until 2004 (Figure 1B). In the mid 2000's the proportion increased to reach 0.6, but after 2009 it decreased back to approximately 0.3. Number of fish caught by Japan increased to 6,000 in 1990's, but peaked in end of 2000's. However, the catch decreased quickly after 2009 to approximately 3,000 fishes (Figure 1 C). Catch-per-unit-effort values as calculated for Japan database were close to 0.00006 fish per hook from 1997 to 2005, but peaked in the end of 2000's and decreased to approximately 0.00010 fish per hook after 2008.

There were two phases in the Korea time series, the first covers from 1975 to 1993, and the second covers from 1994 to 2014. Effort, catch, catch rate and the proportion of positive catches time series seem consistent in the first, but not in the second phase (Figure 1). In the first phase the proportion of positive catches ranged from 0.4 and 0.6, the catches were close to 2500 fishes on average, and most of the CPUE were between 0.00005 and 0.00010 fish per hook (Figures 1 B, C and D).

Monthly variations of catch and CPUE across the years as calculated based on Japan and Korea datasets are shown in Figure 2. Catches of Japan showed similar pattern from 1998 to 2004 (Figure 2 A). Catches were low in the mid of years, but were similar and high in the beginning and in the end of the year. Catches have increased from 2005 to the beginning of 2008, decreased until 2010. In the end of the time series variations of catches showed no seasonal pattern. Estimations of CPUE were usually low in the mid of the year from mid 1990's until 2007 (Figure 2 B). However, the seasonal periodic signal in the CPUE series was not so evident as in the catch time series

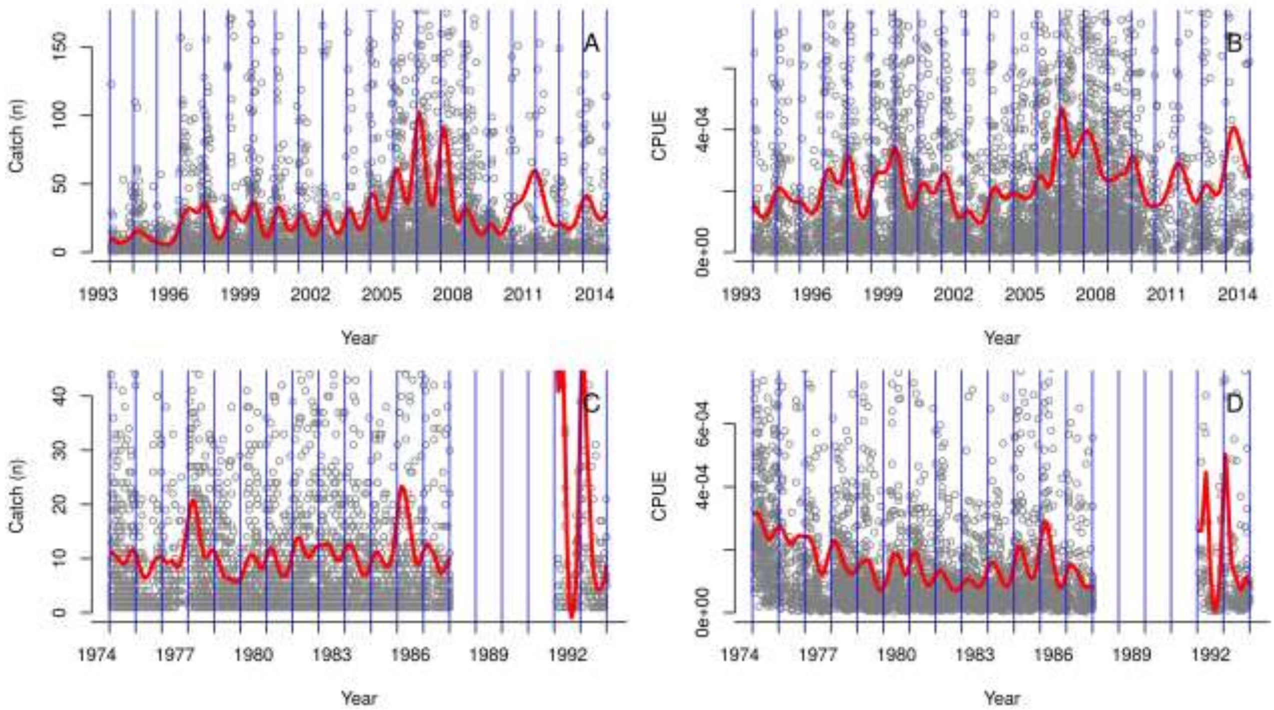

Figure 2 – Catch (number of fish) and CPUE (fish/hook) of Indo Pacific sailfish as calculated based on Japan (A and B) and Korea (C and D) datasets.

Catches as reported in the Korea database were usually low close to the end of the years, but there was not consistent seasonal signal all over the time series (Figure 2 C). In 1992 and 1993 the variability of catch were high with outstanding peaks and plunges. Variations of CPUE of Korea showed seasonal signal in most of the years (Figure 2 D). Peaks and plunges in 1992 and 1993 also appeared in the CPUE time series. Because the variations of catch and CPUE in those two years are quite different, and because there is a four years gap between them and the rest of the time series, data of 1992 and 1993 were discarded in the following analyses.

Spatio-temporal Distribution

Korea

Maps of effort and CPUE of SFA as reported in the Korea dataset are in Figures 3 and 4. Overall most of the longline were set in an equatorial "belt" bounded by 10º N and 10º S latitudinal parallels (Figure 3). In this equatorial area the effort were more and more concentrated in the west as the years went by. In the end of the time series the effort in the east margin of Indian ocean (longitudes eastward of 70ºE) reached very low values. In the 1970's a part of the longline sets were located in the southwest Indian Ocean close to Madagascar, but that southern fishing ground was not explored after 1980.

High CPUE values as calculated based on Korea dataset were scattered all over the Indian Ocean (Figure 4). However, the high values were more often found close to India coast, in the Bay of Bengal, and close to Madagascar, most in the Mozambique Channel whenever the longlines were set there.

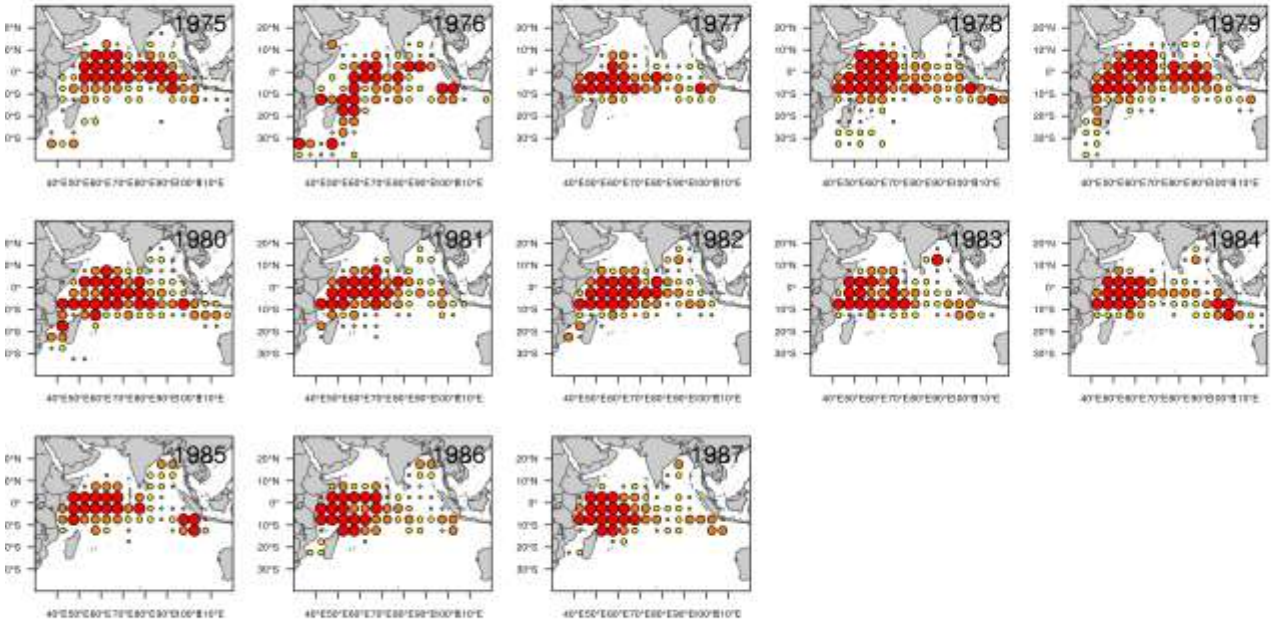

Figure 3 – Effort distribution across the years as reported in the Korea database.

