Historical standardized CPUEs of the blue shark (*Prionace glauca*) from 1966 through 1989

Emma Gee[*], Evgeny V. Romanov[†], David Curnick[‡], Barbara Block[§], Francesco Ferretti[¶]

Abstract

We used an historical longline survey from 1966 to 1989 in the Indian Ocean basin to calculate standardized CPUEs for the blue shark (*Prionace glauca*). CPUEs were generated using a zero-inflated negative binomial (ZINB) generalized additive model (GAM). These CPUEs represent an important basin-wide baseline for blue sharks abundance at the start of industrialization of Indian Ocean fisheries.

Introduction

Like many other shark species, the blue shark suffers from a lack of historical baseline data. Further, the Indian Ocean is one of the least studied ocean sectors for shark exploitation and bycatch (Molina & Cooke, 2012). These spatial and temporal phenomena compound into a significant management challenge for sharks in the Indian Ocean region.

[*]Department of Environmental Studies, University of California Santa Cruz; To whom correspondence should be addressed. Email: emmagee@alumni.stanford.edu

[†]CAP RUN – CITEB (Centre technique de recherche et de valorisation des milieux aquatiques)

[‡]Institute of Zoology, Zoological Society of London

[§]Hopkins Marine Station, Stanford University

[¶]Department of Fish and Wildlife Conservation, Virginia Polytechnic Institute and State University

To address this issue, we utilize data from a scientific longline survey conducted throughout the region from 1966 to 1989 to reconstruct trends of shark abundance (Fig. 1). The historical longline survey was conducted by USSR scientists using gear targeting tuna (Romanov, Sakagawa, Marsac, & Romanova, 2006), making the data comparable to that recorded by modern commercial longline fleets reporting to the IOTC. Shortfin mako stocks in the Indian Ocean were thought to be near-pristine in 1971, five years after the start of the USSR survey (Brunel et al., 2018). While Japan and Taiwan started exploiting these waters in a limited capacity shortly before the survey began, the USSR survey period covers twenty other nations joining the longline fishery (Fonteneau, 2017), along with significant improvements in longline gear and refrigeration technology (Ward & Hindmarsh, 2007), the introduction of industrial purse seining (Fonteneau, 2017), and the start of direct targeting of sharks due to increased global demand for shark fins (Camhi, Valenti, Fordham, Fowler, & Gibson, 2009; Fabinyi, 2012). Thus, the USSR survey reflects shark populations at pre-industrial fishing levels.

Methods

Datasets used

USSR survey data

We estimated historical catch rates from a scientific longline survey carried out by the former Soviet Union (USSR). Longline sets (n = 4,678) were cast throughout the Indian Ocean between 1961 and 1989 as part of the Soviet Indian Ocean Tuna Longline Research Programme (SIOTLLRP) (Romanov et al., 2006). Scientists aboard the ship identified the sharks to species or genus level and measured fork length. We discarded data collected prior to 1966 due to concerns with the reliability of species identification.

SIOTLLRP surveyors recorded for each longline set: date, latitude, longitude, start and end of longline setting and hauling, number of hooks set, basket length, buoyrope length, hookline length, number of baskets, number of hooks per basket, estimated depth of each hook in a basket, and hook number for each capture (Romanov et al., 2006). From these, we derived: soak time, haul time, mean number of hooks per basket, ocean depth, distance from coast, and Longhurst biogeographical province (Bart, 2012), using the "marmap" package to calculate depth and distance to the nearest coast (Pante & Simon-Bouhet, 2013). Month was input into the model as the sum of a sine and cosine transform to linearize the cyclical nature of the seasons (Ferretti, Osio, Jenkins, Rosenberg, & Lotze, 2013). This yielded 14 explanatory variables for shark abundance, which we tested for collinearity using the variance inflation factor (VIF) (Faraway, 2016).

