STANDARDIZATION OF HOOKING RATE (HR) FOR SWORDFISH (*XIPHIUS GLADIUS)* OCCURRING AROUND WESTERN INDIAN OCEAN (AREA 51) AND EASTERN INDIAN OCEAN (AREA57) BASED ON SURVEY DATA COLLECTED THROUGH FSI SURVEYS

D.K.GULATI, S.RAMACHANDRAN, H.D.PRADEEP, L.RAMALINGAM
Fishery Survey of India, Govt. of India, Kochangadi, P.B.No.853, Kochi, Kerala-682005, India
Fishery Survey of India, Govt. of India, Mormugao Zonal Base, Opp. Microwave Station, Mormugao, Goa-403803, India
Fishery Survey of India, Govt. of India, New Fishing Jetty, Sassoon Dock, Colaba, Mumbai, Maharashtra– 400005, India

Abstract

In the present study, swordfish (*Xiphius gladius*) catch in numbers and effort (hooks) data from 'Catch data sheet of FSI large longline fishing vessels operating in FAO area of 51 and 57 from 2007-2019' were analyzed together. These survey data comprise daily geo-referenced fishing positions (latitude and longitude), year, quarters, soaking time (time duration), species number and fishing effort (number of hooks) and catch rate of other coexisting species. Nominal hooking rate was calculated as the number of individuals captured per 100 hooks.

Due to the large percentage of zero swordfish catch in the survey data, the hooking rate (HR) of sword, as the number of fish caught per 100 hooks, was standardized using GLM in R approach with a delta lognormal approximation. The presence/absence and abundance (CPUE) of swordfish were modeled separately. The variables used in the model take into account spatial and temporal variations as well as the abundance of coexisting species. In total, 3056 fishing operations were carried out between 2007 and 2019 of which 1274 and 1782 operations in the FAO area of 57 and 51 respectively.

Abundance indices for swordfish (*Xiphius gladius*) for the period 2007-2019 were estimated using data obtained through oceanic longline surveys conducted by the of FSI large longline fishing vessels. Indices were calculated both for Western Indian Ocean and Eastern Indian Ocean together. Individual longline set catch per unit effort data, collected by scientists, were analyzed to assess effects of factors such as year, quarters, soaking time, latitude, longitude and abundance of coexisting species such as yellow fin tuna, sailfish, marlins and skipjacks besides remotely sensed chlorophyll –a and sea surface temperature derived from satellite imagery.

The main effects considered were temporal (year, quarters), spatial (longitude, latitude) and coexisting species i.e. hooking rate of skipjack tuna, marlins and skipjacks, besides remotely sensed variables chlorophyll a and sea surface temperature in the model. The results suggested that spatial (longitude & latitude), season (year and quarters) and soaking time significantly influenced the nominal hooking rate of swordfish whereas, coexisting species factors and remotely sensed variables particularly surface temperature and chlorophyll turned out to be insignificant and eventually dropped from the model. The high degree of temporal variability

that is still shown in the standardized CPUE trends to suggest that the variables used in the GLM in R do not sufficiently account for all of the confounding factors, or the abundance may indeed be truly variable.

The principal goal of the study was to select the best model that could be used for subsequent prediction of swordfish abundance. The nominal hooking rate of swordfish showed a strong inter-annual fluctuation. However, this variability was reduced in the standardized hooking rate series. This indicated that the standardization process removed certain variability attributes to the explanatory variables. In this study, despite the environmental effects not being included in the model for standardization, our study provides useful information for the swordfish because the analysis based on long time series data which cover a considerable range of western Indian Ocean and Eastern Indian Ocean.

Introduction

Swordfish are a cosmopolitan species, found throughout the world's oceans and seas. Based on data from 'Commercial Long Liners catches, the species' latitudinal range extends from 50°N to 45°S in the western Pacific, from 50°N to 35°S in the eastern Pacific, from 25°N to 45°S in the Indian Ocean, from 50°N to 40°-45°S in the western Atlantic and from 60°N to 45°-50°S in the eastern Atlantic (Nakamura 1985).