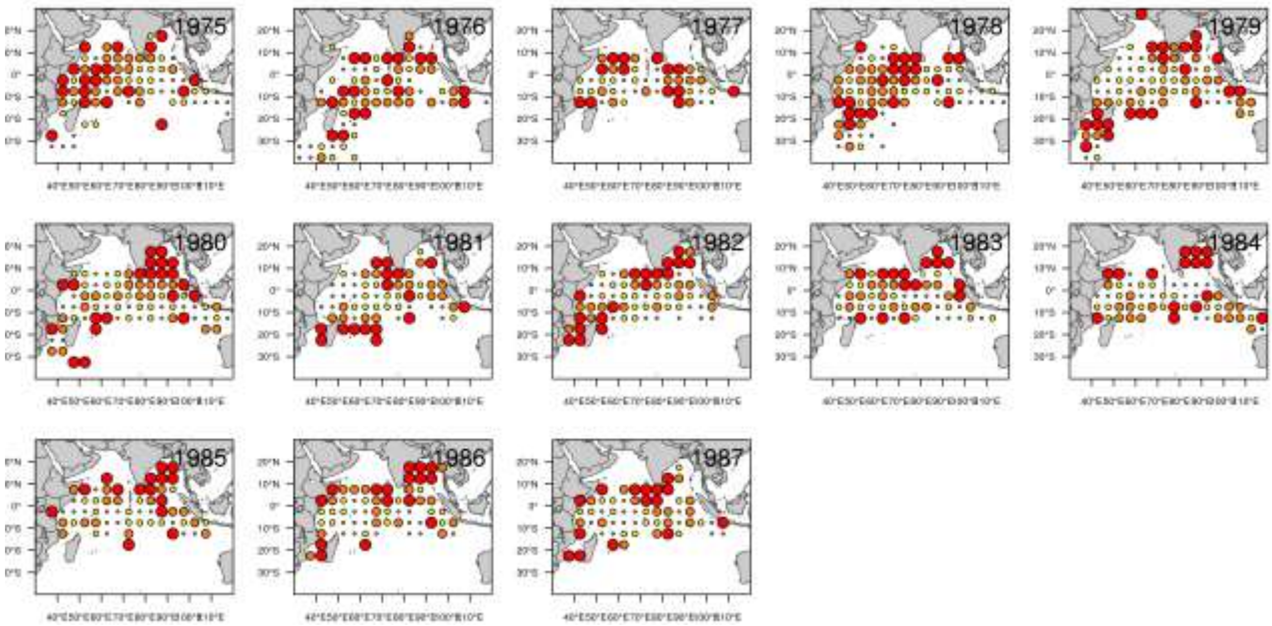

Figure 4 – Catch-per-Unit-Effort distribution across the years as reported in the Korea database.

Spatial distributions of effort and CPUE across the months are summarized in Figures 5 and 6. In most of the months effort was concentrated only in the west of equatorial area, but there were also a focus of effort in the east in November and December (Figure 5). July and August also calls attention because the area in which the effort was high extends all over the equatorial Indian Ocean, from African to the Indonesia coast.

High CPUE were found scattered all over the Indian Ocean in all months (Figure 6). Notice that high values were often found close to India and Sri Lanka, specially from January to April. It is also remarkable that high CPUEs often occurred around Madasgascar all year round.

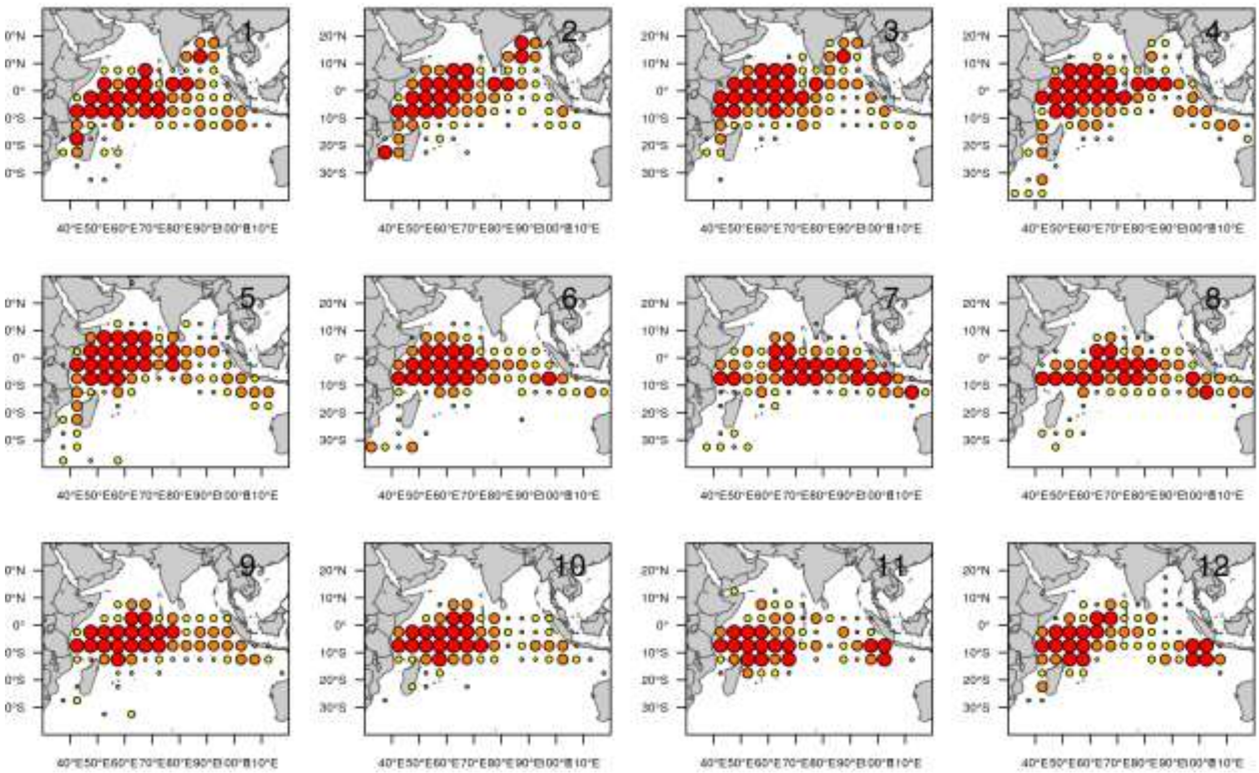

Figure 5 – Effort distribution across months (all years aggregated) as reported in the Korea database.

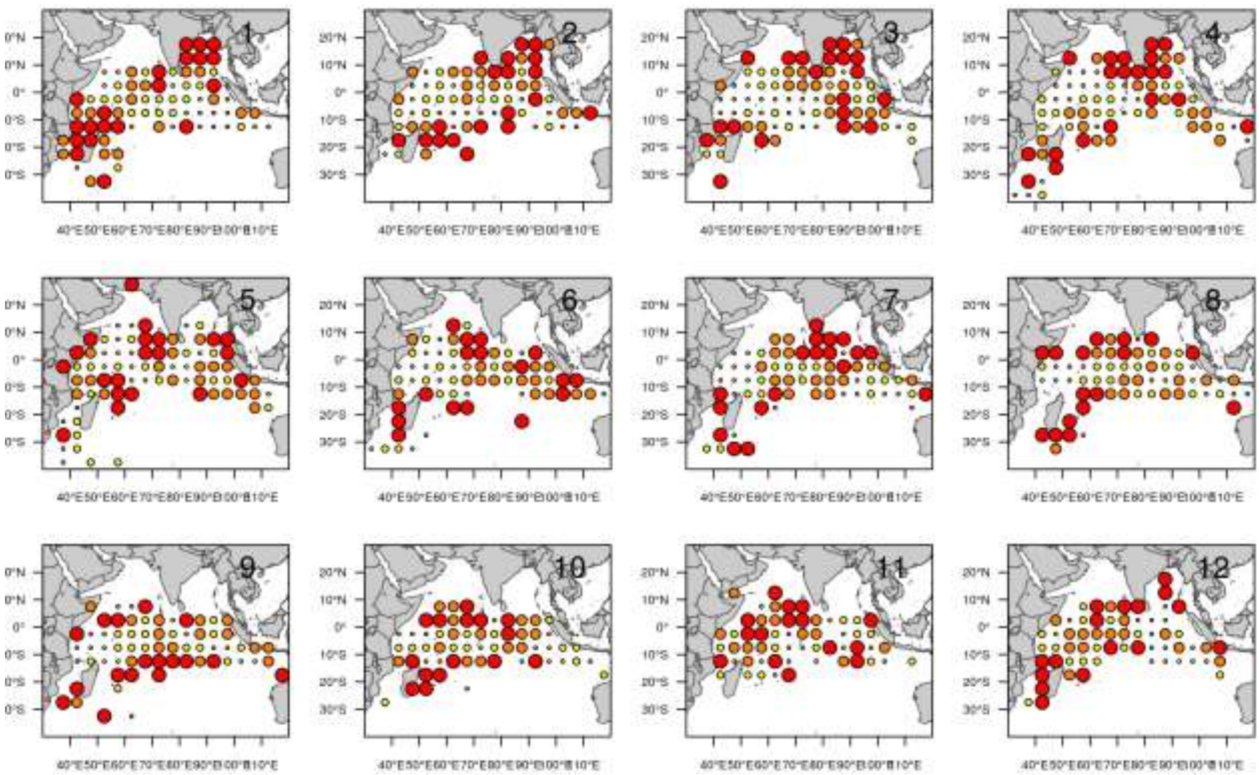

Figure 6 – Catch-per-Unit-Effort distribution across month (all years aggregated) as reported in the Korea database.