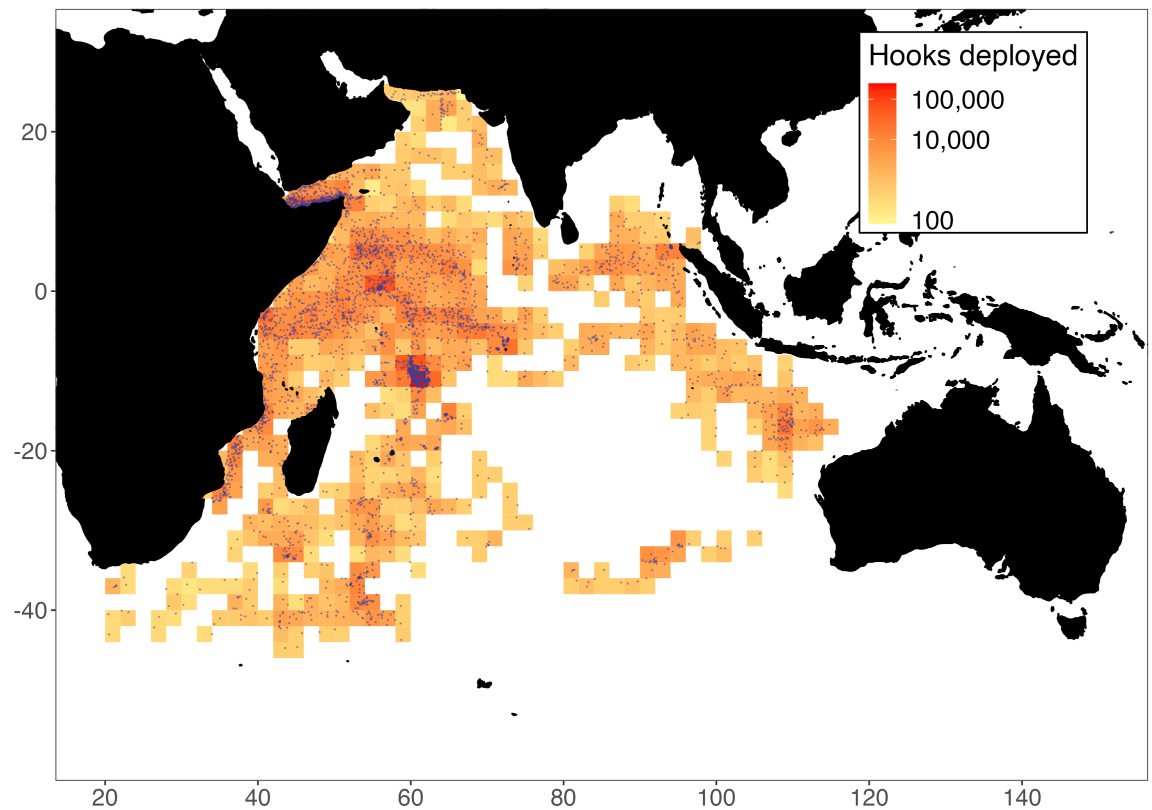

Fig. 1. USSR survey effort by number of hooks (2° by 2° resolution). Purple dots are locations of longline sets.

IOTC data

We obtained fork length data from the IOTC's publicly available catch and effort data (Commission, 2021).

CPUE standardization

We used a frequentist statistical modeling approach to produce standardized CPUEs (individuals per 1000 hooks deployed) from the USSR data. Our modeling approach consisted of three stages: statistical distribution and model framework selection, variable selection, and simulation. We used the blue shark as a model species owing to its status as the most abundant species in the dataset (n = 1,156) and one of the most commonly caught shark species in the Indian Ocean (Tsai & Liu, 2018). We used the blue shark catch data to choose a statistical distribution and model framework to use for all species, but variable selection and simulation were performed for each species for which we produced standardized CPUEs.

We considered 14 statistical distributions and modeling frameworks commonly used in the literature for CPUE standardization (Table 1). We selected the zero-inflated negative binomial (ZINB) generalized additive model (GAM) based on its low Akaike information criterion (AIC) value relative to other models (Table 1) (Akaike, 1998) and the ability of GAMs to model non-linear trends in the data.

