Received: 3 September 2015

Estimation of appropriate reporting ratio for the blue shark caught by Japanese longliner in the Indian Ocean

Mikihiko Kai[1] and Kotaro Yoakawa[1]

[1]National Research Institute of Far Seas Fisheries, Fishery Research Agency
5-7-1 Orido, Shimizu-ku, Shizuoka 424-8633, JAPAN

Summary

This document paper presents an appropriate reporting rate for the blue shark caught by Japanese longline fishery in the Indian Ocean. New statistical approach was applied to choose the best available reporting rate (RR) for blue shark through comparisons of the catch rate between observer data and logbook data. The most appropriate reporting rate was chosen by AIC from the simulation study with the filtering data by the different RR, and the value was 54 %. The value is largely different from the previous study in the ICCAT (80 %). However, this result seems to be reasonable because extreme higher RR can lose the useful information on the logbook data for the CPUE standardization, by contrast, lower RR can include a large number of unreported catch data.

Keywords

Blue shark, Indian Ocean, observer data, Reporting Rate

1. Introduction

Blue shark, *Prionace glauca*, in the Indian Ocean is incidentally caught by Japanese tuna longline fishery. Longliner are obligated to provide logbook data. Japanese logbook data is one of the most important sources of information to estimate the catch-per-unit-of-effort (CPUE) due to the high coverage of wide spatial areas and the long time periods of data. However, reporting rate (RR) of shark species have a tendency to be lower than any other tuna species because sharks are less valuable than tuna species.

Filtering method (Nakano and Honma, 1996; Nakano and Clarke, 2006) is commonly used to remove the bias caused by the systematic missing data and under-reporting. Nakano and Honma (1996) suggested that logbook data more than 70 % of reporting rates (shark catch day/total days operated) included the catch of all sharks, and Nakano and Clarke (2006) suggested that high reporting rates (number of sets with sharks recorded/total number of sets) larger than 80 % filtering avoids large numbers of false zeros and provides the best fit to observer data for blue shark. The logbook data used in those analyses were come from Japanese longline fleet operating in the Atlantic Ocean.

The objectives of this document paper is to estimate an appropriate reporting rate of

Japanese longline fleet operating in the Indian Ocean.

2. Material and Methods

2.1 Data sources

Logbook data of Japanese longliners operating in the Indian Ocean from 1994 to 2013 were compiled by the National Research Institute of Far Seas Fisheries (NRIFSF) with observer data collected through national observer program of CCSBT. Set-by-set data used in this study includes information on number of cruise for each vessels, operational time (year, month, day), catch number, catch weight, amount of effort (number of hooks), number of branch lines between floats (hooks per float: hpf) as a proxy for gear configuration, location (longitude and latitude) of set by resolution of 1 × 1 degree square, names of species caught, and vessel identity.

2.2 Selection of data

The data selection was fundamentally same as those used by Kanaiwa et al. (2015). The data more northern than 25˚S and January-March was removed to keep consistency of fishery because most of the observer data were collected from the operational areas of Japanese tuna longliner targeting for southern Bluefin tuna.

2.3 Definition of reporting rate

Three types of reporting rate were defined as follows;

$$\mathrm{RR1} = \frac{\text{Number of sets with "sharks" recorded}}{\text{Total number of sets}} \qquad (1)$$

$$\mathrm{RR2} = \frac{\text{Number of sets with "blue sharks" recorded}}{\text{Number of sets with "sharks" recorded}} \qquad (2)$$

$$\mathrm{RR3} = \frac{\text{Number of sets with "blue sharks" recorded}}{\text{Total number of sets}} \qquad (3)$$

where the definition of RR1 is the same as that used by Nakano and Clarke (2006). Hereafter, RR1 is referred to as a RR which is used to filter the data for CPUE standardization.