When all the data were aggregated the core area in which the Korea effort was concentrated arises, as indicated by orange an red filled circles in Figure 7 A. In this figure the polygon indicates the area considered for further analyses. The sample size calculated as the number of longline sets, were low outside the “core” area. When one looks at the CPUE inside the core, an heterogeneous pattern appear (Figure 7 B). Hence the core was split into four subareas. The CPUE were high CPUEs in areas 1 (Bay of Bengal, south India and Sri Lanka) and 4 (North of Madagascar and Mozambique Channel), intermediate in area 2 (margin of area 1) and area 3 (west and mid of the Indian Ocean equatorial waters). Catchabilities or abundances of sailfish in the four areas are supposed to be different.

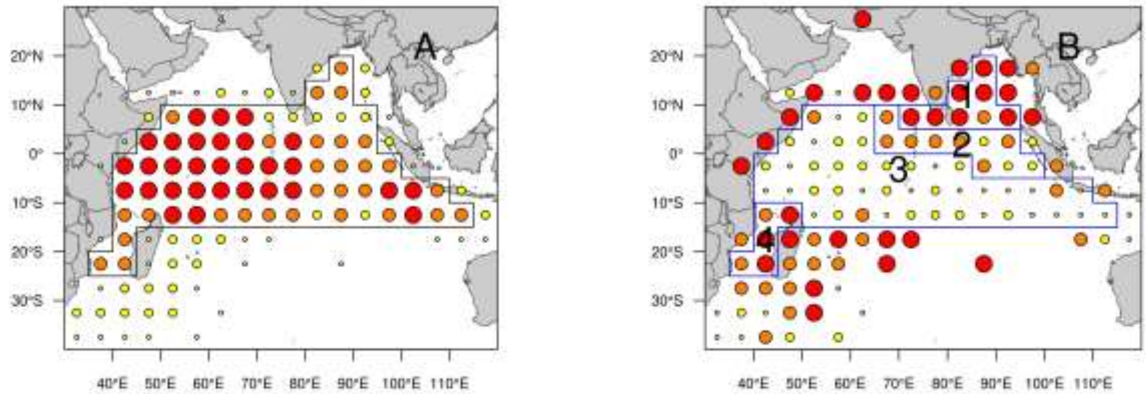

Figure 7 – Effort (A) and catch-per-unit-effort (B) distributions as reported in the Korea database. All data aggregated (1975-1987). Polygons indicate the effort core area in panel A, and area subdivisions in panel B.

A summary of the main results found when analyzing the Korea dataset is warranted. There is not a consistent seasonal pattern in catch rates across the years. Because the effect of month changes from year to year, it may worth the effort to include interactions between those two factors in the models. Area is probably the main factor driving CPUE outcomes inside a given year. Distributions of effort across months, areas and years are not very well balanced, even though the data seem useful.

Japan

In most of the years effort distribution of Japan covered a wide area encompassing the equatorial, central and the western margin of the Indian Ocean (Figure 8). The widest displacement of effort occurred from the end of 1990’s to end of 2000’s. In the beginning of the time series from 1994 to 1996 there were some gaps in the central equatorial region. After 2010 the covered area shrank and effort was concentrated mainly in two fishing spots, the Mozambique Channel, and the central-east part of the Indian Ocean (Figure 8).

Most of the high CPUE values occurred in the west and in the north of the Indian Ocean (Figure 9). Bay of Bengal, the area nearby India and Sri Lanka coasts and the Mozambique Channel arise as fishing spots with high catchability or abundance of sailfish all over the years. In opposition the CPUE values were often low in the central and east part of the Indian Ocean.

Efforts of Japan boats were mostly concentrated in the west margin of Indian Ocean from January to May, and in November and December (Figure 10). Bay of Bengal and north part of the Indian Ocean around Sri Lanka and nearby India were fished mainly in the beginning of the year. In the mid of the year fishing effort was not high close to equatorial coast of Africa, like in the beginning and in the end of the year. Central west area and south part of Mozambique Channel were the main fishing grounds from July to October (Figure 10).

High CPUE values occurred scattered all over the west and north of Indian Ocean, but once again those high values were more often found south of India, nearby Sri Lanka, in the Gulf of Bengal and close to the north of Madagascar and of the Mozambique Channel (Figure 11). High CPUE estimations were more often found from January to April, but high values also occurred all year round. Therefore, the seasonal pattern is unremarkable.

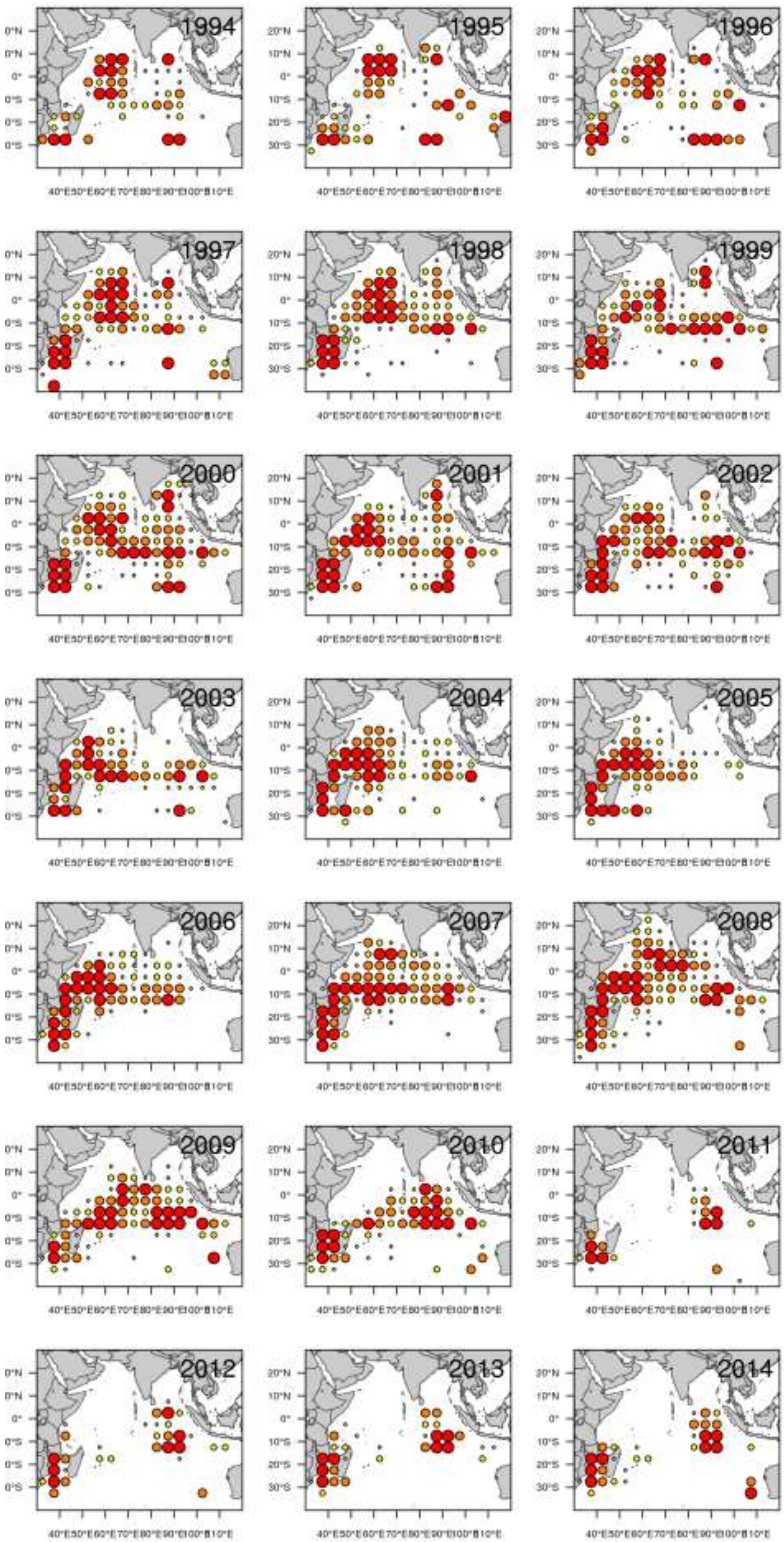

Figure 8 – Effort distribution across the years as reported in the Japan dataset.

