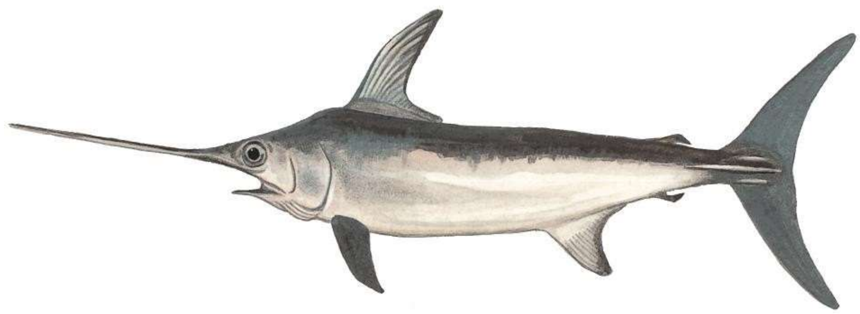
CPUE Standardisation of Swordfish caught in South African waters

da Silva, C[1,2], West, W[1]., and Kerwath, SE[1,3]



SUMMARY

Swordfish, *Xiphias gladius*, is a target species in the South African pelagic longline fleet operating along the west and east coast of South Africa. A standardization of the CPUE of the South African swordfish directed longline fleet for the time series 2004-2024 was carried out with a Generalized Additive Mixed Model (GAMM) with a Tweedie distributed error. Explanatory variables of the final model included Year, Month, geographic position (Lat, Long) and a targeting factor (Fishing Tactic) with two levels, derived by clustering of PCA scores of the root-root transformed, normalized catch composition. Vessel was included as a random effect. Swordfish CPUE had a definitive seasonal trend, with catch rates higher in winter (July - October) than in the rest of the year. Standardized (normalized) CPUE has been relatively stable over the 20-year period, ranging from a low 0.60 in 2014 to a peak of 1.36 in 2023. The CPUE has shown a general increasing trend since 2021.

KEYWORDS

Swordfish, standardized cpue, longline, GAMM, targeting, PCA cluster, random effect

AFFILIATIONS

[1]Department of Forestry, Fisheries and the Environment, [2]Rhodes University, South Africa, [3]University of Cape Town, South Africa

INTRODUCTION

Commercial fishing for large pelagic species in South Africa dates to the 1960s (Welsh, 1968; Nepgen, 1970). Exploitation of large pelagic species in South Africa can be divided into four sectors, 1) pelagic longline, 2) tuna pole-line 3) commercial linefishing (rod and reel) and 4) recreational line- fishing. Pelagic longline vessels are the only vessels that target swordfish, with negligible bycatch being caught in other fisheries. Pelagic longline fishing by South African vessels began in the 1960s with the main target being southern bluefin tuna (*Thunnus maccoyii*) and albacore (*Thunnus alalunga*) (Welsh, 1968; Nepgen, 1970). This South African fishery ceased to exist after the mid 1960's, as a result of a poor market for low quality southern bluefin and albacore (Welsh, 1968). However, foreign vessels, mainly from Japan and Chinese-Taipei, continued to fish in South African waters from the 1970s until 2002 under a series of bilateral agreements. Interest in pelagic longline fishing re-emerged in 1995 when a joint venture with a Japanese vessel confirmed that tuna and swordfish could be profitably exploited within South Africa's waters. Thirty experimental longline permits were subsequently issued in 1997 to target tuna, though substantial catches of swordfish were made during that period (Penney and Griffiths, 1999). The commercial fishery was formalised in 2005 with the issuing of 10-year long term rights to sword- fish- and tuna-directed vessels. The fishery is coastal and swordfish-oriented effort concentrates in the southwest Indian Ocean region (20°- 30°S, 30°- 40°E) and along the South African continental shelf in the southeast Atlantic (30°- 35°S, 15°- 18°E). As such, the fishery straddles two ocean basins, the Indian and Atlantic Ocean (Fig. 1).

The jurisdictions of the Indian Ocean Tuna Commission (IOTC) and International Commission for the Conservation of Atlantic Tuna (ICCAT) are separated by a management boundary at 20°E. The South African caught swordfish originate from an Indian and an Atlantic Ocean stock, with a broad admixture zone between 17°E and 30°E, hence the artificial split at 20°E in reporting stock indices requires further investigation.

South Africa's overall swordfish catches reached a peak in 2002 at 1 559 t, and likewise within the IOTC area of competence (Longitude > 20 degrees) it peaked at 828 t in that year. In recent years South Africa's swordfish catches have declined to around 300 t in the IOTC area of competence.

Here, we present an update of the standardised catch-per-unit-effort (CPUE) time series for swordfish caught in the South African longline fishery. The methodology follows that first introduced by da Silva et al. (2017) and uses a generalised additive mixed model (GAMM) applied to catch and effort data from the South African pelagic longline fleet operating during the period 2004 - 2024. The GAMM was fitted using a Tweedie distribution and included *year, month, latitude, longitude, fishing tactic (targeting)* as fixed factors and *vessel* as random effect. Targeting was determined by clustering PCA scores of the root-root transformed, normalized catch composition.