2.4 Comparisons of data between observer and logbook

Three comparisons were conducted as follows;

- Number of cruise for the three types of reporting rate above were calculated for logbook data and observer data during 1994-2013. Reporting rate was changed by the intervals of 10 %.
- Correlation of nominal CPUE of blue shark were compared between two data set filtered by more than 10, 50, and 90 % RR.
- The following three values were calculated against different RR which were increased by the

interval of 1 %. (1) correlation coefficient of nominal CPUE of blue shark between two data set, (2) number of sets of logbook data, (3) proportion of the number of year to the total number of years of logbook data

- Comparison of annual changes of nominal CPUE of blue shark between logbook data and observer data during 1994-2013 that was filtered by more than 10, 50, and 90 % RR.

2.5 Estimation of appropriate reporting rate

Appropriate reporting rate (RR) is determined using the statistical approach. The procedure is as follows;

(1) The cruise of the logbook data was regarded as “type 1”, if the same cruise is included in the observer data. The remaining cruises were regarded as “type 0”.

(2) For the “type 0” data sets, the data is filtered by the different RR which is increased by the interval of 1 %.

(3) The cruise of the logbook data was regarded as “type 1” if the RR is more than the given value (e.g. 10%, 20% etc.).

(4) CPUE is standardized for the 100 data sets (each data sets have a different number of “type 1”) using the Zero-inflated negative binomial model;

$$Count\ process's\ catch = \beta_1 \cdot year + \beta_2 \cdot area + \beta_3 \cdot season + \beta_4 \cdot year:area + \beta_5 \cdot area:season + \beta_6 \cdot type + intercept + \varepsilon_1$$

$$False\ zero\ prob = \beta_7 \cdot year + \beta_8 \cdot area + \beta_9 \cdot type + intercept + \varepsilon_2$$

where B_1~B_7 are coefficients for each factors, “year” is a year effect from 1994 to 2013, “area” is a spatial effects where the area is separated into two areas eastern than 90˚E (area 1) and western (area 2), “season” is separated into season 1 (April-July) and season 2 (August-December), “year:area” and “area:season” are two way interaction of each factor, ε_1 is error terms followed by negative binomial model with log link function, and ε_2 is error terms followed by binomial distribution with logit link function.

(5) Akaike information criterion (AIC) (Akaike 1973) is used to choose the most appropriate RR from the 1-100% RR.

3. Results

For three types of reporting rates (RR1, RR2, RR3), the number of cruise for observer data tended to more increase than those for logbook data with higher reporting rate (Fig. 1). For the case of RR1, the number of cruise for logbook data increased with lower reporting rate (upper figure in fig.1). For RR2 and RR3, the number of cruise for logbook data were remarkably large when the reporting rates were 0-10% and 91-100%, while the number of cruise for observer data have a tendency to gradually increase when the reporting rates exceed more than 41-50% (middle and bottom figure in fig.1).

The relation of nominal CPUE of blue shark between logbook data and observed data seemed to be similar among three different RR (10%, 50%, 90%) except that the number of observer data is clearly different among them when the nominal CPUE of the logbook data is 0 (Fig. 2). These results suggested that it is possible to remove the unreported catch for blue shark. With regards to the 0 values of nominal CPUE for observer data, observer frequently records only a part of the whole operation, so that the data set can include the zero-catch.

A difference of the nominal CPUEs between logbook data and observer data become small with higher correlation of nominal CPUE (Fig. 3). However, there is a tradeoff between correlations of nominal CPUE and number of sets of logbook data. In addition, the proportion of the number of year to the total number of years of logbook data also decrease with the increase of the RR. These results indicate that higher RR may lose the more information of the important data for the CPUE standardization.

Annual changes of nominal CPUE of blue shark indicates that the data filtering with higher RR can remove some nominal CPUE of zero catch. However, it cause the loss of the information on some year (Fig. 4). The most appropriate reporting rate was 54 % (Fig. 5).