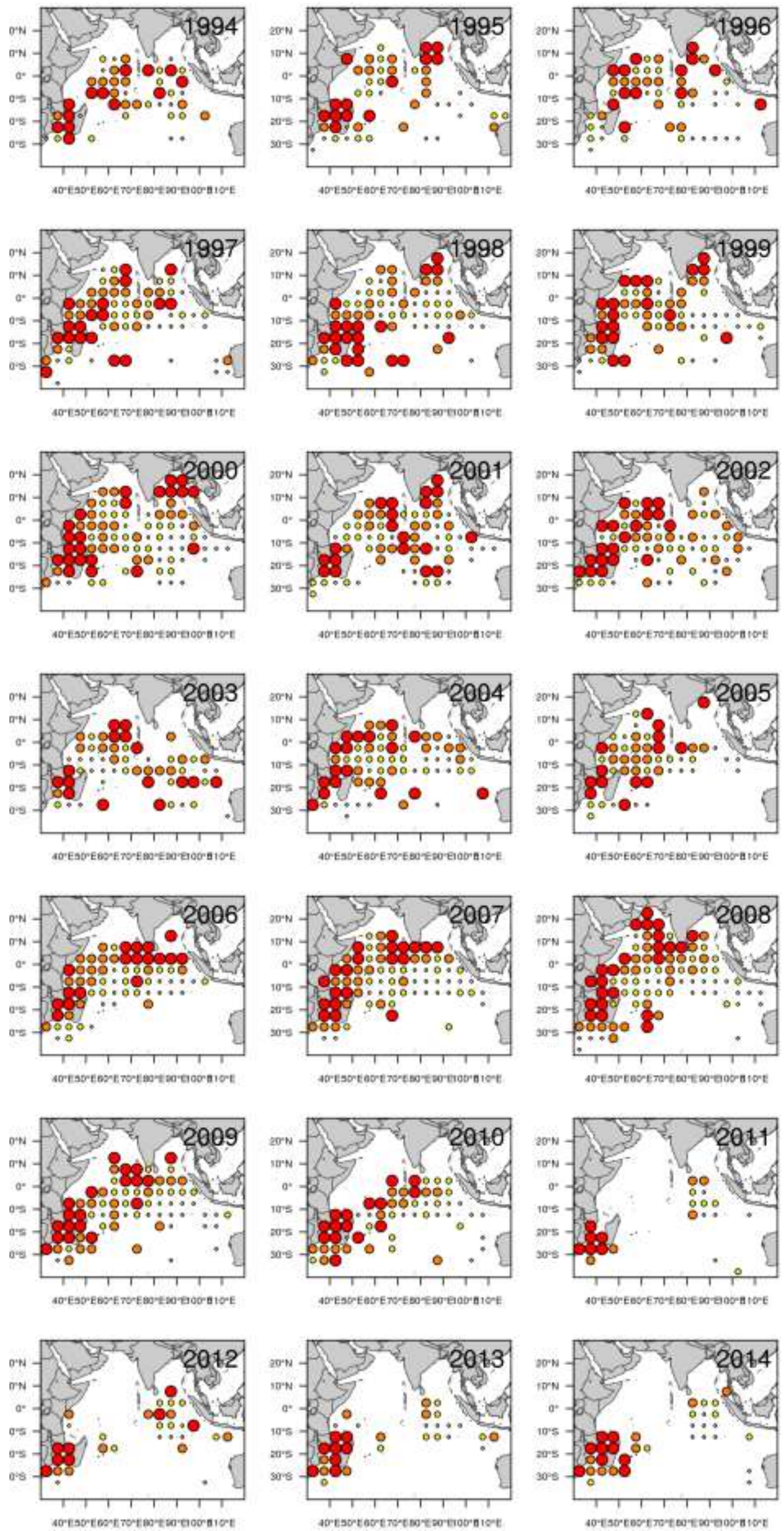

Figure 9 – Catch-per-Unit-Effort distribution across the years as reported in the Japan dataset.

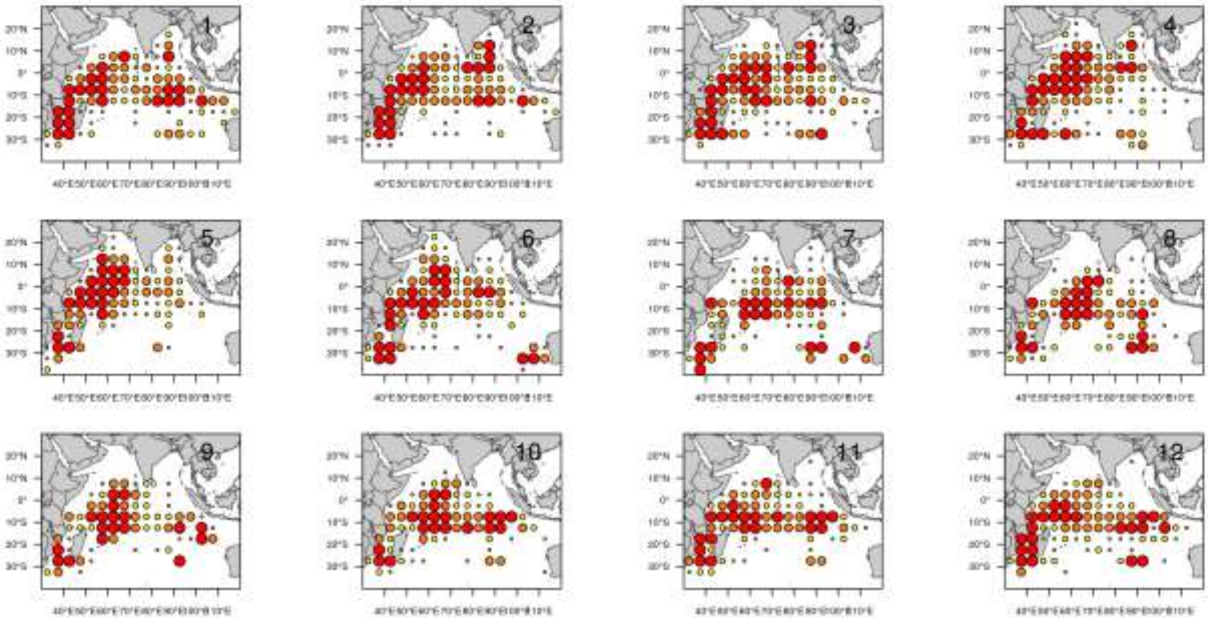

Figure 10 – Effort distribution across months (all years aggregated) as reported in the Japan database.

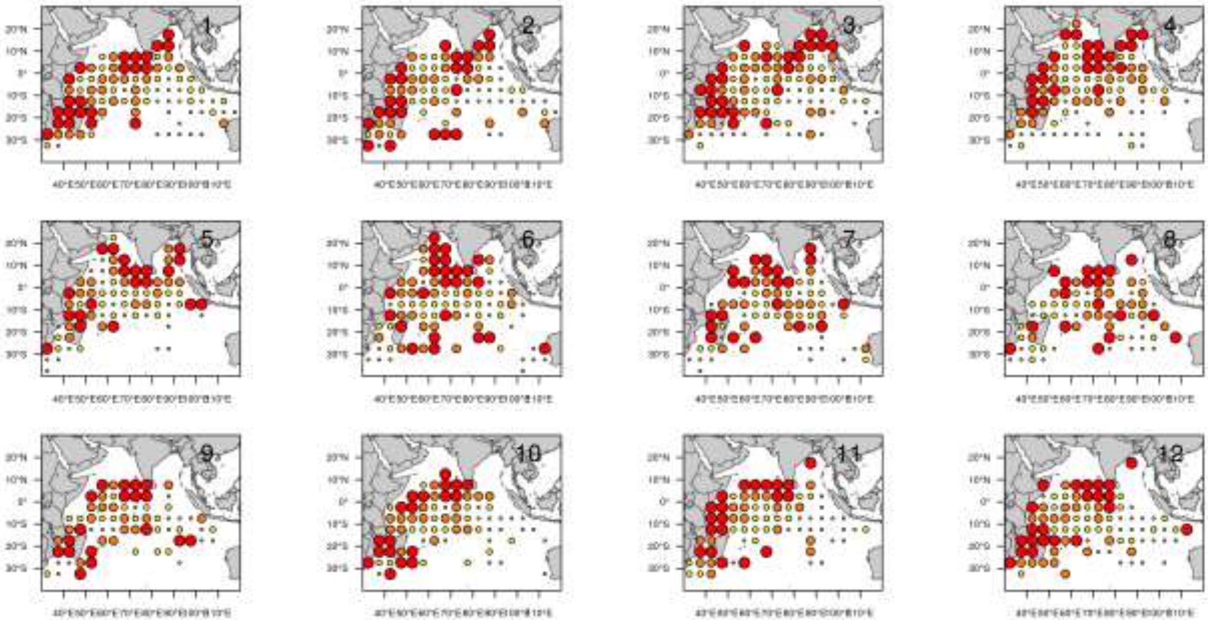

Figure 11 – Catch-per-Unit-Effort distribution across month (all years aggregated) as reported in the Korea database.

Total effort and CPUE as calculated with aggregated data are shown in Figure 12. The polygon in Figure 12 A indicates the core area selected for the analysis, and the smaller polygons in Figure 12 B stand for subdivisions of the core motivated by the differences of CPUE. Effort of Japan was concentrated in equatorial area (10ºN – 15ºN) and in the Mozambique Channel (Figure 12 A). Efforts were low outside the selected area except in the very south part of the Mozambique Channel and in a small area between 25ºS and 30ºS close to 90ºE (Figure 12 A). The later area was not selected because it is tiny and far from the main core. The two squares in the south of Mozambique Channel were not included in the selected area because the CPUE were different from the rest of Mozambique Channel (Figure 8 B). The five subdivisions selected for analyses were: 1 – Bay of Bengal, Sri Lanka and south and southwest of India; 2 – equatorial central part of Indian ocean; 3 – western equatorial area; 4 – north of Madagascar and Mozambique Channel; and 5 – eastern part of Indian Ocean (Figure 12 B).