MATERIALS AND METHODS

CATCH AND EFFORT DATA PREPARATION

Catch and effort data for the period 2004-2024 were extracted from the South African longline logbook database. Each record included the following information: (1) date, (2) unique vessel number, (3) catch position at a 1 x 1 degree latitude and longitude resolution and (4) mandatory catch reports in kilogram per set and (5) hooks per set. Maps of the distribution of effort that distinguish between sets with zero swordfish catch ($y = 0$) and those with positive catch ($y > 0$), with proportional bubble sizes indicating relative CPUE (Figure 1). Criteria for filtering out potentially misreported logbook entries or outliers included the following; records without catch position (6 850 records removed), sets in which < 500 hooks were deployed (218 records removed) , and only data from east of 29 degrees (Longitude > 29) was considered to exclude the area of admixture, where stock originating in the two ocean basins cannot be distinguished (West 2016) (29 091 records removed). The final dataset contained 10 504 sets, 18 580 388 hooks and 68 distinct vessels.

MODEL FRAMEWORK

Swordfish CPUE was standardized using Generalized Additive Mixed Models (GAMMs), which included the covariates year, month, 1 x 1 degree latitude (Lat) and longitude (Long) coordinates and vessel as random effect. To account for variation in fishing tactics, we considered an additional factor for targeting derived from a cluster analysis of the catch composition (He et al., 1997; Carvalho et al., 2010; Winker et al., 2013). For the clustering analysis, all CPUE was modelled as catch in metric tons per species per vessel per day. All the following analysis was conducted within the statistical environment R. The R package 'cluster' was used to perform the CLARA analysis, while all GAMMs were fitted using the '*mgcv*' and '*nlme'* libraries described in Wood (2006).

Clustering of the catch composition data was conducted by applying a non-hierarchical clustering technique known as *CLARA* (Struyf et al., 1997) to the catch composition matrix. To obtain the input data matrix for *CLARA*, we transformed the $CPUE_{i,j}$ matrix of record $_i$ and species $_j$ into its Principal Components (PCs) using Principal Component Analysis (PCA). For this purpose, the data matrix comprising the $CPUE_{i,j}$ records for all reported species was extracted from the dataset. The CPUE records were normalized into relative proportions by weight to eliminate the influence of catch volume, fourth-root transformed and PCA-transformed. Subsequently, the identified cluster for each catch composition record was aligned with the original dataset and treated as categorical variable Fishing Tactic (FT) in the model (Winker et al., 2013). To select the number of meaningful clusters we followed the PCA-based approach outlined and simulation-tested in Winker et al. (2014). This approach is based on the selection of non-trivial PCs through non-graphical solutions (as opposed to the Catell's Scree test), called the Optimal Coordinate test alongside the Kaiser-Guttman rule (Eigenvalue > 1). The Optimal Coordinate test is available in the R package '*nFactors*' (Raiche et al., 2013). The optimal number of clusters considered is then taken as the number of retained PCs plus one (Winker et al., 2014). The results suggest that only the first PC is non-trivial (Fig. 2) and correspondingly two clusters were selected as optimal for the CLARA clustering.

The CPUE records were fitted by assuming Tweedie distribution (Tascheri et al., 2010; Winker et al., 2014). The Tweedie distribution belongs to the family of exponential dispersion models and is characterized by a two-parameter power mean-variance function of the form $Var(Y) = \phi\mu p$, where ϕ is the dispersion parameter, μ is the mean and p is the power parameter (Dunn and Smyth, 2005). Here, we considered the case of 1 < p < 2, which represents the special case of a Poisson (p = 1) and gamma (p = 2) mixed distribution with an added mass at 0. This makes it possible to accommodate high frequencies of zeros in combination with right-skewed continuous numbers in a natural way when modeling CPUE data (Winker et al., 2014; Ono et al., 2015). As it is not possible to estimate the optimal power parameter p internally within GAMMs, p was optimized by iteratively maximizing the profile log-likelihood of the GAMM for 1 < p < 2 (Fig. 3). This resulted in a power parameter p = 1.5 with an associated dispersion parameter of ϕ = 9.49 for the full GAMM. The full GAMM expressed swordfish CPUE as:

$$CPUE = exp(\beta_0 + Year + s1(Month) + s_2(Long, Lat) + FT + {}_{\alpha}V)$$

where s_1() denotes cyclic cubic smoothing function for Month, s_2() a thin plate smoothing function for the two-dimensional covariate of Lat and Long, FT is the vector of cluster numbers treated as categorical variable for 'Fishing Tactic', and α_v is the random effect for Vessel v (Helser et al., 2004). The inclusion of individual vessels as random effects term provides an efficient way to combine CPUE recorded from various vessels (n = 17) into a single, continuous CPUE time series, despite discontinuity of individual vessels over the time series (Helser et al., 2004). The main reason for treating vessel as a random effect was because of concerns that multiple CPUE records produced by the same vessel may violate the assumption of independence caused by variations in fishing power, skipper skills and behaviour, which can result in overestimated precision and significance levels of the predicted CPUE trends if not accounted for (Thorson and Minto, 2014). The significance of the random-effects structure of the GAMM was supported by both Akaike's Information Criterion (AIC) and the more conservative Bayesian Information Criterion (BIC). Sequential F-tests were used to determine the covariates that contributed significantly ($p < 0.001$) to the deviance explained.