At the 54% filtering ratio of log-book data, detailed check of log-book data revealed two possibilities of difference of CPUE between filtered log-book and observer data. One is that no or rather few data is available for the filtered log-book data, those are ones for years of 1995, 1996, 1999, 2000, 2002-2004. In these years, blue shark CPUEs of filtered log-book data are zero and not comparable with those of observers. Another possible reason is that Japanese longliners targets in the south of 25S in the Indian Ocean not only southern bluefin tuna but bigeye and albacore tunas. Species composition of catch including blue shark are largely different between southern bluefin tuna targeting sets and other tuna targeting sets (Fig. 6). Because CCSBT observer data basically does not include data of sets targeting bigeye and albacore tunas, and ratio of blue shark is usually higher for southern bluefin tuna targeting sets than other tunas targeting sets, data of sets targeting bigeye and albacore tunas targeting sets should be excluded from the calculation of blue shark CPUE of filtered log-book for the comparison purpose. In Figure 7, nominal CPUE calculated from 54% filtering ratio and southern bluefin tuna targeting sets and CCSBT observer data are compared. Southern bluefin tuna targeting sets was assumed that sets of single consecutive operations with more than 90% of them recorded southern bluefin tuna catches. The values in years that no filtered log-book data were available are eliminated from this comparison. Values of nominal CPUE calculated from two different data sets are agreed quite well with each other,

4. Discussions

Intermediate value (54 %) of reporting rate (RR) was chosen by the statistical approach. This results seem to be reasonable because extreme higher RR can lose the useful information of the logbook data for the CPUE standardization, on the other hand, lower RR can include a large number of unreported catch data. Therefore, the most appropriate RR can provide the best available data for the CPUE standardization, although it is impossible to remove all the unreported catch data.

Regarding to the different values of RR between this study and precious studies (Nakano and Honma, 1996; Nakano and Clarke, 2006), we cannot conclude whether which is the best RR because the statistical approach is different and the data source come from different Ocean. In future, same statistical approach will apply for the Atlantic data. Presumably, similar results to this study would be gained from the perspective of the balance between data quality and quantity.

The distribution areas of blue shark is entire water of Indian Ocean, in particular, larger sizes tended to occur in equatorial and tropical regions (Rui et al. 2015). However, most of the data were collected from the southern parts of the Indian Ocean where blue shark is known to be abundant in higher latitudinal area than tropical (FAO, 1984). The distribution and the values of Japanese longline CPUE also support these facts (Hiraoka and Yokawa 2012). Therefore, the area coverage of observer data is considered to be enough for the validation of the logbook data (Fig. A1).

Comparison of nominal CPUE calculated from southern bluefin tuna targeting sets of the 54% filtered log-book data and CCSBT observer data agreed with each other (Fig. 7). Observed small difference of CPUEs between the filtered log-book data and CCSBT observer data should be due to the coverage of CCSBT observer data (about 7%). Even though the log-book data was filtered by shark reporting ratio, number of sets by the filtered log-book data usually larger than that by CCBST observer data. Considering these things written above, the log-book data filtered by the 54% shark reporting ratio is well representing the CPUE of blue shark.

The reporting ration of 54% use for the log-book data filtering was lower than that in the Atlantic (Nakano and Honma, 1969). This would be due to the fact that the observer data used in this study is limited to the one for CCSBT, and part of Japanese southern bluefin tuna longline sets only catches southern bluefin tuna and no shark bycatch (Itou, personal communication, for example, some of SBT targeting sets operating for southern bluefin tunas feeding for Krill aggregations with no bycatch).

Reference

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. *In* Petrov, B.N., Csaki, F. (Eds.) Second International Symposium on Information Theory, Budapest, Akademiai Kiado, pp 267-281.

FAO 1984. Prionace glauca (Linnaeus, 1758), 1849. *In Sharks of the world. An annotated and illustrated catalogue of shark species known to date. Part 2. Carcharhiniformes*. FAO SPECIES CATALOGUE. Vol.4, 521-524.

Hiraoka, Y. and K. Yokawa 2012. Update of CPUE of blue shark caught by Japanese longliner and estimation of annual catch series in the Indian Ocean. IOTC-2012-WPEB07-33.

Kanaiwa, M., Semba, Y. and K. Yokawa 2014. Standardized CPUE of blue shark in the Indian Ocean estimated from observer data in the period between 1992 and 2012. IOTC-2014-WPEB010-26.

Nakano, H. and M. Honma 1996. Historical CPUE of pelagic sharks caught by the Japanese longline fishery in the Atlantic Ocean. ICCAT SCRS 35: 393-398.