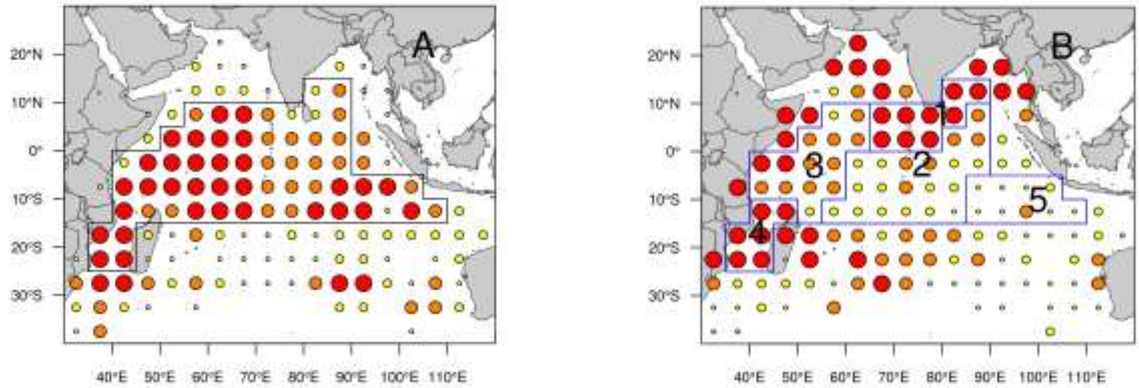

Figure 12 – Effort (A) and catch-per-unit-effort (B) distributions as reported in the Japan database. All data aggregated (1994-2014). Polygons indicate the effort core area in panel A, and area subdivisions in panel B.

Target proxy

Korea

Proportions of yellowfin, bigeye and albacore in the total catch as reported by Korea are shown in Figure 13. Proportions of the three species were similar in “catch-effort” and “nominal catch” databases. Yellowfin and bigeye summed up to almost all the catches all across the years. Proportions of both species were on average close to 0.45 from 1970’s to 1990’s. Proportions of yellowfin were also high in 2000’s, but the proportion of bigeye decreased quickly in the end of the time series. Proportion of albacore were close to 0.15 on average in the 1960’s and in the beginning of the 1970’s, but decreased to very low values after 1974. In the very end of the time series the proportion of albacore peaks to 0.3.

In the time span of interest (1975-1987) proportions of yellowfin and bigeye were steadily close to 0.45. Hence, variations of the proportions of the two species will probably not very useful as proxies of changes of fishermen targets. However, changes in the proportions of albacore in 1975-1987 were remarkable, though the proportions were always below 0.2. Proportion of albacore may worth the effort as proxy of the Korean fishermen strategy.

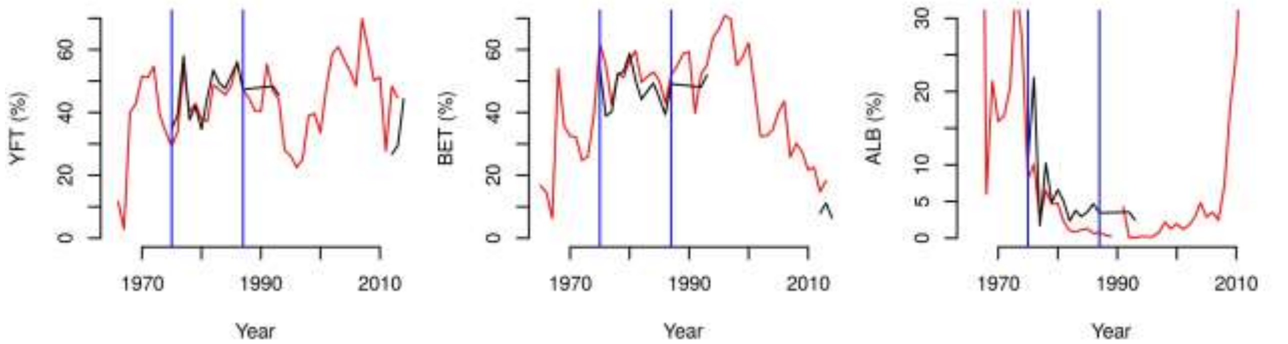

Figure 13 – Proportion of yellowfin (YFT), bigeye (BET) and albacore (ALB) in the total catch as reported in the Korea datasets "catch-effort" (black line) and "nominal catch" (red line). Vertical blue lines indicate the period (1975-1987) with informative data about sailfish.

Japan

Proportions of yellowfin in the Japan datasets decreased from 1950's to 1990's, increased from 1995 to 2003, and decreased again in the end of the time series (Figure 14). The participation of Bigeye in the total catches increased steadily from 1950 to 1977, but decreased slightly from 1978 to 1992, and then more quickly after 1993. Proportion of albacore increased fast from 1950 to mid 1960's, but decreased until the end of 1970's. Participation of albacore did not change much in 1980's and 1990's, but increase in the end of the time series.

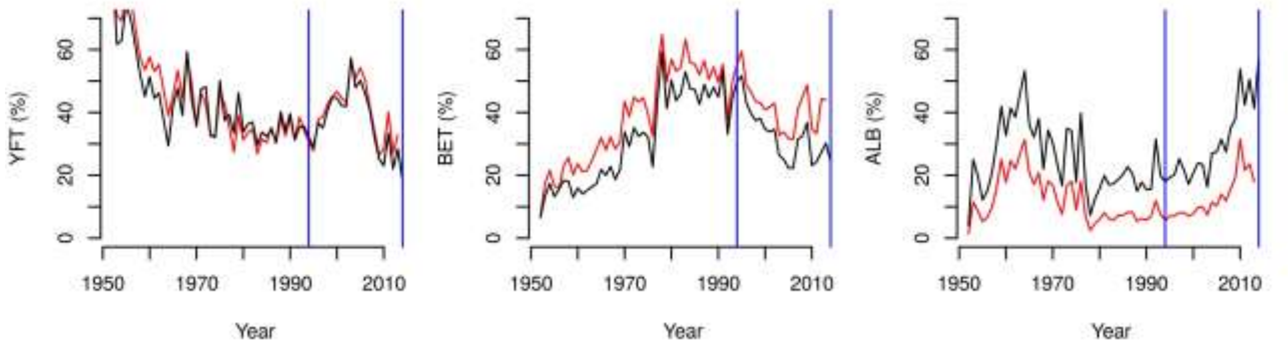

Figure 14 – Proportion of yellowfin (YFT), bigeye (BET) and albacore (ALB) in the total catch as reported in the Japan datasets "catch-effort" (black line) and "nominal catch" (red line). Vertical blue lines indicate the period (1975-1987) with informative data about sailfish.

Proportions of yellowfin as reported in "nominal catch" and in "catch-effort" databases were similar all across the years. However proportions of bigeye as reported in "catch-effort" database was higher than in "nominal catch" database. The opposite pattern arises in the calculation for albacore. Differences between the proportions of albacore as calculated for the two databases are not low, but overall time trends showed by the two series were not quite different. In this sense the behavior of the fraction of the fleet that reported catch-effort is similar to the whole fleet behavior.

Sailfish data is available from 1994 to 2014. In this period proportions of yellowfin peaked in mid 2000's, but decreased fast in the end of the time series. In opposition, the proportion of bigeye has decreased until mid 2000's, but it has increased slightly from 2005 to 2014. Proportion of albacore has increased monotonously from the beginning of 1990's to 2014. There where strong contrasts of proportions of the three tuna species, which indicate that targets have been changing in a very dynamic and fast fashion.

Selected Models

Korea

Fixed effects models did not converged when the proportion of albacore where included as explanatory variables, probably due to the high number of missing data for albacore. However the model has converged when albacore proportion and some interactions were dropped. Explanatory variables included in the selected model and the analysis of deviance is shown in Table 1. All main effects were included in the model. If we rely in the proportional reduction of deviance, the Area arises as one of the most important factors to explain the variability of the CPUE, though the effects of year and month were also of note. Among interactions those including year are the more important to explain the variability of CPUE, but notice they the number of parameters to be calculated is high when Year is in the interactions. Overall the proportion reduction of deviance of the selected model was 0.425.

Table 1 – Analysis of deviance for the fixed effects model fitted to the Korea database. pYFT and pBET stand for the proportions of yellowfin and bigeye in the total catch.

	Df	Deviance	Resid. Df	Resid. Dev	F	Pr(>F)
NULL	NA	NA	3281	4445.65	NA	NA
Year	12	360.69	3269	4084.96	25.85	3.26E-56
Month	11	280.03	3258	3804.92	21.89	1.48E-43
Area	3	474.16	3255	3330.77	135.93	1.09E-82
pBET	1	47.03	3254	3283.74	40.44	2.32E-10
pYFT	1	2.10	3253	3281.64	1.81	0.179032313
Year:Month	124	520.32	3129	2761.32	3.61	9.54E-35
Year:Area	34	105.93	3095	2655.39	2.68	5.45E-07
Year:pBET	12	44.13	3083	2611.26	3.16	0.000166836
Month:pYFT	11	38.76	3072	2572.50	3.03	0.000487205
pBET:pYFT	1	14.82	3071	2557.68	12.74	0.000362982

Diagnostic plots of residuals for the fixed effect model are shown in Figure 15. The model fitted seems unbiased and the residual distribution seems homocedastic. Notice also that standardized residuals distribution is approximately normal, though there left tail is of concern. However, there are not outliers or very strong influential data points.