Annual CPUE was standardized by fixing all covariates other than *Year* and *Lat* and *Long* to a vector of standardized values X_0. The choices made were that *Month* was fixed to July (*Month* = 11), representative of the high catch quarter and FT was fixed to the fishing tactic produced the highest average catch rates (FT = 1). The expected yearly mean $CPUE_y$ and standard-error of the expected

$log\ (CPUEy)$ for the vector of standardized covariates X_0 were then calculated as average across all

Lat-Long combinations (here forth grid cells) *a*, such that

$$E[CPUE(X_0^T\widehat{\beta})] = \frac{1}{A}\sum_{a}^{A} exp(\widehat{\mu}_{y,a})$$

and

$$\hat{\mu}(X_0^T\hat{\beta}) = \sqrt{\frac{1}{A}\sum_{a}^{A} \widehat{\sigma}_y^2\ , a}$$

where $\hat{\mu}_{y,a}$ is the standardized, model-predicted $log(CPUE_{y,a})$ for *Year y* and *Lat* and *Long* for grid cell a, $\hat{\sigma}_{y,a}$ is the estimated model standard error associated with $log(CPUE_{y,a})$), *A* is the total number of grid cells and *T* denotes the matrix in which *X* is transposed.

RESULTS AND DISCUSSION

Predicted values from the GAMM were back transformed to obtain yearly mean CPUE with 95% confidence intervals (Table 2). A normalised index was calculated by dividing each year's mean by the overall mean CPUE. Outputs included (1) A time series of standardized and normalized CPUE indices, (2) Random vessel effect estimates, (3) Residual vs. fitted plots and (4) Simulated quantile-quantile (QQ) plot using gratia:qq_plot (). These diagnostics confirmed model adequacy, with residuals displaying no major patterns and QQ plots falling within simulation envelopes.

The analysis of deviance for the stepwise regression procedure showed that all the covariates considered were highly significant ($p < 0.001$) and the inclusion of all considered fixed effects were supported by both the AIC and BIC (Table 1). Seasonality (*Month*) accounted for most of the deviance explained by the model (Table 1), followed by the inclusion of the effect of targeting other species (*Fishing tactic*), particularly tuna (Fig. 2). The inclusion of targeting, and the justifiable use of the Tweedie distribution (Figs. 3) has improved the model fit, however, further analyses could be considered.

The final model took the form:

SWFkg~factor(Year)+s(Month,bs="cc") + s(Lat,Lon)+(1|rv)

Determining the vessel specifications according to crew size and trip length are both deemed to be poor indicators of vessel type (Leslie et al., 2004; Smith and Glazer, 2007). It is challenging to obtain this vessel information (gross registered tonnage (GRT), length, use of live bait and sonar information) for the entire fleet, but a classification into vessel type was attempted in the past (Kerwath et al., 2012) based on maximum and average number of crew. However, there was no significant improvement in explanatory power by including vessel type as categorical variable or by using a subset of vessels from each class as indicator vessels. To include vessel as a random effect was deemed the most appropriate solution. There was notable variation among vessels (Fig. 4), and the inclusion of the random vessel effect produced the most parsimonious error structure. The random effect did not have a large effect on the confidence intervals.

The diagnostic plots (Figs 5-8) indicate that the model provides an acceptable fit for CPUE standardization, although some departures from model assumptions are evident. Deviations in the QQ plot suggest that the extreme tails of the distribution are not well captured, while the residuals versus fitted values show evidence of heteroscedasticity and localized underprediction. Despite these limitations, residuals are randomly distributed over time, with no indication of systematic temporal bias. This supports the use of the model outputs for deriving a relative abundance index.

Such departures are common in CPUE standardizations, particularly when analyzing highly variable and zero-inflated fisheries data using the Tweedie distribution (Shono 2008; Candy 2004). The level of misfit observed here is consistent with that reported in many standardization exercises and does not invalidate the resulting indices. CPUE data are inherently noisy, and some degree of lack of fit is expected. Provided that the standardized series is interpreted as a relative, rather than absolute, measure of abundance, the results remain scientifically defensible and appropriate for use in stock assessment and management advice (Maunder & Punt 2004).

In accordance with the previous analyses (da Silva et al., 2017; Parker et al 2017; Parker and Kerwath, 2020 and Parker, 2022), swordfish CPUE from the South African pelagic longline fishery displayed a definitive seasonal trend, with higher catch rates in austral winter (July - October) than the rest of the year (Fig. 6a). This may in part be due to the seasonal operations of Joint-Venture vessels operating predominantly off the East coast of South Africa, which pre-dominantly fish in the South African EEZ during the winter period. Standardized (normalized) CPUE has been relatively stable over the 20-year period, ranging from a low 0.60 in 2014 to a peak of 1.36 in 2023. The CPUE has shown a general increasing trend since 2021.

Future work should nevertheless explore whether alternative model formulations might provide a better fit to the data. Delta-lognormal or delta-gamma approaches (Stefánsson 1996) represent plausible alternatives that may reduce the influence of extreme residuals. In addition, the change in model fit compared with earlier analyses warrants further investigation. Potential explanations include changes in the temporal or spatial structure of the data (e.g., shifts in fishing effort or species composition), the inclusion or reparameterization of covariates, sensitivity to the estimated Tweedie power parameter, and the effect of larger or updated datasets that alter the balance between zero and positive catch rates. Evaluating these hypotheses within a structured model comparison framework would strengthen confidence in the robustness of the standardized CPUE index. The amendment of the catch return forms to include the target per catch day, sea surface temperature, bait type, hooks between floats and soak time could further improve the standardization of the CPUE data in this fishery

Table 1. Results from the GAMM applied to swordfish (*Xiphias gladius*) indicating the deviance explained by parameters selected for the final model.