Nakano, H. and S. Clarke 2006. Filtering method for obtaining stock indices by shark species from species-combined logbook data in tuna longline fisheries. Fish. Sci. 72: 322-332.

Rui et al. 2015. Distribution patterns of sizes and sex-ratios of blue shark in the Indian Ocean. IOTC-2015-WPEB11-XX.

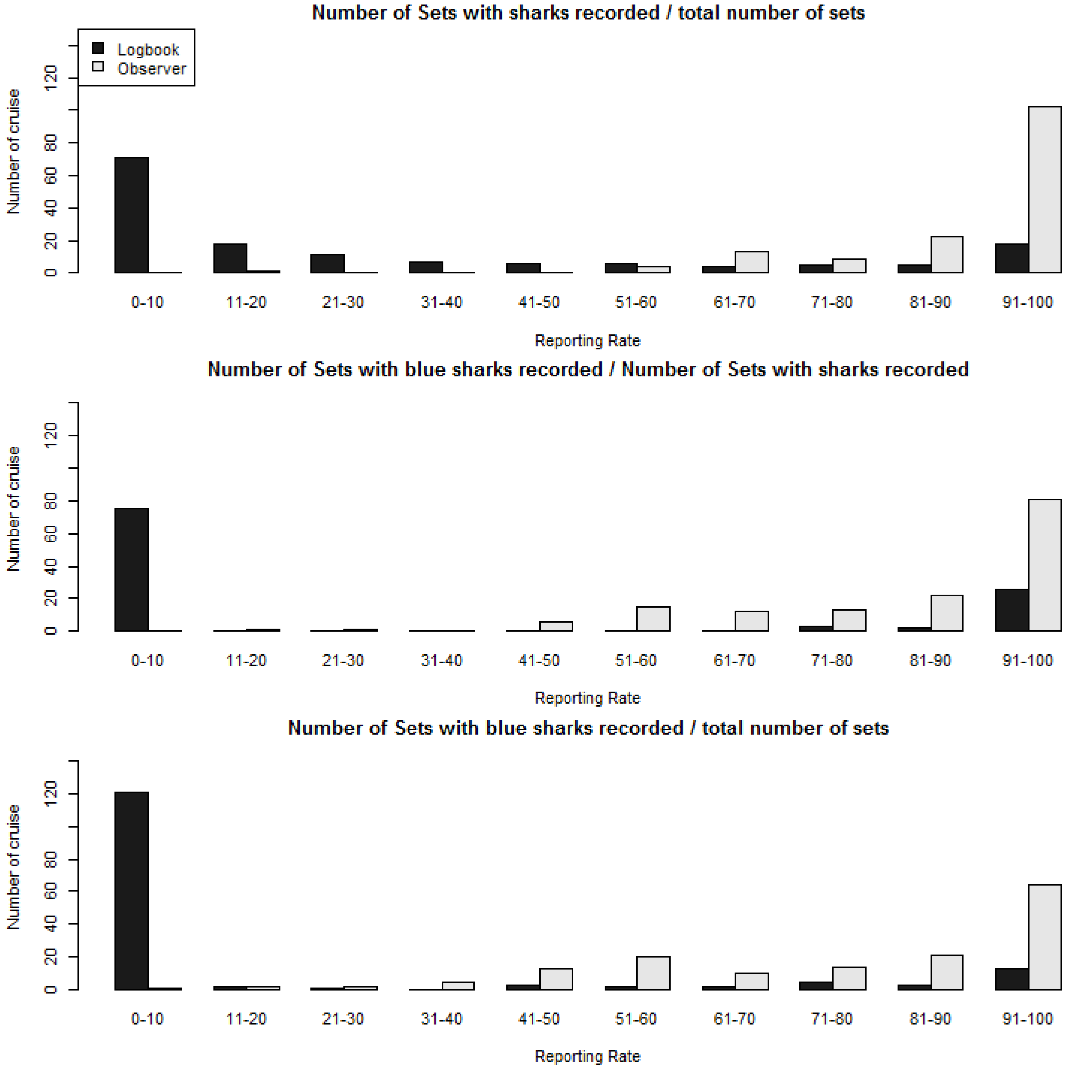

Fig.1 Number of cruise against reporting rate (RR) intervals of 10 % for logbook data (black bar) and observer data (gray bar) during 1994-2013.