Table 1. AIC values of candidate models for catch rate standardization. Selected statistical distribution and model framework is in bold.

Model	AIC	R function
Poisson GLM	9066.62	glm()
Negative binomial GLM	7124.69	glm.nb()
Zero-inflated Poisson GLM	8488.14	zeroinfl()
Zero-inflated negative binomial GLM	7013.7	zeroinfl()
Poisson GAM	9066.62	gam()
Negative binomial GAM	7120.81	gam()
Zero-inflated Poisson GAM	8171.72	zipgam()
Zero-inflated negative binomial GAM	**7034.77**	**zinbgam()**
Zero-inflated Poisson GLMM	Did not converge	glmmTMB()
Zero-inflated negative binomial GLMM	Did not converge	glmmTMB()
Tweedie GLM	Did not converge	glm()
Tweedie GAM	7837.2	gam()
Tweedie GLMM	Did not converge	glmmTMB()
Delta-lognormal	7424.85	deltaLN()

The ZINB GAM is a mixture model with two component models: a negative binomial GAM predicting counts and a binomial GAM predicting the probability of a false zero. We used the "zigam" package in R to fit ZINB GAM models (Wotherspoon & Burch, 2017). We modified the package's source code to produce confidence intervals for predicted values using a Monte Carlo approach (Preacher & Selig, 2012).

For variable selection, we followed Babyak's (Babyak, 2004) rule of having at least 10 non-zero counts in the data for each variable. To select variables under this limit, we conducted variable selection in two steps, first permuting the variables to find which produced the best models, and then determining whether any of those variables could be dropped from a preliminary model. For the first step, we tried all possible combinations of the 14 candidate variables in a process known as dredging (Barton, 2020). We tested the component models of the ZINB GAM separately and used GLMs because of the computationally expensive nature of dredging. Variables appearing in every model in the 95th percentile confidence set of model performance were then considered for their respective component of the ZINB GAM model. In our second step, to reduce the risk of overparameterization, we tested the negative binomial and binomial GAMs to see if any variables could be removed without significant ($> 1\%$) loss of % deviance explained.

In our final step of model development, we performed simulations to test the statistical power of our ZINB GAM and its ability to capture the underlying biological processes in the data. We generated simulated counts for the survey data using the ZINB distribution from the model 100 times. A new model was fit to each simulated dataset and the coefficients of the variables recorded. We plotted a histogram of each coefficient for each variable and examined the distribution for approximate normality. If the coefficients were not centered on the estimate generated from the real data, we concluded that the model did not successfully capture the process of the data. In cases where this was true, we added or removed variables until the coefficient distributions were centered and approximately normal. The final parameterization of the blue shark ZINB GAM was:

Count process: $count \sim year + lat + lon + depth + Longhurstprovince + month + soaktime + basketlength + offset(log(nhooks))$

Zero process: $w \sim year + lat + depth + month + Longhurstprovince + soaktime + offset(log(nhooks))$

Results

Standardized CPUEs

A total of 3,181 blue sharks were caught throughout the USSR survey. Figure 2 and Table 2 show the standardized CPUEs for blue sharks from the USSR survey. Residual plots can be found in the Appendix.