Model	DF	AIC	BIC	Deviance	Deviance Explained	Percent Deviance Explained	P value
NULL	1	128354.7	128369.2	240394.1	2.91E-11	0	<0.001
Year	21	127716.7	127876.1	228340	12054.12	5.01	<0.001
Month	27	126894.2	127095.2	214038	26356.1	10.96	<0.001
Spatial	35	125135.5	125399.4	183503.7	56890.37	23.67	<0.001
Cluster	36	123404.2	123674.9	159762.6	80631.47	33.54	<0.001

Table 2. Normalised nominal and standardised CPUE values, including standard error (SE) and confidence intervals (LCI, UCI) for swordfish (*Xiphias gladius*) for the period 2004 - 2024.

Year	Nominal	CPUE	CV	LCI	UCI
2004	1.17	1.07	0.04	0.76	1.49
2005	0.63	1.11	0.04	0.80	1.55
2006	0.83	1.00	0.04	0.72	1.40
2007	0.78	0.83	0.04	0.60	1.16
2008	1.37	1.23	0.03	0.89	1.71
2009	0.84	0.83	0.04	0.59	1.15
2010	1.57	1.03	0.04	0.74	1.43
2011	1.63	1.06	0.04	0.77	1.47
2012	1.36	0.90	0.04	0.65	1.25
2013	1.30	0.89	0.04	0.64	1.23
2014	0.42	0.62	0.04	0.45	0.87
2015	0.75	0.83	0.04	0.60	1.16
2016	0.19	0.90	0.04	0.63	1.28
2017	0.09	1.11	0.04	0.76	1.60
2018	0.34	0.86	0.04	0.61	1.21
2019	0.71	0.93	0.04	0.67	1.30
2020	0.52	0.92	0.04	0.66	1.29
2021	1.20	1.07	0.04	0.77	1.48
2022	1.78	1.27	0.03	0.92	1.75
2023	2.00	1.36	0.03	0.98	1.89
2024	1.51	1.18	0.03	0.85	1.63

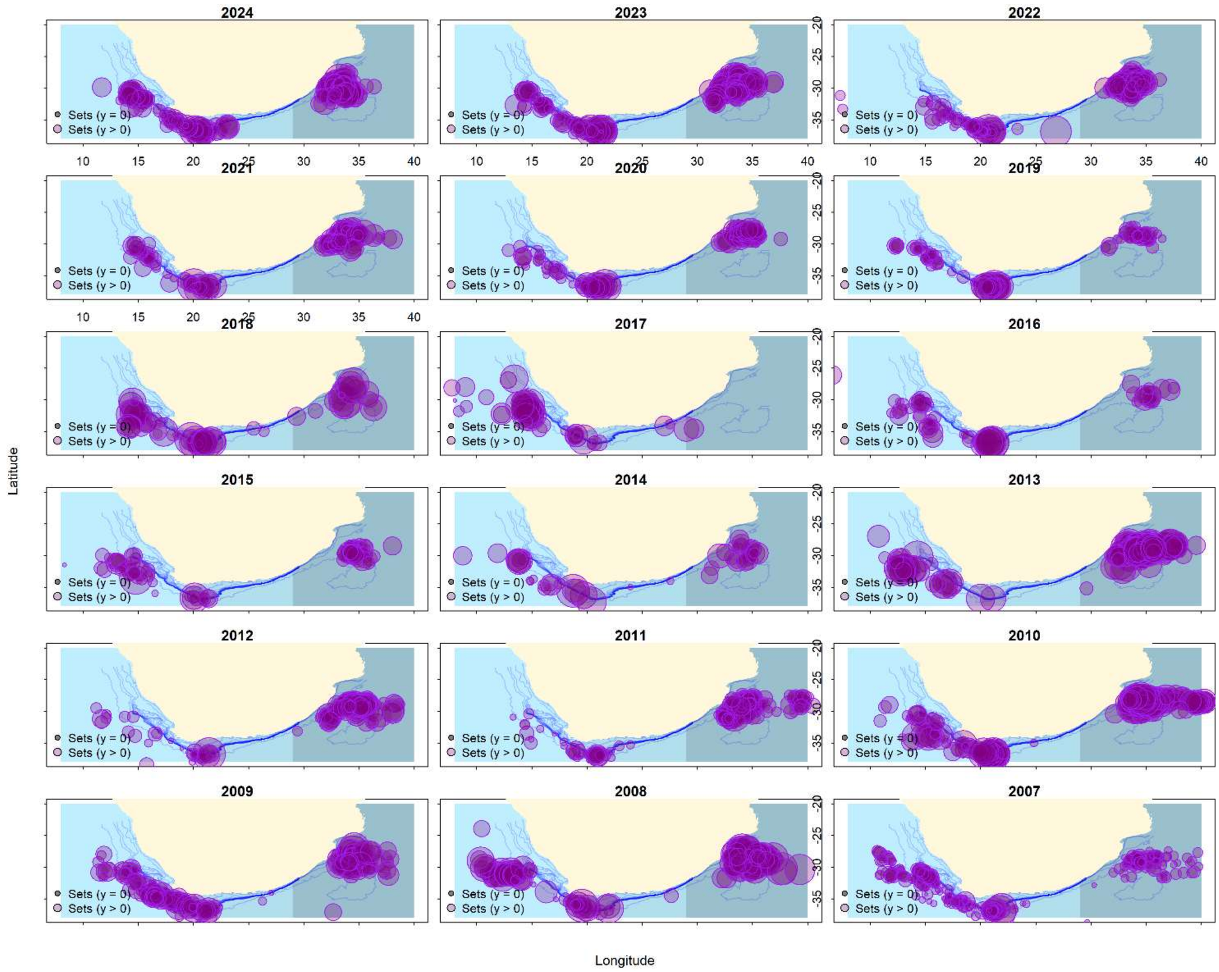


Figure 1. Annual effort distribution for the South African longline fleet. Longline sets that did not encounter a swordfish are the smallest circles, and the circle diameter increases proportional to the weight of swordfish caught per set. The black line indicates the ICCAT/IOTC boundary.