Since FAO and international environmental groups have concerned on the conservation of swordfish in recent years, it is necessary to examine the recent trend of swordfish by examining the survey data of sword fisheries. Catch and effort data are being increasingly used to construct indices of relative abundance for commercial and recreational fisheries (Hoey et al. 1996; Brown 2002; Goñi et al. 1999). However, nominal catch rates obtained from fishery statistics or observer programs also require standardization to correct for the effect of factors not related to regional fish abundance but assumed to affect fish availability and vulnerability (Bigelow et al. 1999). The use of generalized linear models (GLM in R) is becoming standard practice in catch rate standardization because this approach allows identification of the factors that influence catch rates and calculation of standardized abundance indices through the year effect (Goñi et al. 1999). Long liners harvest older and larger fish than purse seiners (Cole 1980). For longlining, the effort measure used here is the number of hooks set (in hundreds). Catch is reported as the number of fish caught. The delta-lognormal modeling, which can account for a large proportion of zero values, is an appropriate approach to model zero-heavy data (Vignaux 1994; Starr, 2012).

Key words:

Generalized Linear Modeling **(**GLM in R), ***Xiphius gladius***, billfish, Hooking rate, DLN(Delta lognormal non-linear model), sst (Sea surface temperature)

Material and methods

Fishery Survey of India longline fishery data from the Western (1274sampling stations) and Eastern (1782) Indian Ocean for the period from 2007 to 2019 were used for standardization hooking rate (number of fish per 100 hooks) of swordfish (figure1).

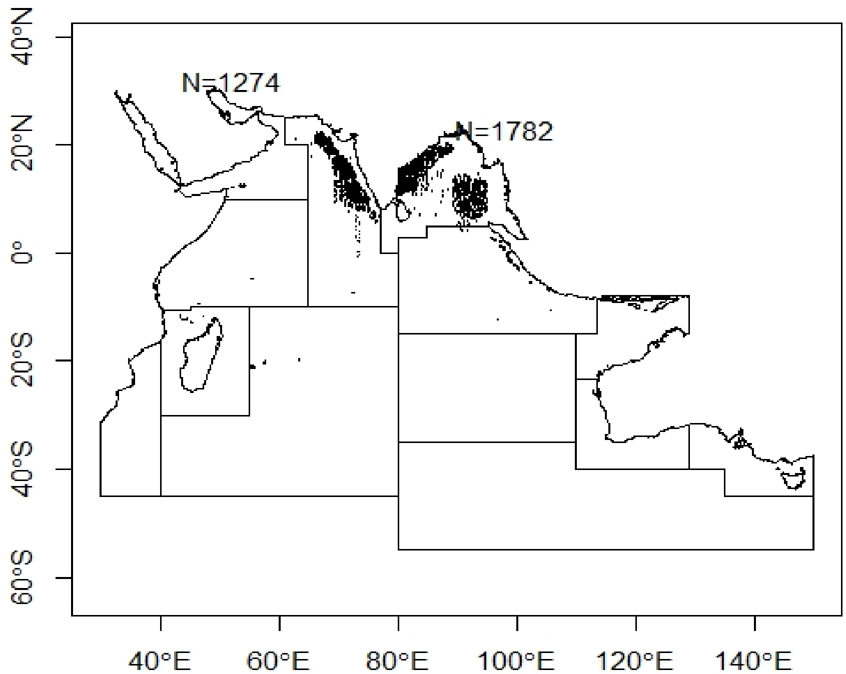

Figure 1 Sampling distribution in FAO Area 51 and 57

These data comprise daily geo-referenced fishing positions (latitude and longitude), species number, fishing effort (number of hooks), environmental parameters such as sea surface temperature and Chlorophyll-a derived from satellite imagery (https://oceancolor.gsfc.nasa.gov). A large proportion of zero values are found in the catch data obtained from FSI surveys of highly migratory species swordfish conducted in the far distant waters (figure2).