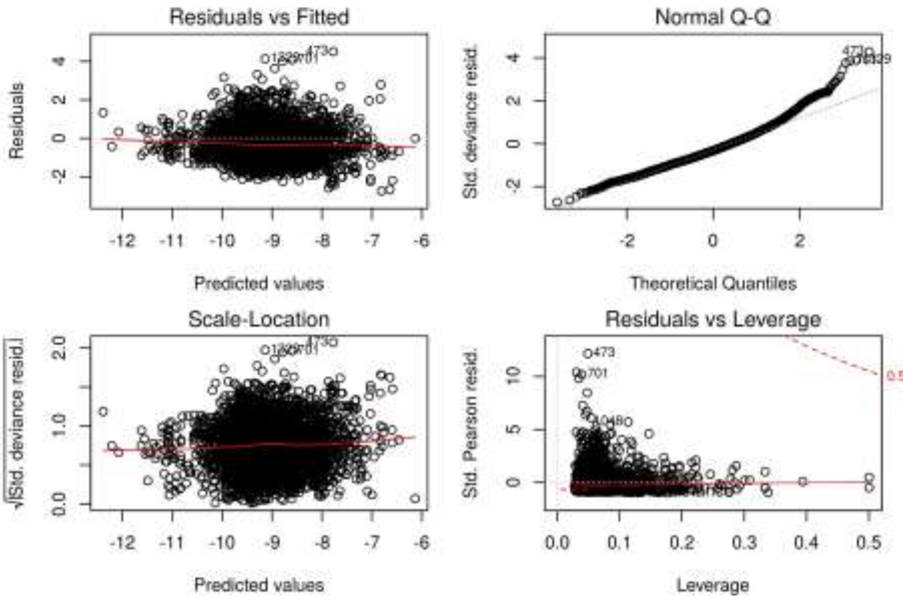

Figure 15 – Residuals standard diagnostics plots for Korea dataset and year interactions as fixed effects.

Mixed effect model did not converge when all the variables and interactions were included in.

However, convergence was achieved when the proportion of yellowfin and some interactions were dropped off. Several different nested models were compared based on AIC. Type II chi-square tests for the sequential inclusion of the factors as calculated for the selected model are shown in Table 2. Notice that most of explanatory were kept in the model as indicated by AIC, though the main effect of proportion of albacore was not significant in the hypothesis test. However the proportion of albacore is in a significant interaction.

Table 2 – Chi-squared test type II for the mixed effects model fitted to the Japan database. pALB, BET and pYFT stand for the proportions of albacore, yellowfin and bigeye in the total catch.

	Chisq	Df	Pr(>Chisq)
Year	2573.65	13	0^+
Area	346.48	3	8.64E-75
Month	30.16	11	0.001494226
pBET	49.12	1	2.40E-12
pALB	0.35	1	0.552849648
pBET:pALB	33.90	1	5.81E-09

Standard diagnostic plots of residuals for the mixed effect model are showed in Figure 16. Computational calculations of influential and discrepant measurements (e.g leverage) for mixed effect models are very time demanding hence they were not carried out. Results showed in Figure 16 indicate that the fitted model seems unbiased and that the residual distribution seems homocedastic. Notice also that standardized residuals distribution is grossly symmetrical, but normality assumption was violated specially in the left tail.

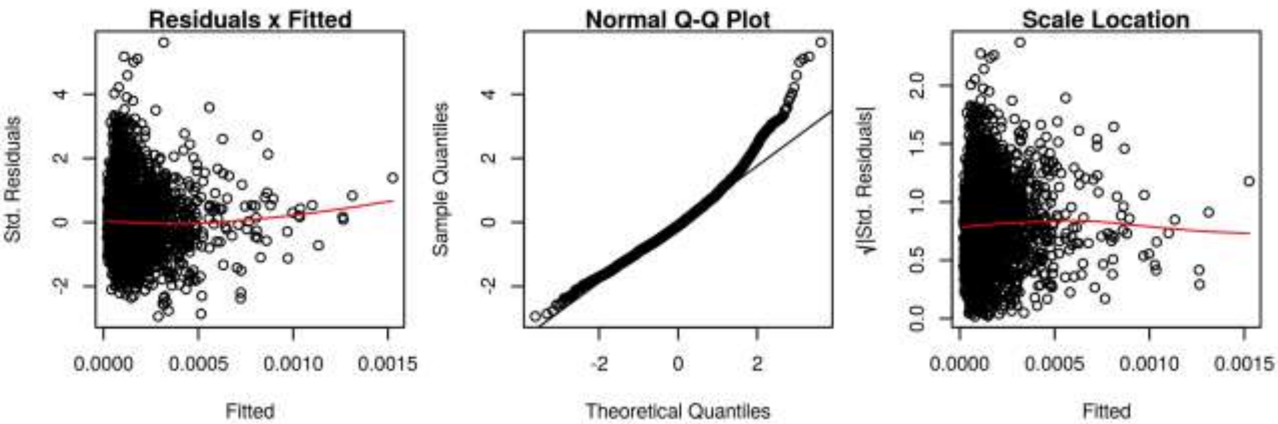

Figure 16 – Residuals standard diagnostics plots for Korea dataset and year interactions as fixed effects.

Japan

Analysis of deviance and exploratory variables included in the selected fixed effect model fitted to Japan database are shown in Table 3. Area was the main factor to explain variability of SFA CPUE if we rely on proportional reduction of deviance as criterion. Effects of year, month and proportion of yellowfin were weak. Effects of proportions of the other tuna species were even lower. Overall the effects of interactions were also weak. Proportion reduction of deviance of the selected model was 0.471.

Table 3 – Analysis of deviance for the fixed effects model fitted to the Japan database. pYFT, pBET and pALB stand for the proportions of yellowfin, bigeye and albacore in the total catch.

	Df	Deviance	Resid. Df	Resid. Dev	F	Pr(>F)

NULL	NA	NA	3684	6458.98	NA	NA
Year	20	363.38	3664	6095.59	14.26	9.81E-47
Month	11	236.16	3653	5859.43	16.85	6.64E-33
Area	4	1764.22	3649	4095.22	346.17	5.72E-252
pBET	1	10.91	3648	4084.31	8.56	0.00345232
pYFT	1	150.05	3647	3934.26	117.77	5.14E-27
pALB	1	17.65	3646	3916.60	13.86	0.000200393
Year:Area	74	272.16	3572	3644.44	2.89	6.90E-15
Month:pBET	11	64.96	3561	3579.48	4.64	4.70E-07
Area:pYFT	4	97.07	3557	3482.40	19.05	1.64E-15
pBET:pALB	1	64.26	3556	3418.15	50.43	1.48E-12

Residual diagnostic plots indicate that the model is acceptable in the sense it is not biased (Figure 17). Residuals distribution was approximately symmetrical though normality assumption was violated. Residuals were homocedastic, and there are not outliers or strong influential points.

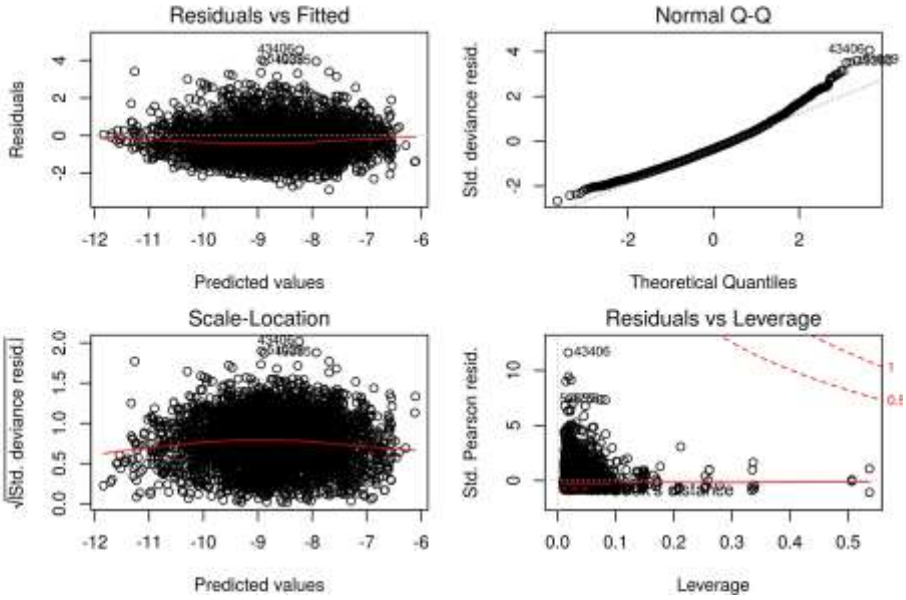

Figure 17 – Residuals standard diagnostics plots for Japan dataset and year interactions as fixed effects.

Likewise in the Korea analysis, convergence of mixed effect model for the Japan database was not straightforward. Nevertheless, models have converged when all fixed effect interactions and random interaction between year and month were dropped off. The terms of the selected model as well as the chi-square tests components are shown in Table 4. All main effects were kept in the selected model based on AIC calculations. However, the proportions of bigeye and of yellowfin were not significant whenever we rely in the chi-square tests.

Table 4 – Chi-squared test type II for the mixed effects model fitted to the Japan database. pALB, BET and pYFT stand for the proportions of albacore, yellowfin and bigeye in the total catch.

	Chisq	Df	Pr(>Chisq)
Year	224.20	21	5.90E-36
Area	495.24	4	7.18E-106
Month	68.39	11	2.46E-10
pBET	1.07	1	0.301791451
pYFT	0.01	1	0.92976248
pALB	14.84	1	0.000117333

Diagnostic plots of residuals for the mixed effect model showed in Figure 18 indicate that the fitted model seems unbiased and that the residual distribution seems homocedastic. However the residuals distribution can not be fairly approximated by the normal distribution.