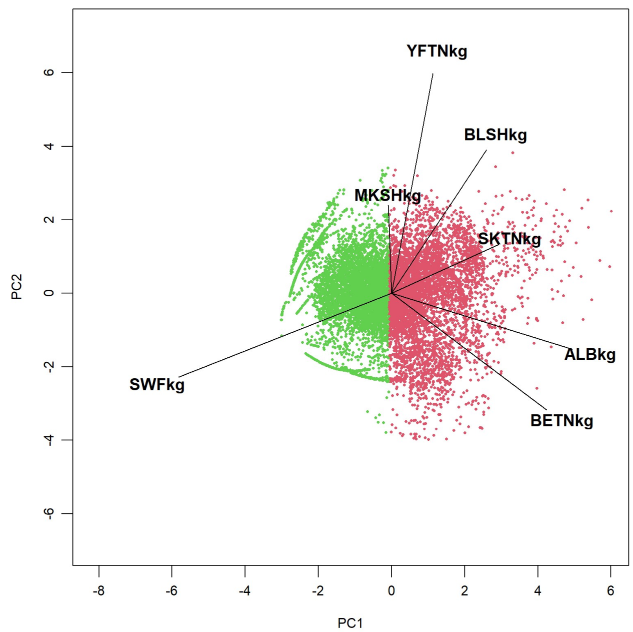


Figure 2. A graphical representation of the two clusters that characterise the different fishing tactics projected over the first two Principal Components (PCs), where only PC1 was determined to be non- trivial. Fishing Tactic 2 (FT1): Cluster one (red) is predominantly swordfish catches. Fishing Tactic 1 (FT2): Cluster two (green) is pre- dominantly tuna (ALB, BETN, SBT) with a mixture shark (blue and shortfin mako) catch.

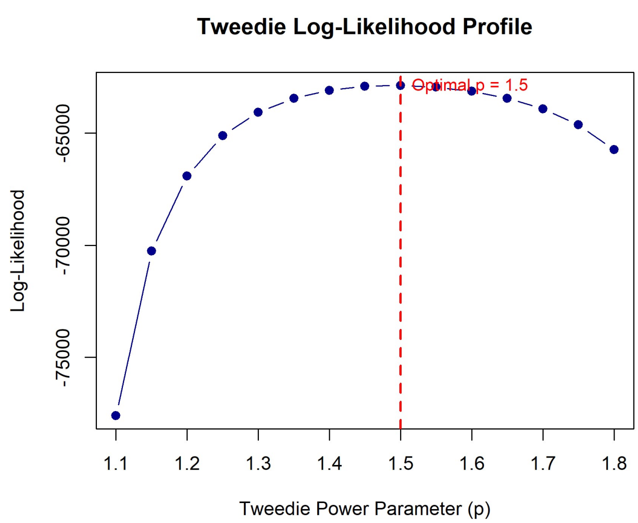


Figure 3. Log-likelihood profile for over the grid of power parameters values (1 < p < 2) of the Tweedie distribution. The vertical dashed line denotes the optimized p used in the final standardization GAMM.