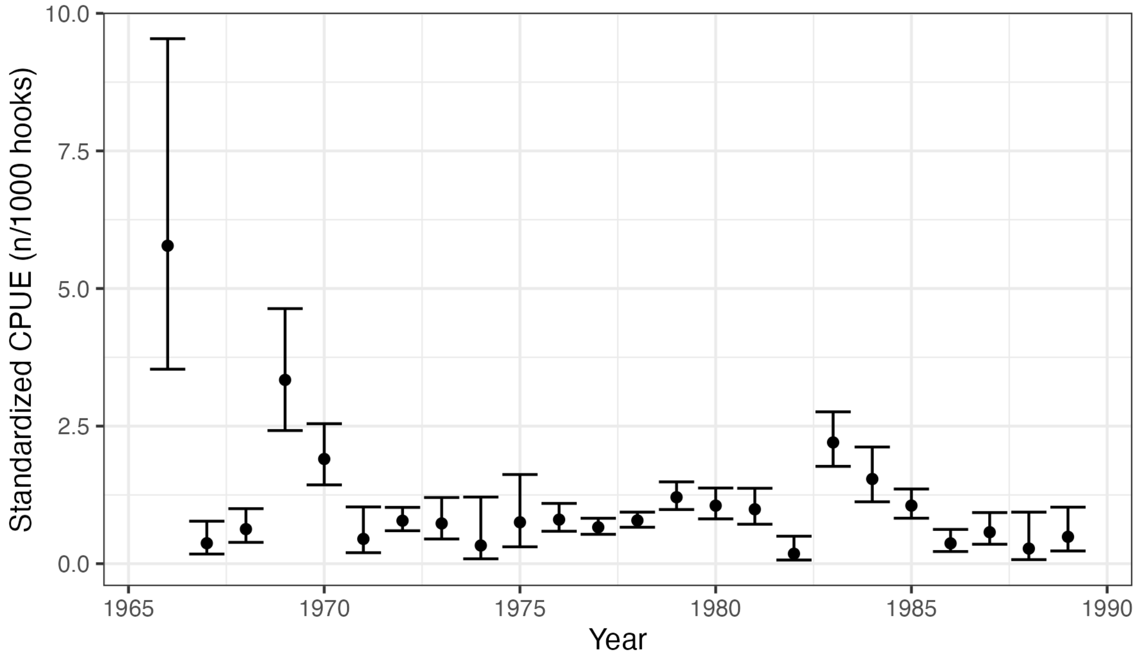

Fig. 2. Standardized CPUEs of the blue shark with 95% confidence intervals.

Table 2. Standardized CPUEs of the blue shark with 95% confidence interval bounds.

year	Standardized CPUE	Lower bound	Upper bound
1966	5.7769	3.5340	9.5390
1967	0.3717	0.1758	0.7725
1968	0.6281	0.3881	1.0000
1969	3.3390	2.4198	4.6344
1970	1.9024	1.4322	2.5438
1971	0.4500	0.1996	1.0313
1972	0.7809	0.5991	1.0233
1973	0.7313	0.4492	1.2027
1974	0.3306	0.0889	1.2111
1975	0.7529	0.3072	1.6202
1976	0.8012	0.5898	1.0959
1977	0.6619	0.5340	0.8255
1978	0.7860	0.6623	0.9373
1979	1.2075	0.9840	1.4867
1980	1.0552	0.8143	1.3751
1981	0.9887	0.7176	1.3705
1982	0.1812	0.0671	0.4994
1983	2.2034	1.7687	2.7595
1984	1.5388	1.1248	2.1204
1985	1.0561	0.8264	1.3572
1986	0.3695	0.2216	0.6228
1987	0.5719	0.3538	0.9291
1988	0.2751	0.0737	0.9369
1989	0.4882	0.2315	1.0267

Change in size

Mean fork length of blue sharks declined from 193.8 cm in the USSR data to 166.4 cm in the IOTC data, which covers 2006 through 2018 (Fig. 3).