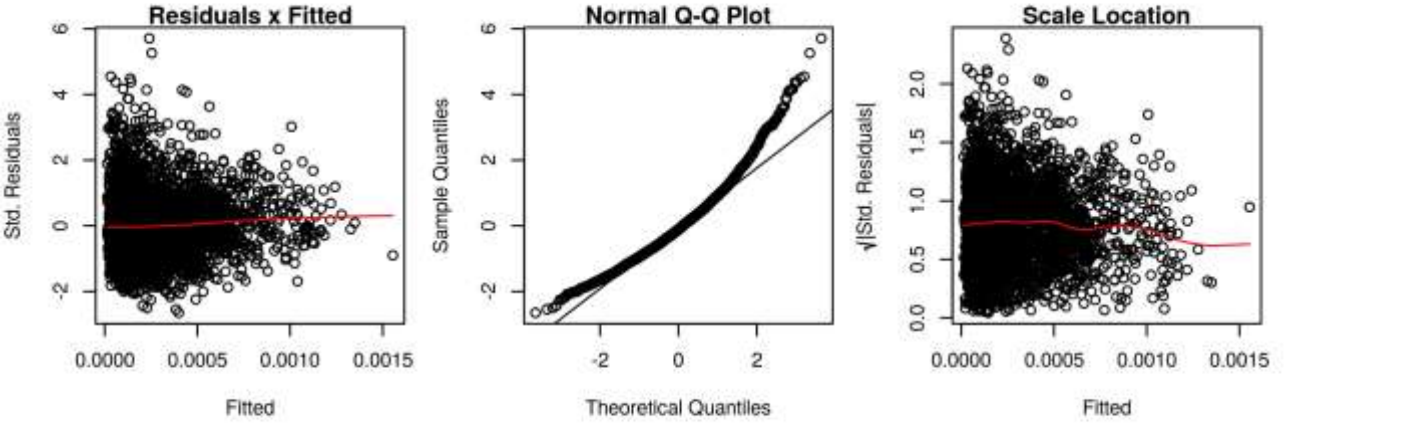

Figure 18 – Residuals standard diagnostics plots for Korea dataset and year interactions as fixed effects.

Standardized CPUE

Korea

Standardized and nominal CPUE calculations were scaled by dividing them by their means in order to make comparisons easier. Estimations of standardized and nominal CPUE as calculated for Korea dataset and fixed effect model are shown in Figure 19 A. Nominal CPUE has decreased from 1975 to 1983, but increased until 1986 and decreased in 1987. Time trend line of nominal CPUE is smooth, while the standardized CPUE shows peaks and plunges. Overall standardized CPUE time trend is similar to that of the nominal CPUE from 1978 onwards, but the nominal CPUE is high in the beginning of the time series, while the standardized CPUE estimations for 1975 and 1976 were low. Dotted lines in Figure 19 stand for simple linear regressions to make easier to figure general time trends of CPUE series. Both, nominal and standardized CPUEs point for decreasing trends though the slope calculated for nominal CPUE is more negative.

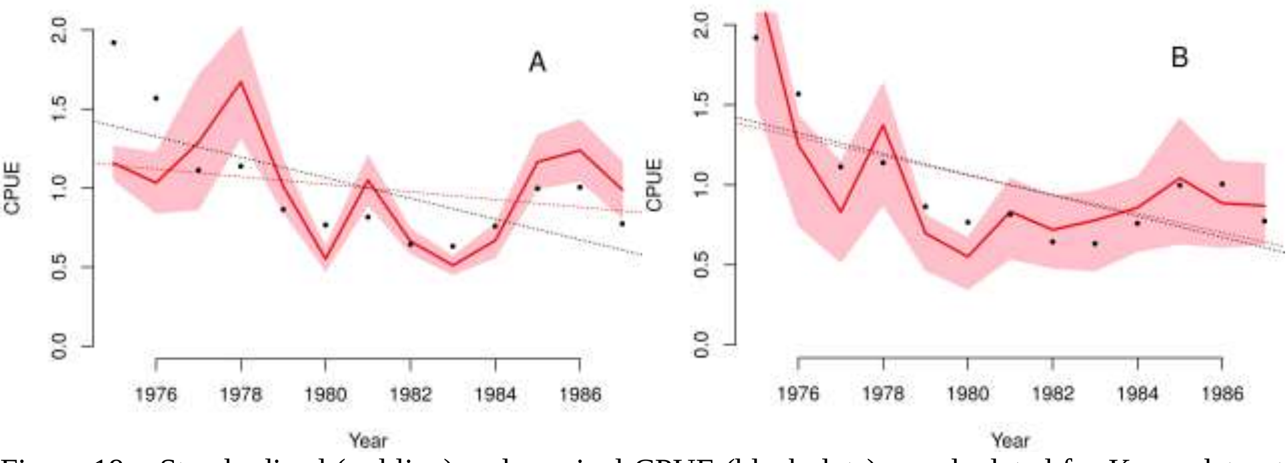

Figure 19 – Standardized (red line) and nominal CPUE (black dots) as calculated for Korea dataset using fixed (A) and mixed (B) effect model. Polygon filled pink stands for the 95% confidence interval of the estimation of standardized CPUE. Dotted lines represent simple linear regressions fitted to standardized (red line) and to nominal CPUE (black line).

Estimations of standardized CPUE as calculated using mixed effect model are showed in Figure 19

B. Confidence intervals were wider as calculated using mixed model. The main differences between the estimations using fixed and mixed models appear in the beginning of the time series (1975 and 1976) (Figure 19). In the beginning of the time series estimations calculated using mixed effect models were high in the beginning of the time series. Standardized CPUE as calculated using mixed effect standardized CPUE had decreased across the years like the nominal CPUE. Negative slopes of the regressions fitted to standardized and to nominal CPUE estimations were similar.

Japan

Estimations of standardized and nominal CPUE as calculated for Japan dataset using fixed effect model were similar all across the years (Figure 20 A). The two CPUE time series showed an increasing trend from 1994 to 2008. After 2009 the CPUE decreased and reached minimum values in 2011, but they have increased again in the recent years. Slopes of the linear regressions fitted to the standardized (fixed effect) and nominal CPUE were positive and they were similar.

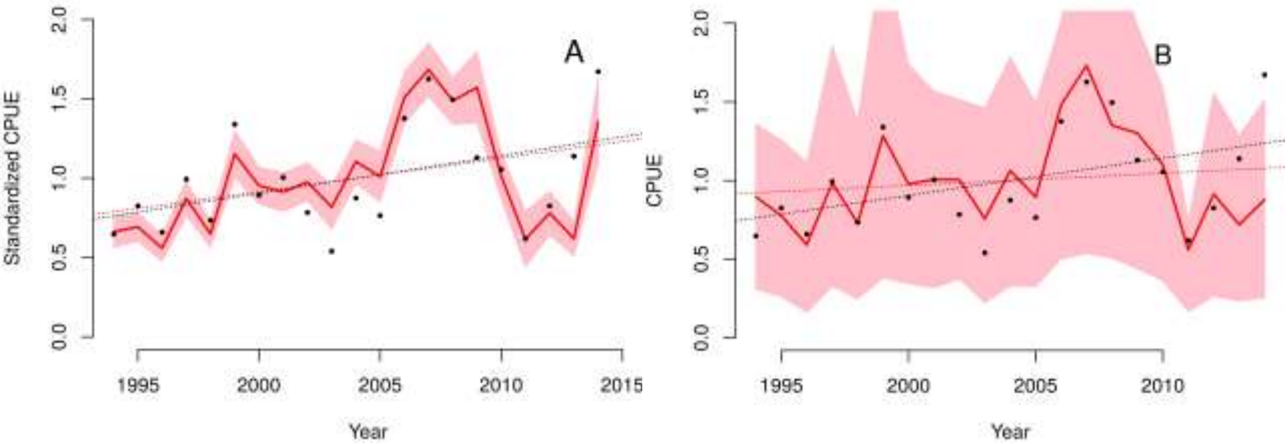

Figure 20 – Standardized (red line) and nominal CPUE (black dots) as calculated for Japan dataset using fixed (A) and random effect (B) model. Polygon filled pink stands for the 95% confidence interval of the estimation of standardized CPUE. Dotted lines represent simple linear regressions fitted to standardized (red line) and to nominal CPUE (black line).

Confidence intervals estimated using mixed effects model were wide (Figure 20 B). Overall time trends of fixed and mixed models were similarly positive. However, standardized catch rates calculated using mixed models did not increase in the end of the time series.

Discussion

Almost all fleets did not report catches of sailfish in 1970's and 1980's, but Korea is an exception. In that timespan Korean data of SFA seems consistent. Zero catches were not reported but probably most of all missing data were true zero catches. The argument in favor of this speculation is that if one assumes that empty cell were true zero catches, calculations of the proportion of positive seem consistent in the 1975-1987 time span. However, the quantity and the quality of the information decreased after 1987, and finally data were not reported in recent decades.