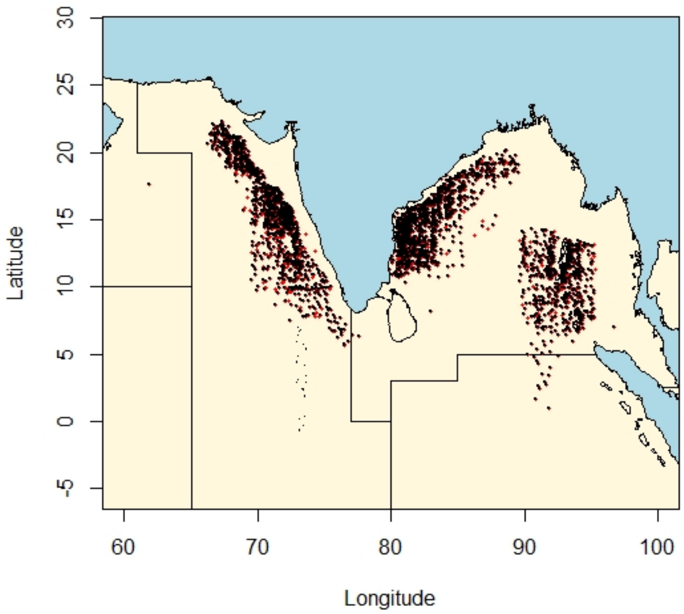

Figure 2 High Proportion of zeroes (black points) with positive effort

Many species, especially highly migratory have a high proportion of zero catches with positive effort. In order to deal with this type of highly skewed data, methods that deal with zero catch observations are required. The Poisson or the negative binomial distributions explicitly include zeros in the probability density distribution, however the Poisson function is usually restrictive, as the observed variance is normally greater than the mean (Ortiz and Arocha 2004). Other approaches include the use of the delta type two-step models, such as the delta lognormal (Pennington 1996; Lo et al. 1992; Stefansson 1996; Ortiz and Arocha 2004) or the zero-inflated models (Lambert 1992; Shono 2008). The delta model analyzes separately the positive observations and the probability that a null or positive observation occurs, and consists of two GLM, one assuming a lognormal and the other a binomial distribution. Given that the proportion of zero observations for the swordfish is moderate (20–60% annual average), we opted to use the delta-lognormal model as recommended by (Vignaux 1994; Starr, 2012).

Survey data for swordfish typically have statistical distributions with a large fraction of "zeros" (samples in which no swordfish are observed) and may be so distorted that conventional methods based on the sample mean yield inefficient estimates of abundance.

The delta distribution avoids problems with contagion by treating zero and nonzero data separately; final estimates of abundance are obtained from the product of the proportion and mean for nonzero observations. The delta distribution is used in many disciplines to model processes that generate more zero observations than might be expected on the basis of distributional assumptions (Lambert 1992).

Factors and interactions that explained 0.5% or more of the variability were considered significant and included in the final model. In the present study the environmental effects were not included in the model for standardization. The results obtained in this study can be improved if long time series observers' data on physical oceanic variables were included in the model. The best model for both GLM and PA models were selected using the Stepwise AIC method (Venables and Ripley, 2002).

The following multiplicative model was applied to the data in this study:

The catch rates of the positive catch events (hauls with positive catch) were modeled assuming a lognormal error distribution.

The best model for both GLM and Delta models were selected using stepwise AIC method.
bestmodel<-
glm(success(HR)~factor(year)+factor(q)+factor(latstrat)+factor(lonstrat)+factor(soaking time),family =binomial(link=logit),data=swordfish)

To calculate the proportion of positive records we used a model assuming a binomial error
best.glm<-
glm(log(hr.swordfish.)~factor(year)+factor(q)+factor(latstrat)+factor(lonstrat)+factor(soaking time), data=swordfish).

The log normal model is applicable for positive catch data. Zero catches are also encountered in observer samples. A GLM in R model based on a binomial distribution and using the presence/absence of swordfish (success = 1/0)) as the dependent variable was also fitted to the same set of data using the same set of explanatory variables. The binomial model will provide another series of standardized annual CPUE coefficients that is similar to the series estimated from the lognormal GLM in R. A combined model which integrates the two series of relative annual changes estimated by the lognormal and binomial models was estimated using the delta distribution which allowed zero and positive catches (Vignaux 1994; Starr, 2012).

$$Y_y^{Comb} = \frac{Y_y^{Ln}}{\left[1 - P_0\left[1 - \frac{1}{Y_y^{Binom}}\right]\right]}$$

where
Y_y^{Comb} = combined CPUE index for year y,
Y_y^{Ln} = lognormal CPUE index for year y,
Y_y^{Binom} = binomial CPUE index for year y, and
P_0 = proportion of zeros for base year 0.