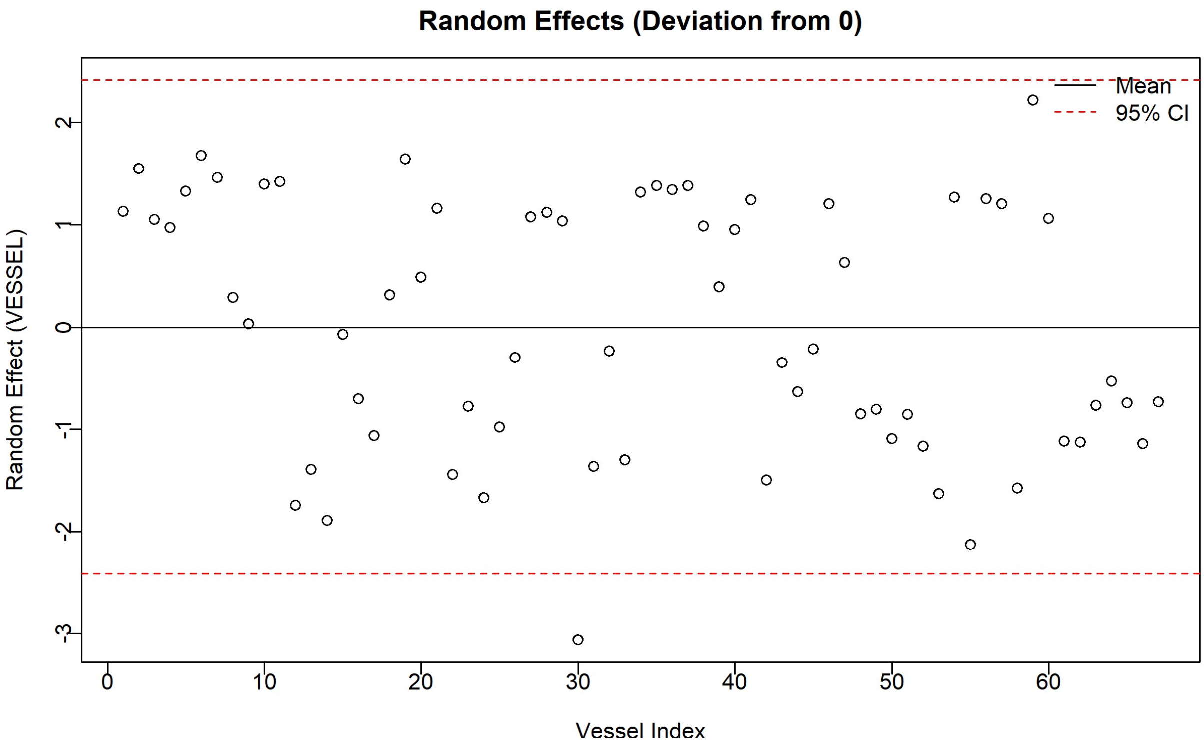


Figure 4. Random effects coefficients (dots) illustrating the deviation from the mean of zero across the 68 vessels retained for the analysis. Dashed lines denote the 95% confidence interval of the mean.

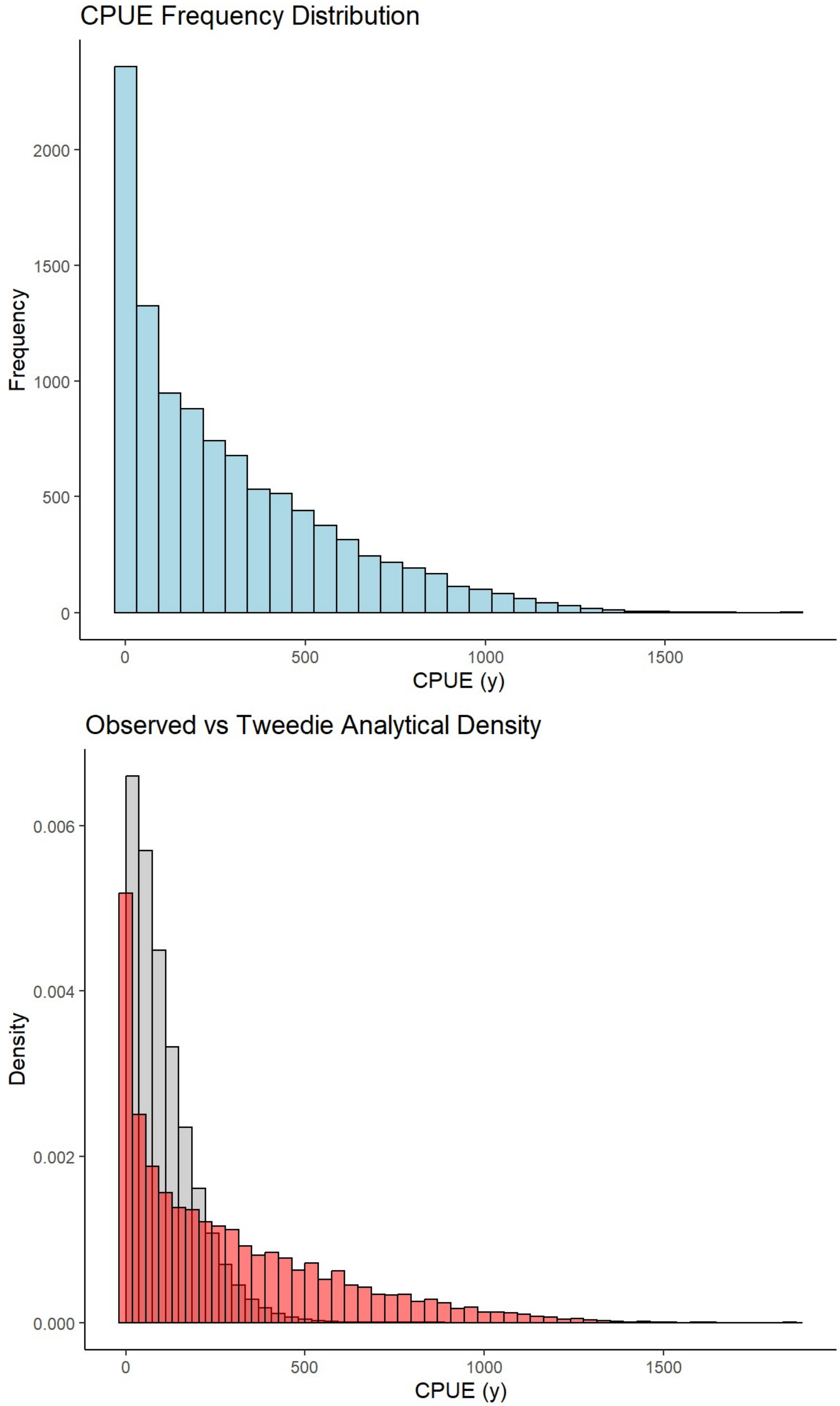


Figure 5. CPUE frequency, and density, distributions for the South African swordfish directed long- line fishery. The red line denotes the expected density of the response for the Tweedie GAMM, and supports the use of the Tweedie distribution form in the GAMMs.