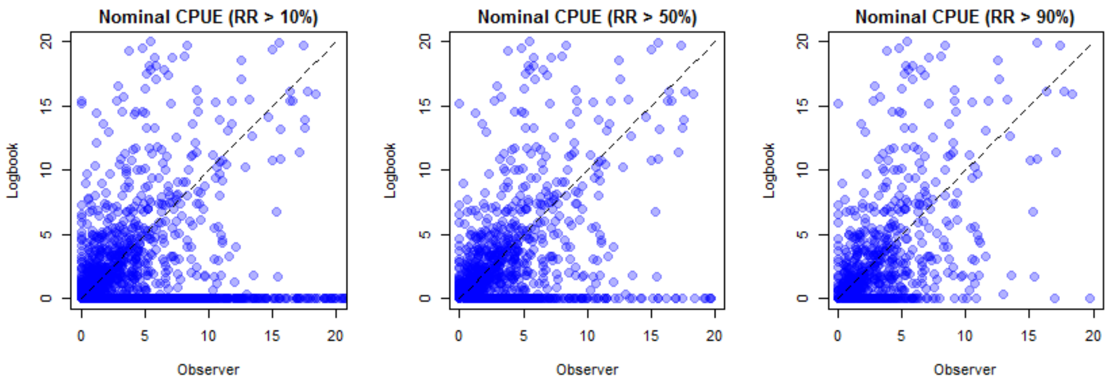

Fig.2 Correlation of nominal CPUE of blue shark between logbook data and observer data during 1994-2013 with filtering by different value of reporting rate (RR; number of sets with “sharks” recorded / total number of sets). Dotted line denotes the ratio of 1:1.