For comparison with the standardized CPUE index, we also estimated the nominal CPUE (Arithmetic CPUE) and scaled it to the level of the standardized CPUE index.

$$A_y = \frac{\sum_{i=1}^{n_y} C_{iy}}{\sum_{i=1}^{n_y} E_{iy}}$$

$$\bar{A} = \sqrt[n_y]{\prod_{y=1}^{n_y} A_y}$$

$$A'_y = \frac{A_y}{\bar{A}}$$

where $C_{iy}$is catch and $E_{iy}$is effort for each record $i$in year y; $\bar{A}$ is the geometric mean of the Arithmetic CPUE; and $A_y$and A'_y are Arithmetic CPUE and scaled Arithmetic CPUE for year y, respectively. The indices of abundance by year were determined based upon the standardized year effects. For comparative purposes, each relative index of abundance was obtained dividing the standardized catch rates by the mean value in each series

Results and discussion

Abundance and distribution of marine organisms depends upon the oceanic conditions and hydrodynamics structures contributing toward its nutrient richness which in turn attract forage fishes and cephalopods and becomes the feeding ground for top predators. Probably the higher hooking rate for the swordfish shown in the fig.3 is attributable to the oceanic conditions and relatively higher concentrations of nutrients due to inflow of river water carrying nutrients.

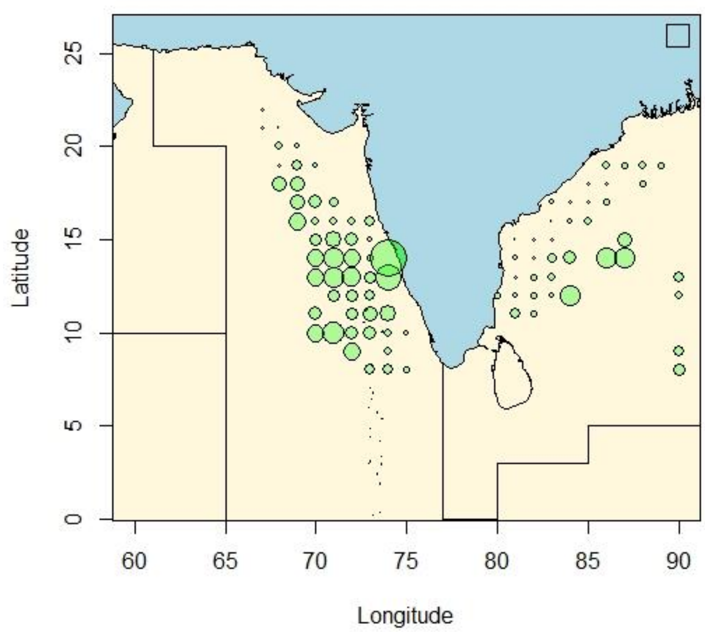

Figure 3 Map showing hooking rate for Swordfish caught in FSI surveys

The marlins, oceanic sharks, sail fish, sword fish etc. are not usually targeted by the Tuna long liner fleets of FSI operating in Indian Ocean but are caught and retained as a by-product as it fetches high value. Data collected during the period 2007-2019 from the operations of FSI Tuna long liners operating in the Indian ocean clearly shows swordfish are less abundant than the targeted species like yellowfin tuna.