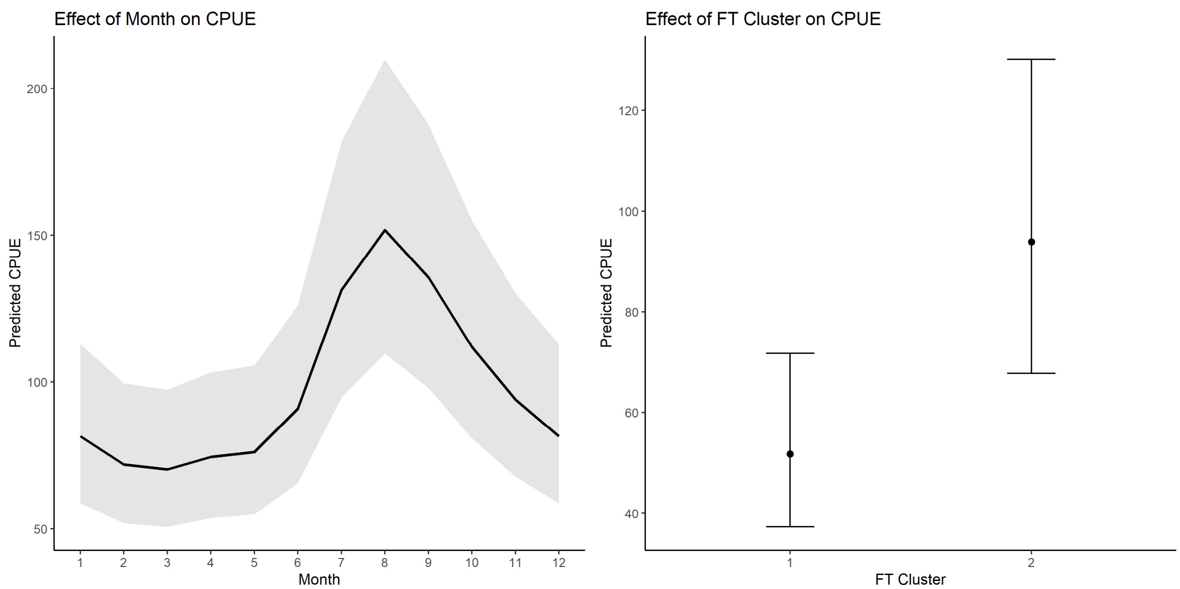


Figure 6. The influence of the fixed effects Month and Fishing Tactic on the CPUE of swordfish when modelled using the GAMM applied to the South African swordfish directed longline data.

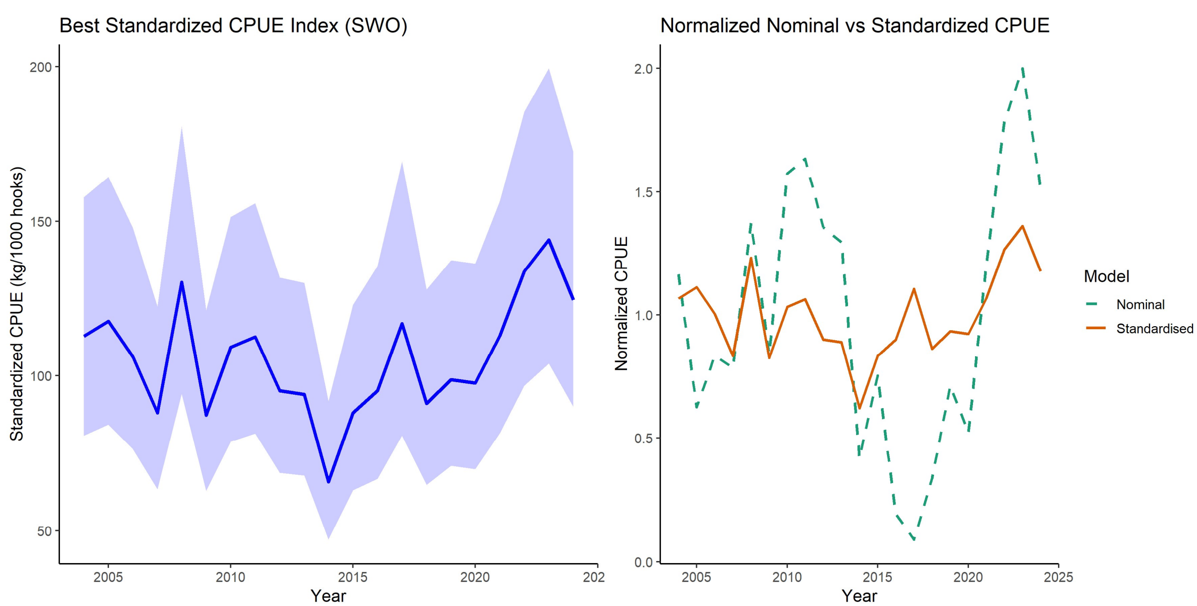


Figure 7. Standardized swordfish CPUE for the South African pelagic longline fishery for the period 2004 to 2024 (left panel). The 95% confidence intervals for the nominal CPUE are denoted by grey shaded areas. A comparison of nominal and the standardized CPUE models (right panel).