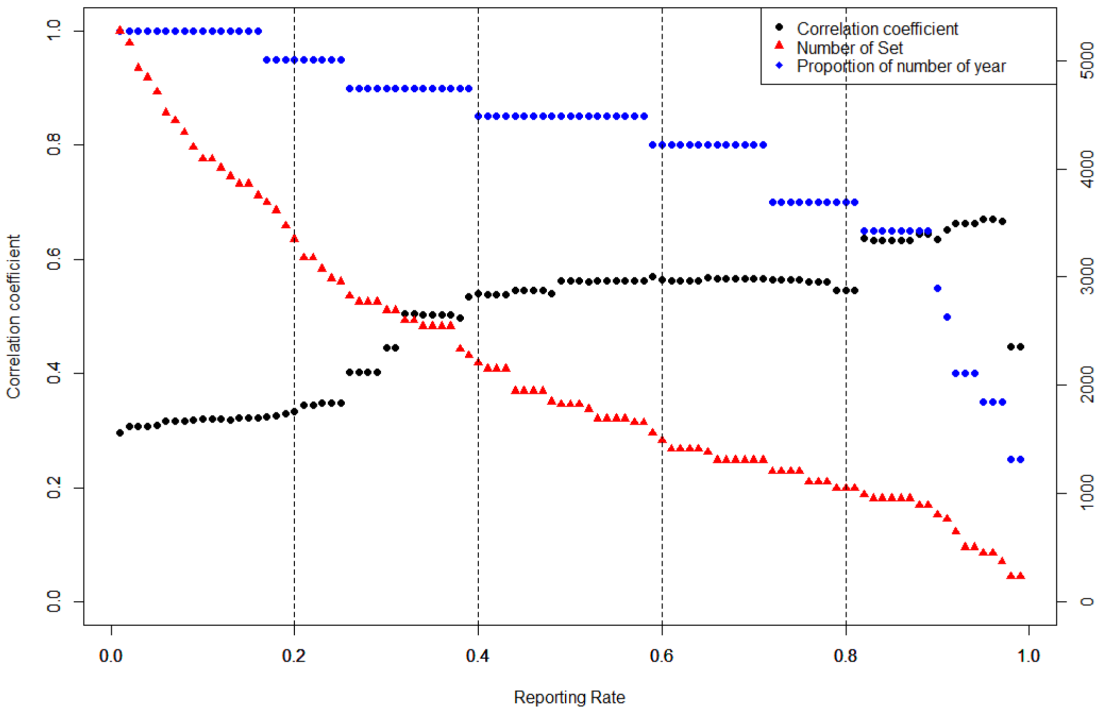

Fig.3 Correlation coefficient of nominal CPUE of blue shark between logbook data and observer data during 1994-2013 (black circle), number of sets of logbook data (red triangle), proportion of the number of year to the total number of years of logbook data (blue circle) for filtered data with reporting rate (RR; number of sets with “sharks” recorded / total number of sets) intervals of 1 %

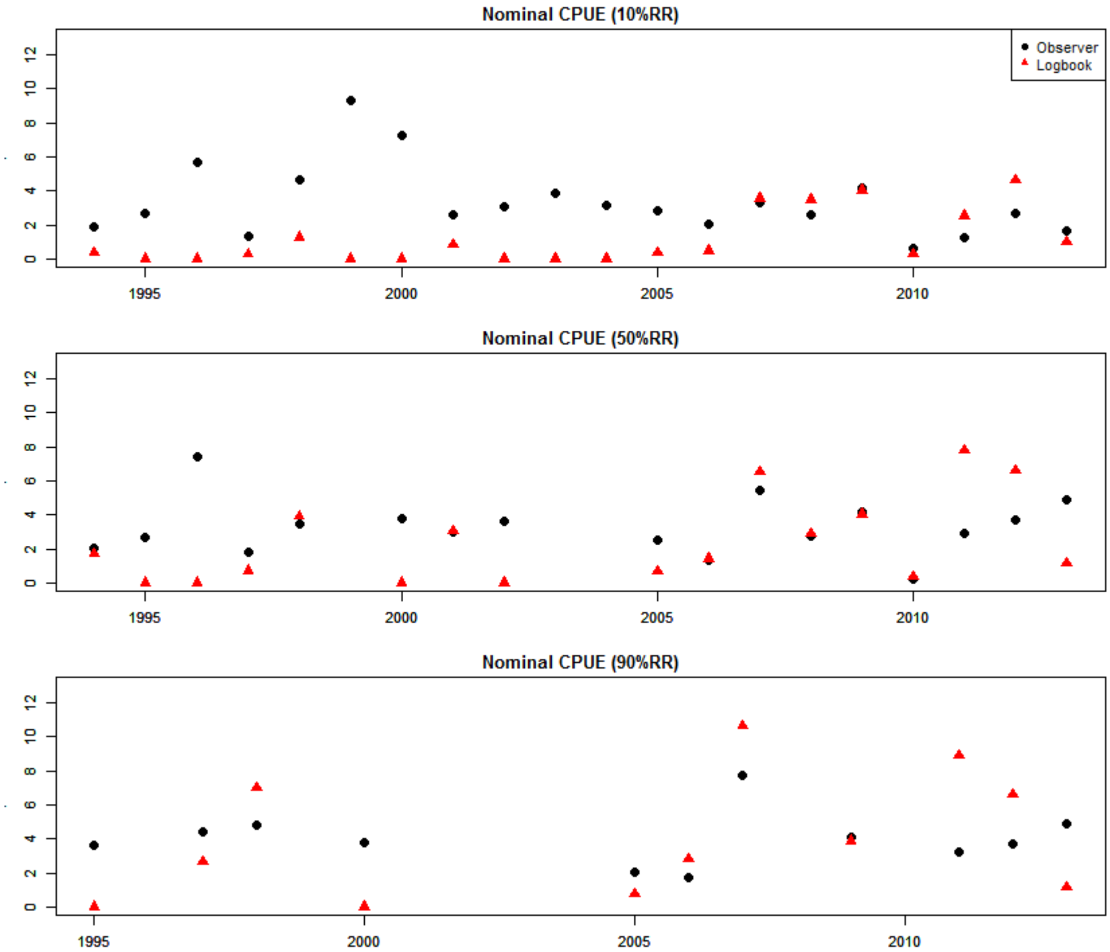

Fig.4 Comparison of annual changes of nominal CPUE of blue shark between logbook data (red triangle) and observer data (black circle) during 1994-2013 for filtered data with different reporting rate (RR; number of sets with “sharks” recorded / total number of sets).