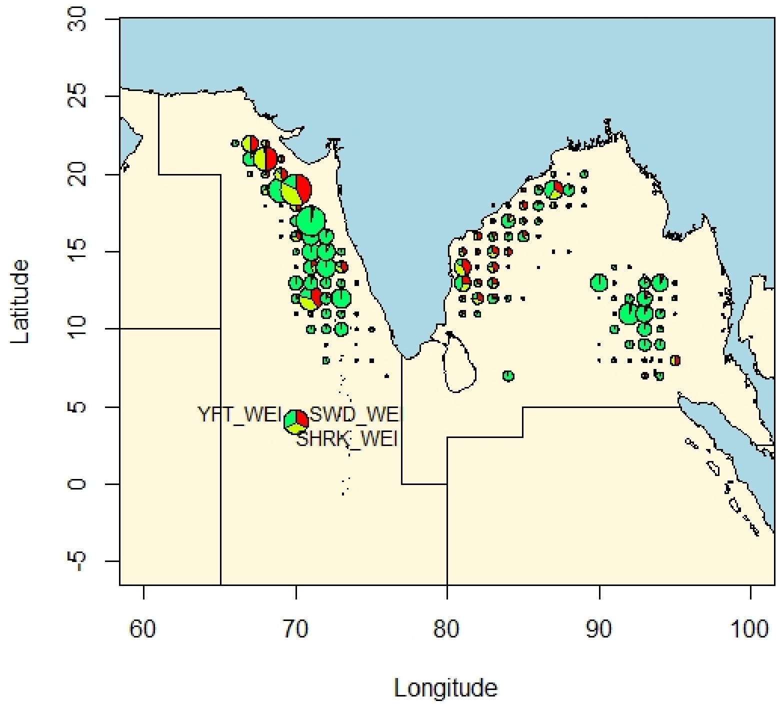

Figure 4 Map showing abundance of swordfish

The standardized hooking rate trend contains the combined effects from two models, one that calculates the probability of a zero observation and the other one that estimates the count per year. The nominal hooking rate of swordfish showed a strong inter-annual oscillation. This high variability was reduced in the standardized hooking rate series. In general, the standardized hooking rate series of the swordfish caught by FSI longliners showed a stable trend whereas, the temporal trends estimated by the models suggest a decline in catch rates in the year between 2011and 2014, thereafter remaining relatively stable at these lower levels in the series fig.5

The GLM approaches are more traditional and the delta-lognormal models explicitly model the zero catches. Progressive improvements in expertise and technological improvements in the gear will also affect fishing power, but are particularly difficult to quantify.