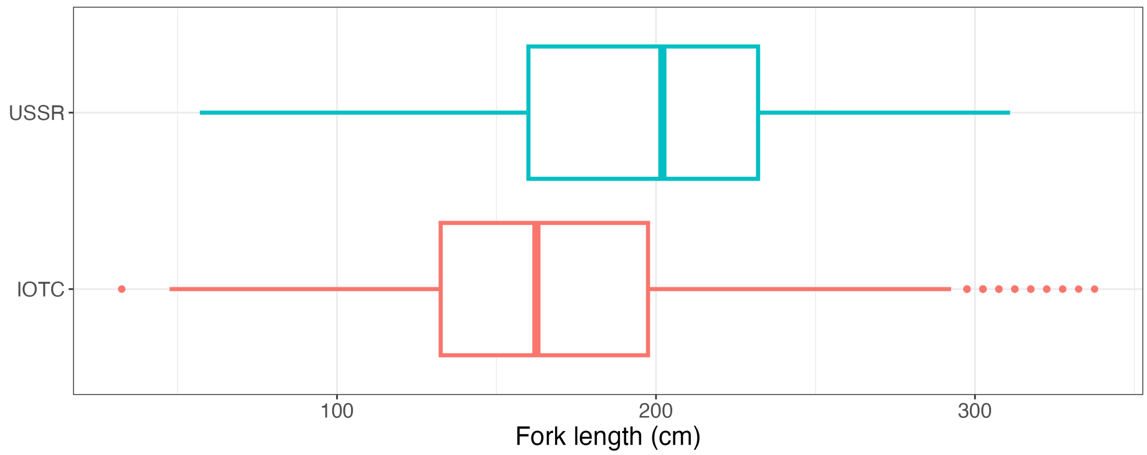

Fig. 3. Fork length of blue sharks in the USSR and IOTC data.

Discussion

A USSR survey that spanned 24 years provided a rare record of shark initial abundances in the Indian Ocean. Over the survey period, the abundance of blue sharks did not significantly change. However, fork length has declined from the USSR survey to present IOTC data, suggesting that fishing has had a long-term impact on the population.

References

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike* (pp. 199–213). Springer.

Babyak, M. A. (2004). What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic medicine*, *66*(3), 411–421.

Bart, V. (2012). *Longhurst province.* `https://www.marineregions.org/gazetteer.php?p=details&id=22538`.

Barton, K. (2020). Mumin: Multi-model inference. *R package version 1.43.17*. Retrieved from `\url{https://CRAN.R-project.org/package=MuMIn}`

Brunel, T., Coelho, R., Merino, G., Ortiz de Urbina, J., Rosa, D., Santos, C., & Macias, D. (2018). *A preliminary stock assessment for the shortfin mako shark in the indian ocean using data-limited approaches* (Tech. Rep. No. IOTC-WPEB14-2018-037). Indian Ocean Tuna Commission.

Camhi, M. D., Valenti, S., Fordham, S., Fowler, S., & Gibson, C. (2009). The conservation status of pelagic sharks and rays: report of the iucn shark specialist group pelagic shark red list workshop. *IUCN Species Survival Commission Shark Specialist Group. Newbury, UK. x+ 78p.*

Commission, I. O. T. (2021). *Catch and effort data - vessels using drifting longlines (iotc-2021-wpeb17(as)-data04_rev1).* `https://iotc.org/WPEB/17AS/Data/04-CELL`.

Fabinyi, M. (2012). Historical, cultural and social perspectives on luxury seafood consumption in china. *Environmental Conservation*, *39*(1), 83–92.

Faraway, J. (2016). faraway: Functions and datasets for books by julian faraway. *R package version*, *1*(7).

Ferretti, F., Osio, G. C., Jenkins, C. J., Rosenberg, A. A., & Lotze, H. K. (2013). Long-term change in a meso-predator community in response to prolonged and heterogeneous human impact. *Scientific reports*, *3*(1), 1–11.

Fonteneau, A. (2017). *Atlas des pêcheries thonières de l'océan indien/atlas of indian ocean tuna fisheries*. IRD éditions.

Molina, J. M., & Cooke, S. J. (2012). Trends in shark bycatch research: current status and research needs. *Reviews in Fish Biology and Fisheries*, *22*(3), 719–737.

Pante, E., & Simon-Bouhet, B. (2013). marmap: A package for importing, plotting and analyzing bathymetric and topographic data in r. *PLoS ONE*, *8*(9), e73051. (doi:10.1371/journal.pone.0073051)

Preacher, K. J., & Selig, J. P. (2012). Advantages of monte carlo confidence intervals for indirect effects. *Communication Methods and Measures*, *6*(2), 77–98.