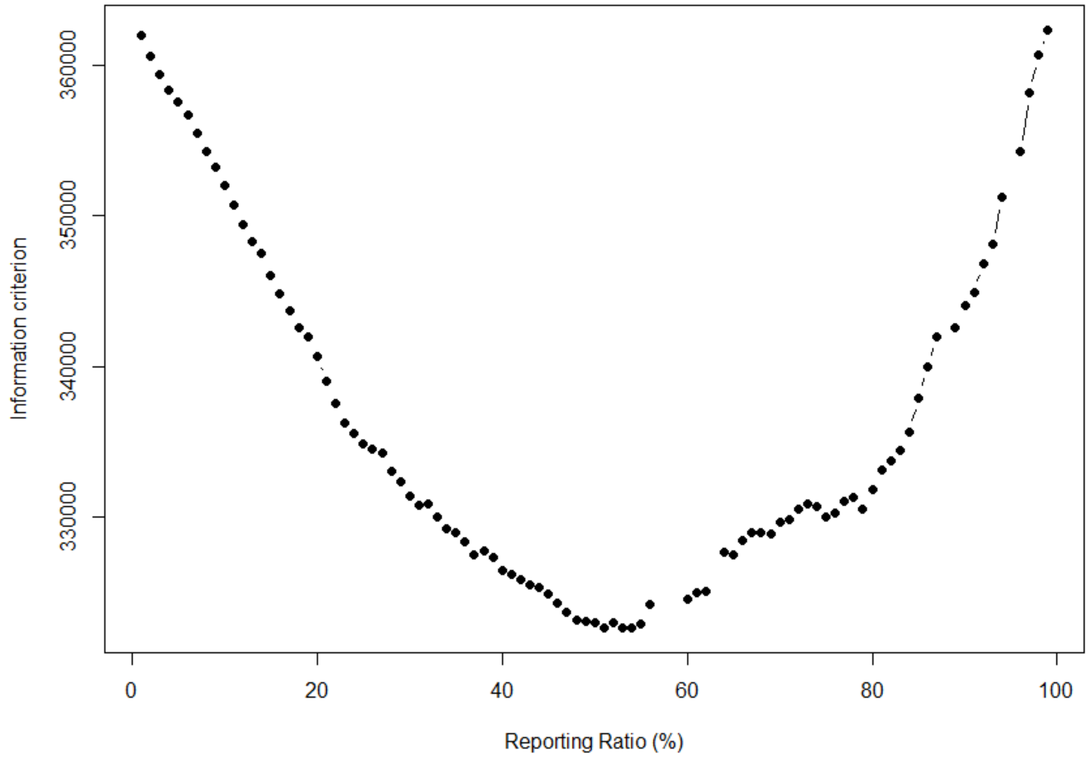

Fig. 5 AIC against different reporting rate (RR; number of sets with “sharks” recorded / total number of sets). AIC is calculated from CPUE standardization of blue shark for logbook data during 1994-2013.