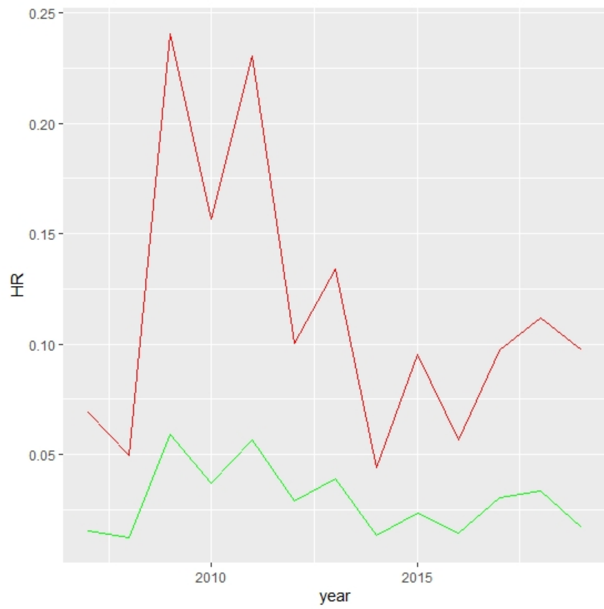

Figure 5 The standardized hooking rate series for swordfish using DLN model

The GLM in R approaches are more traditional and the delta-lognormal models explicitly model the zero catches. The variable with the highest contribution to the explained deviance was lonstrat (69%) followed by latstrat (22.2%), year (15.8%), quarter(15.8%), soaking time (3.4%),while the explained deviance of chlorophyll-a, sst, marlin, skipjack and sailfish was less than or slightly more one percentage case of Swordfish has been dropped from the GLM in R model. Most of the explained deviance came from the spatial and temporal effects (Table1).

```
Analysis of Deviance Table

Model: gaussian, link: identity

Response: log(HOOKRATE)

Terms added sequentially (first to last)


                     Df Deviance Resid. Df Resid. Dev      F
Pr(>F)
NULL                                   245     71.384
factor(YEAR)          9   2.5944       236     68.789 1.1566
0.3242815
factor(q)             3   2.5606       233     66.229 3.4247
0.0180147 *
factor(latStrat)      3   3.3520       230     62.877 4.4831
0.0044525 **
factor(soaking time)  2   0.5574       228     62.319 1.1183
0.3286655
factor(lonStrat)      6   6.9898       222     55.329 4.6742
0.0001687 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 1 Analysis of Deviance

The GLM in R identified that the four categorical longitude, latitude, year and quarter contributed more than 99%. The diagnostic results from the DLN model do not indicate severe departure from model assumptions (fig.).

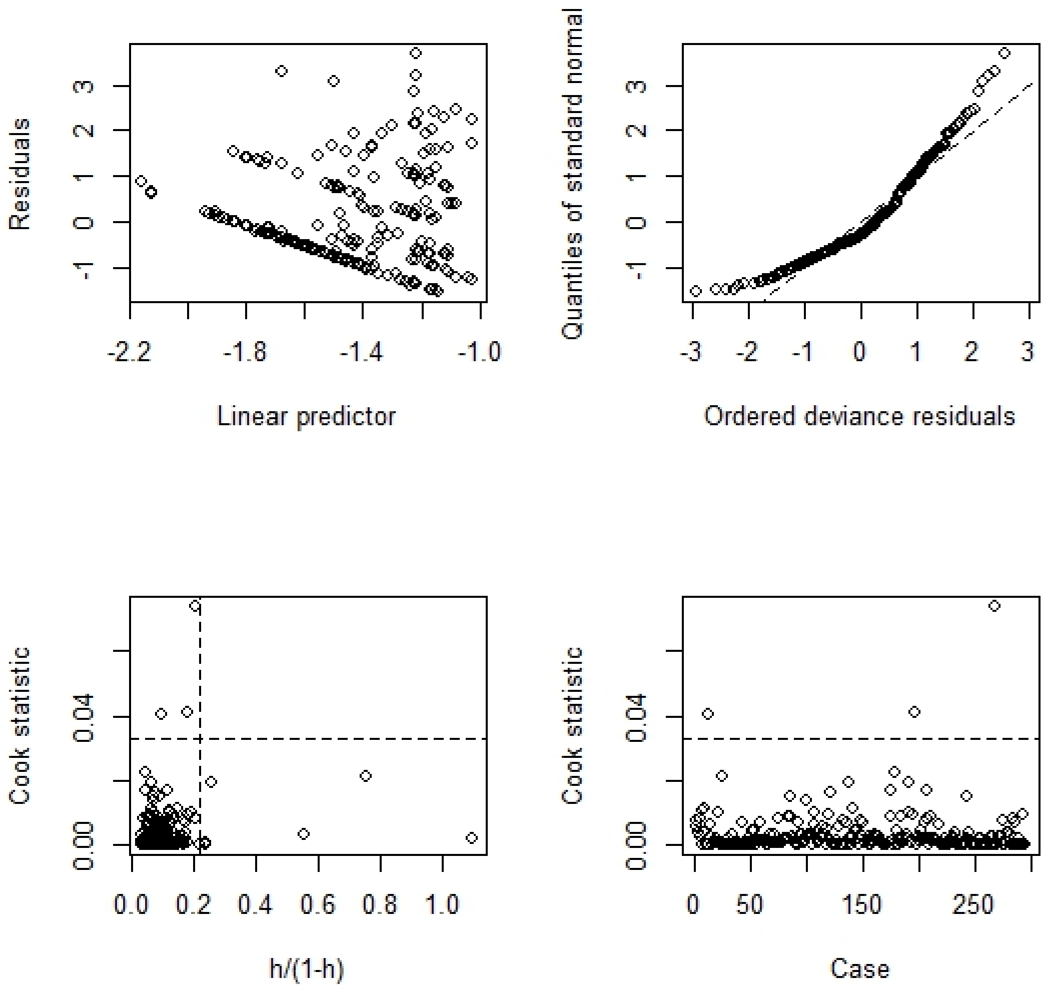

Figure 6 Diagnostic results from DLN

The analysis of deviance in case of Indian Ocean shows that the model could explain only22% of the deviance. The information on other explanatory variables such as environmental factors, some of the interactions, mechanization, more information on seasonality and hydrodynamic structures such fronts and eddies becomes the hot spots of biological activities through local concentration of nutrients which serve the foraging grounds for top predators as these structure generally attract forage fish and cephalopods, etc. is needed to reduce the unexplained deviance and to enhance the reliability of the model.

The aims of this paper were to provide a statistical modeling framework for conducting hooking rate standardizations using Fishery Survey of India long line fishery data collected from the Western and Eastern Indian Ocean for swordfish and provide a comparison in the trends between the nominal CPUEs and their standardized. For the binomial model, the error distribution plot, link function plot, and variance function plot all followed the expected patterns confirming the model assumptions (McCullagh and Nelder 1989). For the lognormal model, the link function plot and the error distribution plot also followed the expected patterns. Overall model diagnostics confirmed the model fit and estimate indices obtained. Since FSI data set is long time series, meaningful temporal trends are anticipated.

The standardized CPUE is similar to the nominal CPUE with no overall significant upward or downward trends. Further improvements are possible on all levels i.e., better reporting of species targets and fishing position, better classification of fishing time and vessel power and possibly the inclusion of water parameters.

The high degree of temporal variability that is still shown in the standardized CPUE trends suggest that the data are too sparse to give any meaningful indication of proxy abundance. Nevertheless, this may also suggest that variables used in the GLM in R s do not sufficiently account for all of the confounding factors, or abundance may indeed be truly variable. Standardization of the indices revealed several effects significantly associated with catch rate. Standardizing the data removed these effects from the indices of relative abundance, resulting in what may be considered more representative indices.