Romanov, E. R., Sakagawa, G., Marsac, F., & Romanova, N. (2006). *Historical database on soviet tuna longline tuna research in the indian and atlantic oceans (first results of yugniro-nmfs data rescue project)* (Tech. Rep. No. IOTC-2006-WPTT-10). Indian Ocean Tuna Commission.

Tsai, W.-P., & Liu, K.-M. (2018). *Updated standardized cpue of blue shark by taiwanese large-scale tuna longline fishery in the indian ocean* (Tech. Rep. No. IOTC-2018-WPEB14-34). Indian Ocean Tuna Commission.

Ward, P., & Hindmarsh, S. (2007). An overview of historical changes in the fishing gear and practices of pelagic longliners, with particular reference to japan's pacific fleet. *Reviews in Fish Biology and Fisheries*, *17*(4), 501–516.

Wotherspoon, S., & Burch, P. (2017). zigam: Em implementation of zero-inflated gams. *R package version*, *0.1*.

Acknowledgments

We thank the work of YugNIRO scientists in collecting the historical survey data on sharks in the Indian Ocean funded by the Ministry of Fisheries of the former USSR. Additionally, we thank the work of the IOTC in collecting and compiling fisheries data for the region.

Appendix

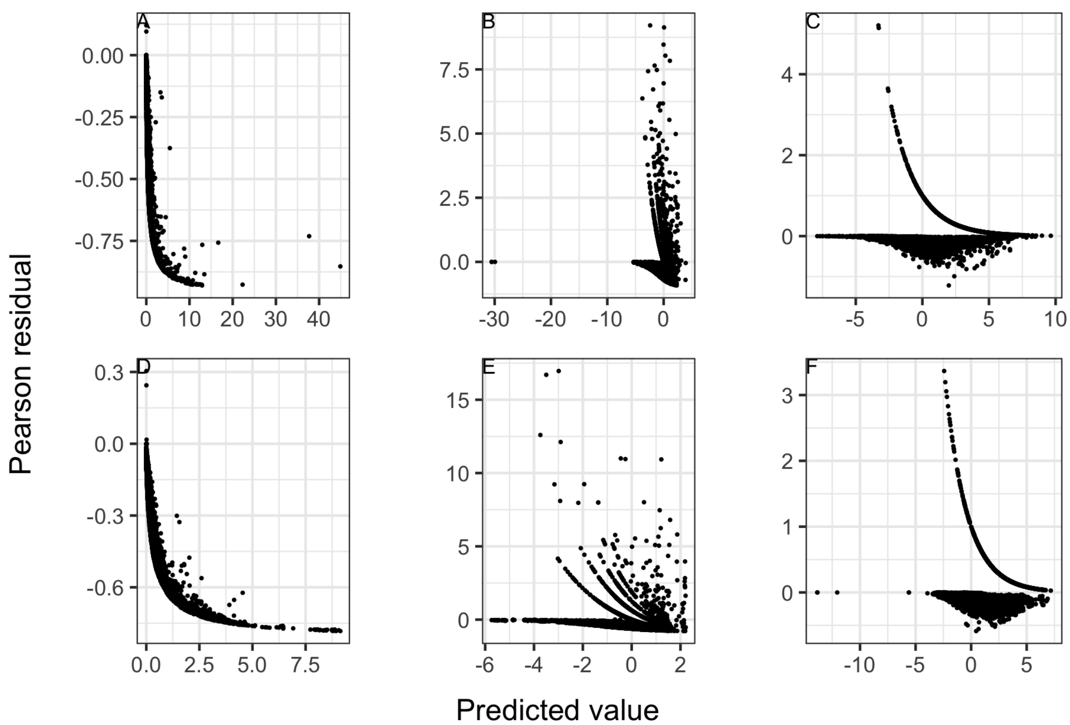

Fig. S1. Residuals of (A) the whole model with year as a factor, (B) the count process with year as a factor, (C) the zero process with year as a factor, (D) the whole model with year as a continuous variable, (E) the count process with continuous year, and (F) the zero process with continuous year for the blue shark.