Most of fish caught by Korea longline fleet from 1975 to 1987 were yellowfin and bigeye. The proportions of the two species in the total catches were similar. In opposition the proportions of bigeye caught by Korean longline fleet in the Pacific were often higher than those of yellowfin from the end of 1980's onward (Moon and Know, 1996), because the regular longlines were gradually replaced by deep longlines in most of the fishing grounds after 1977 (Gong et al., 1989). However, the effect of the gradual change of regular to deep longline was not perceptible in the Indian Ocean in 1980's. The proportion of yellowfin decreased only in the beginning of 1990's when the proportion of bigeye had increased. Therefore, the time span 1975-1987 analyzed in this paper seems a relatively stable period, in the sense the fish target did not change much across the years. Efforts were systematically concentrated in the same areas and the variations of the proportions of the main tuna species (*i.e.* yellowfin, bigeye and albacore) did not draw attention. Fishing operational variables often used as target proxies (e.g. length of branch lines and number of hooks between floats) were not available. However, the apparent stability of the proportions of tuna species in the catches indicate that the modus operandi of Korea fleet did not change much in the period analyzed. Hence the relative abundance time series estimated is probably not strongly biased due to changes of fishing strategies in the 1975-1987 timespan.

Some billfish species (striped and blue marlin) were among the targets of Japanese longline fleets in the very beginning of the fisheries (Nishida et al, 2012). However, all billfish species have been bycatch during the last decades. The proportions of target species in the total catch and fishing operational procedures of Japan have changed much in the last decades (Nishida and Wang, 2014; Okamoto and Ijima, 2015). In addition there are also spatial differences concerning fishing operational characteristics. For example, the proportions of deep longline sets (i.e. high number of hooks between floats) are lower in the southwest part of the Indian Ocean (Okamoto and Ijima, 2015).

Estimations of CPUE standardization of billfish caught by longline Japanese vessels were calculated based on detailed set by set recently (e.g. Nishida et al., 2012; Nishida and Wang, 2014; Okamoto and Ijima, 2015). In some of the previous approaches operational characteristics (e.g. number of hooks between floats), environmental (e.g. temperature) or calculations based on the contributions of species to the total catch (e.g. cluster analysis) were used build the explanatory variables.

In the current year Okamoto and Ijima (2015) have used fishing operational characteristics and environmental as explanatory variables to calculate the very first estimations of standardized CPUE for sailfish based on a comprehensive set by set database. On the other hand, in this paper the calculations were based on aggregated dataset, on positive catches only, and no fishing operational characteristics were used as explanatory variables. Instead, the gross proportions of tuna species were the “target proxy” variables in the model. Calculations of Okamoto and Ijima (2015) based on the detailed database and with operational explanatory variables are probably less biased. However, comparative analyses of the results are warranted in order to uncover the differences and uncover the risks of using simple (maybe naïve) poor models to standardize the catch rates.

In this paper estimations were calculated for areas integrated, while Dr. Okamoto and Dr. Ijima kindly provided separated estimations of each of the three areas they have considered (see Okamoto and Ijima, 2015). Hence their estimations were averaged (equal weights by area) in order to make comparisons easier. All estimations available are shown in Figure 21. Nominal and standardized estimations were remarkable similar from 1994 to 2007, and from 2009 to 2012. However, the estimation of 2008 as calculated set by set database (Okamoto and Ijima, 2015) was higher than those calculate using the aggregate dataset. In the end of the time series the nominal CPUE and the estimations calculated using set by set database are very much conflicting, the former had increased while the later had decreased. Standardized estimations as calculated using aggregated dataset were in between the nominal and the “set by set” calculations in 2013 and 2014. Overall time trends of all CPUE calculations were very much the same before 1995. However, there were important differences in the end of the time series. The 2008 peak and the low estimations of 2013 and 2014 as calculated using the set by set database rendered very strong negative slope of CPUE in the end of the time series. In opposition the nominal and the calculations base on aggregated dataset indicate an increasing time trend of CPUE in the recent years. As far as the set by set calculations are less biased the outstanding differences among the calculations is a warning on the risks of using standardized CPUE as calculated based on aggregated databases, at least in the case of Japan longline fisheries, which have experienced major changes concerning operational characteristics and targets.

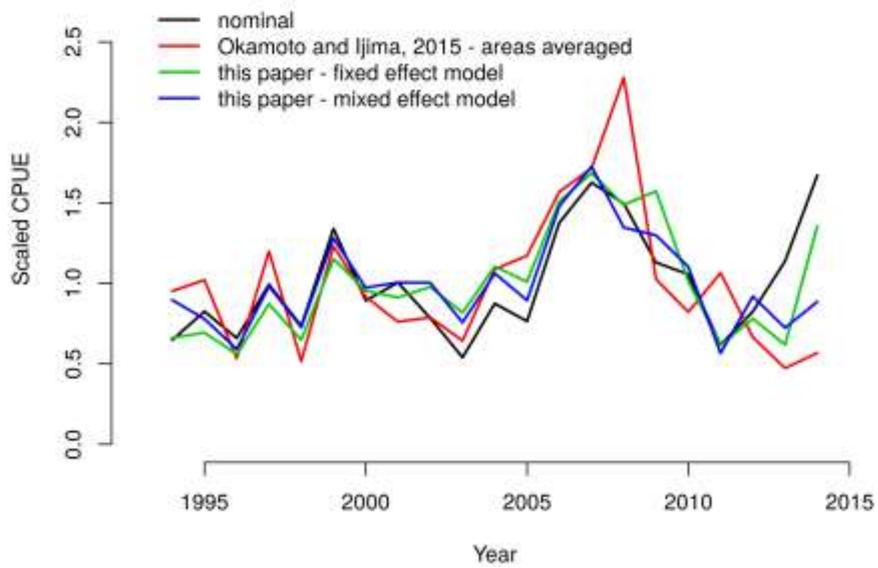

Figure 21 – Estimations of catch-per-unit-effort (CPUE) of Indo Pacific sailfish caught by longline Japanese fleet. Black line stands for the gross nominal CPUE, while the other lines stand for standardized CPUE calculated using set by set dataset (red), and aggregated dataset (green and blue).

Acknowledgements

The author is debt with Dr. Hiroaki Okamoto and Dr. Hirotaka Ijima who kindly provided before the meeting the estimations of standardized CPUE calculated based on set by set database.

References

Akaike, H. 1974. A new look at the statistical identification model. IEEE Transactions on Automatic Control, 19: 716-723.

Baum, J., and W. Blanchard. 2010. Inferring shark population trends from generalized linear mixed models of pelagic longline catch and effort data. Fisheries Research **102:**229–239.

Baum, J., R. Myers, D. Kehler, B. Worm, S. Harley, and P. Doherty. 2003. Collapse and conservation of shark populations in the Northwest Atlantic. Science **299:** 389–392.

DOBSON, A. J. 2002. An introduction to generalized linear models. 2nd Edition. Chapman & Hall/CRC. 225 pp.

Gong, Y.; Lee, L-U; Kim, Y. S. and Yang, W. S. 1989. Fishing efficiency of Korean regular and deep longline gears and vertical distribution of tunas in the Indian Ocean. Bull. Korean Fish. Soc. 22(2): 86-94.

IOTC. 2015. Catch and effort data. <<http://www.iotc.org/sites/default/files/documents/2015/07/IOTC- 2015-WPB13-DATA07_CELongline.zip>> *Accessed* (July 31, 2015).

Maunder, M. N. and Punt, A. E. 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res. 70: 141-159.

McCulagh, P. and Nelder, J. A. 1989. Generalized Linear Models. Chapman and Hall, London. 513p.

Moon, D. Y. and Kwon, J. N. 1996. Korean tuna fisheries in the Pacific Ocean and interaction between the fisheries. In: Shomura, R. S.; Majkowski, J. and Harman, R. F. (Eds.). Status of Interactions of Pacific Tuna Fisheries in 1995. Proceedings of the Second FAO Expert

Consultation on Interactions of Pacific Tuna Fisheries.

Nishida, T.; Shiba, Y.; Matsuura, H. and Wang, S-P. 2012. Standardization of catch rates for Striped marlin (*Tetrapturus audax*) and Blue marlin (*Makaira nigricans*) in the Indian Ocean based on the operational catch and effort data of the Japanese tuna longline fisheries incorporating time-lag environmental effects (1971-2011). IOTC–2012–WPB10–19 Rev_2.

Nishida, T and Wang, S-P. 2014. CPUE standardization of swordfish (*Xiphias gladius*) exploited by Japanese tuna longline fisheries in the Indian Ocean using cluster analysis for targeting effect. IOTC–2014–WPB12–21.

Pilling G. M.; Apostolaki, P.; Failler, P; Floros, C.; Large, P.A.; Morales-Nin, B.; Reglero, P.; Stergiou, K. I. and Tsikliras, A. C. 2008. Assessment and management of data-poor fisheries. In: A. Payne, J. Cotter, T. Potter (eds) Advances in Fisheries science: 50 years on from Beverton and Holt, pp. 280-305. Blackwell Publishing, CEFAS.

Pinheiro, J. C. and Bates, D. M. 2000. Mixed Effects Models in S and S-Plus. Springer-Verlag. New York.

Richards, L. J. and Maguire, J-J. 1998. Recent international agreements and the precautionary approach: new directions for fisheries management science. Canadian Journal of Fisheries and Aquatic Sciences 55, 1545–1552.

Searle, S. R.; Speed, F. M.; and Milliken, A. G. 1980. Population marginal means in the linear model. An alternative to least squares means. The American Statistician. 34(4): 216-221.