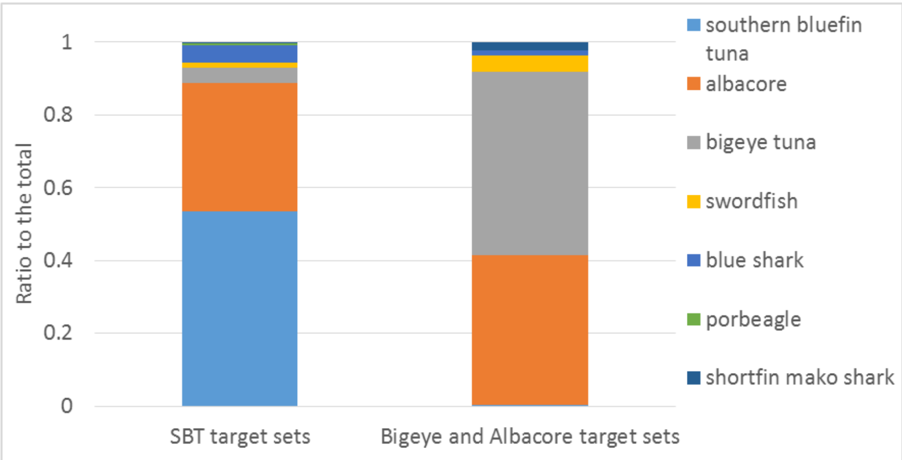

Fig. 6. Example of species composition of southern bluefin tuna targeting sets, and bigeye and albacore tunas targeting sets. Data in 1997 in the area south of 25S were used.

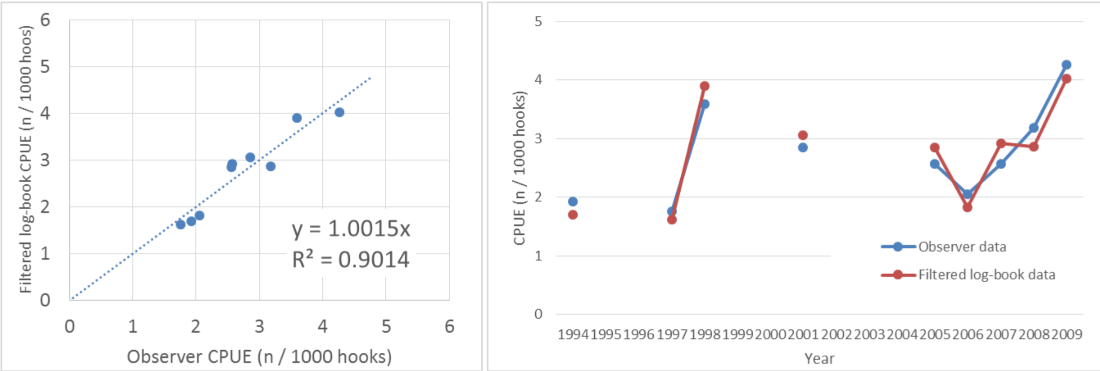

Fig. 7. Comparison of nominal CPUE of blue shark by 54% filtered log-book data of southern bluefin tuna targeting data and the one by CCSBT observer data. Data after 2009 were eliminated due to the fact of Japanese longliners mandated best utilization of blue shark meet and prohibited shark fining.

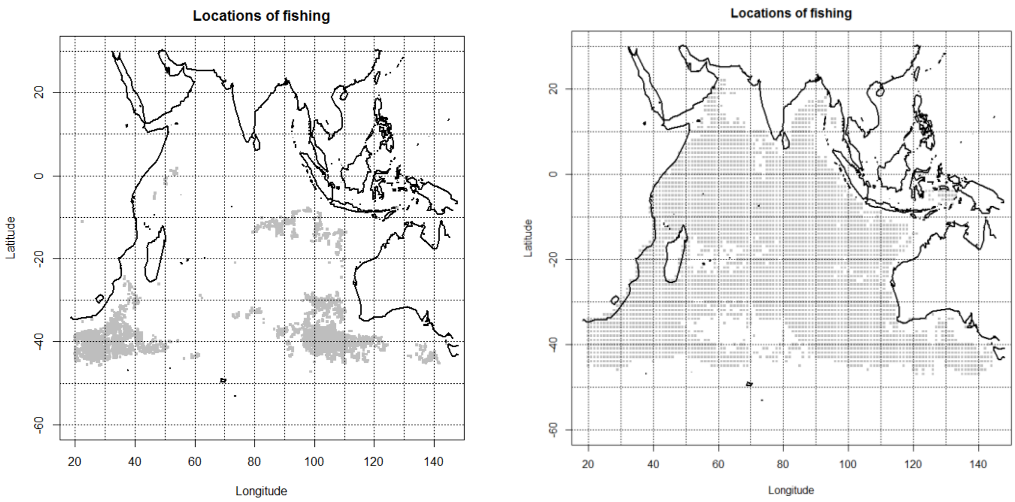

Fig. A1 Operational locations of Japanese longliner for observer data (left panel) and logbook data (right panel).