The present study fully agrees with IOTC recommendation that in order to avoid further decline in the catch rate of the swordfish, no further increases in catch and effort in this region. The declining in the catch rate of survey data is not directly attributable to depleted stocks and may be because of local overfishing.

Our present findings revealed that the catch rate did not increase with soak time, where minimum soak time, from end of setting to start of hauling, used in swordfish GLM in R models found to be in the agreement with the previous findings that total soak time is inappropriate for catch models because it includes haul back time, which increases as a function of catch. If landed catch does not increase as a function of soak time, then limiting longline soak time to reduce bycatch mortalities would not cause decreased swordfish catch nor result in economic losses for fishers. While minimum soak time limits would likely decrease bycatch mortality rates in swordfish longline fisheries, impacts on other aspects of the fishing process would need to be considered, such as negative impacts on fisher safety ((Carruthers,2011*et.al)*.

The recent advances in tracking technologies have facilitated mapping of where the oceanic activities like hydrodynamics, eddies, etc. are most likely, utilizing such technologies may help improve the reliability of the models.

References

BIGELOW, K.A., C.H. Boggs, and X. He. 1999. Environmental effects on swordfish and blue shark catch rates in the US North Pacific longline fishery. Fisheries Oceanography 8: 178-198.

Carruthers, Erin & Neilson, John & Smith, Sean. (2011). Overlooked bycatch mitigation opportunities in pelagic longline fisheries: Soak time and temperature effects on swordfish (*Xiphias gladius*) and blue shark (*Prionace glauca*) catch. Fisheries Research - FISH RES. 108. 112-120. 10.1016/j.fishres.2010.12.008

COLE. J. S.1980. Synopsis of biological data on the yellowfin tuna, *TkunmtSalbacares*(Bonnaterre, 1788), in the Pacific Ocean.In W. H. Bayliff (editor), Synopses of biological data on eightspecies of Scombrids, p. 71-150. IATTC Spec. Rep. 2. Inter-American Tropical Tuna Commission, La Jolla, CA, U.S.A.

FAO Fisheries Department, “World review of highly migratory species and straddling stocks,” ***FAO Fisheries Technical Papers*, Vol. 337, pp.1-70 (1994)**.

GOÑI, R., F. Alvarez and S. Adlerstein. 1999. Application of generalized linear modeling to catch rate analysi of Western Mediterranean fisheries: the Castellón trawl fleet as a case study. Fisheries Research 42: 291-302.

HOEY, J.J., J. Mejuto, J.M. Porter, H.H. Stone and Y. Uozomi. 1996. An updated biomass index of abundance for North Atlantic swordfish. Collect. Vol. Sci. Pap. ICCAT, 46(3): 354-361.

LAMBERTD, 1992. Zero-inflated Poisson regression. with an application to defects in manufactunng. Technometrics 34: 1-14.

McCullough, p., Nelder, J. A., 1989. Generalized linear models, 2nd ed. New York: Chapman and Hall. Nakamura, I. 1985. FAO species catalogue. Vol 5. Billfishes of the world. An annotated and illustrated catalogue of marlins, sailfishes, spearfishes and swordfishes known to date. FAO fisheries symposium. 5(125): 65pp.

ORTIZ, M. and G.A. Diaz. 2004. Standardized catch rates for yellowfin tuna (Thunnus albacares) from the U.S. pelagic longline fleet. Collect. Vol. Sci. Pap. ICCIAT, 56(2): 660-675.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
URL http://www.R-project.org/.

Venables, W. N. and Ripley, B. D. 2002. Modern Applied Statistics with S. Fourth edition. Springer.

Vignaux, M. 1994: Catch per unit effort (CPUE) analysis of west coast South Island and Cook Strait spawning hoki fisheries, 1987–93. N.Z. Fisheries Assessment Research Document 94/11. 29 p. (Unpublished report held in NIWA library, Wellington, New Zealand).

Acknowledgements

The authors wish to express their sincere thanks to the Fishery survey of India, Ministry of Fisheries (Dept. of Fisheries) for utilizing the data and according permission to present the paper in the IOTC Eighteenth Working Party on Billfish (WPB18) scheduled to be held through VC during 2-4 September,2020.

.