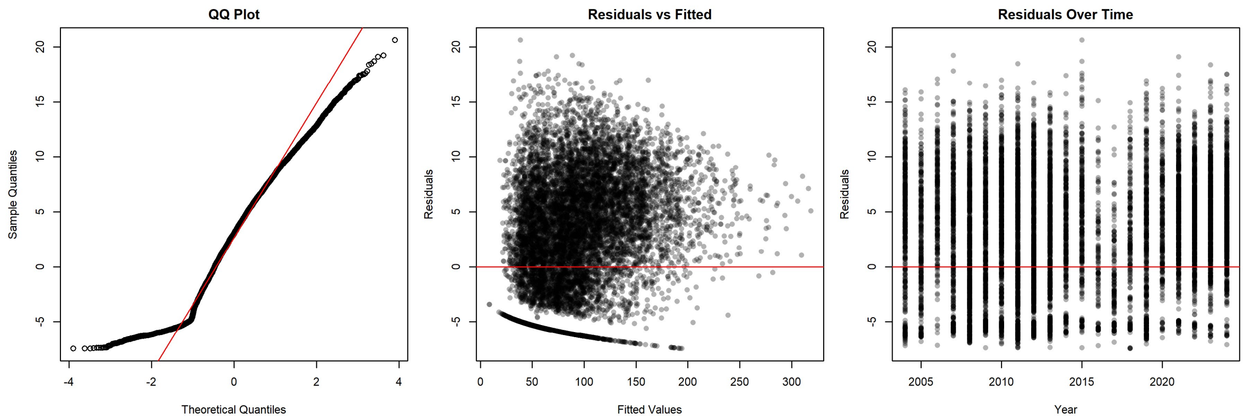


Figure 8. Diagnostic plots for the fitted CPUE standardization model: (a) normal Q–Q plot of residuals assessing normality, (b) residuals versus fitted values assessing homogeneity and potential nonlinearity, and (c) residuals plotted over time (by year) to check for temporal patterns or autocorrelation.

REFERENCES

Candy, S.G., 2004. Modelling catch and effort data using generalised linear models, the Tweedie distribution, random vessel effects and random stratum-by-year effects. *Ccamlr Science*, *11*(0), pp.59-80.

Carvalho, F. C., Murie, D. J., Hazin, F. H. V., Hazin, H. G., Leite-Mourato, B., Travassos, P., and Burgess, G. H. et al. 2010. Catch rates and size composition of blue sharks (*Prionace glauca*) caught by the Brazilian pelagic longline fleet in the southwestern Atlantic Ocean. *Aquatic Living Resources*, 23: 373–385. http://www.alr- journal.org/10.1051/alr/2011005.

da Silva, C., Parker, D., Winker, H., West, W., and Kerwath, S.E. 2017. Standardization of the catch per unit effort for swordfish (*Xiphias gladius*) for the South African longline fishery. IOTC-2017-WPB15-37.

Dunn, P. K., and Smyth, G. K. 2005. Series evaluation of Tweedie exponential dispersion model densities. Statis- tics and Computing, 15: 267–280.

He, X., Bigelow, K. A., and Boggs, C. H. 1997. Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery. *Fisheries Research*, 31: 147–158.

Helser, T. E. T. E., Punt, A. E., and Methot, R. D. 2004. A generalized linear mixed model analysis of a multi-vessel fishery resource survey. *Fisheries Research*, 70: 251–264. http://linkinghub.elsevier.com/retrieve/pii/S0165783604001705.

Kerwath, S. E., Winker, H., and West, W. M. 2012. Standardization of the catch per unit effort for albacore (*Thunnus alalunga*) for the South African tuna-pole (baitboat) fleet for the time series 1999-2010. *Col. Vol. Sci. Pap. ICCAT.,* 68: 604–614.

Leslie, R. W., Restrepo, V., and Antony, L. L. 2004. Standardised South Atlantic Albacore CPUE for the South African baitboat fishery., 1985-2002. *Col. Vol. Sci. Pap. ICCAT*, 56: 1504–1524.

Maunder, M.N. and Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries research, 70(2-3), pp.141-159.

Nepgen, C. S. 1970. Exploratory fishing for tuna off the South African west coast. Investigational report (South Africa. Division of Sea Fisheries).

Ono, K., Punt, A. E., and Hilborn, R. 2015. Think outside the grids: An objective approach to define spatial strata for catch and effort analysis. *Fisheries Research*, 170: 89–101. Elsevier B.V. http://linkinghub.elsevier.com/retrieve/pii/S0165783615001708.

Parker, D. and Kerwath, S.E., 2020. Standardized catch per unit effort of swordfish (*Xiphias gladius*) for the South African longline fishery. IOTC–2020–WPB18–13 Ward, P., Elscot, S. 2000. Broadbill swordfish: Status of world fisheries: Bureau of Rural Sciences, Canberra.

Parker, D., 2022.Standardised catch rates of swordfish (*Xiphias gladius*) for the South African pelagic longline fishery (2004-2020). *Collect. Vol. Sci. Pap. ICCAT*, *79*(2), pp.211-224.

Parker, D., Winker, H., West, W. and Kerwath, S.E., 2017. Standardization of the catch per unit effort for swordfish (*Xiphias gladius*) for the South African longline fishery. *Sci. Pap. ICCAT*, *74*, pp.1295-1305.

Penney, A. J., and Griffiths, M. H. 1999. A first description of the developing South African longline fishery. Coll. Vol. Sci. Pap., 49: 162–173.

Raiche, G., Walls, T. A., Magis, D., Riopel, M., and Blais, J. G. 2013. Non-graphical solutions for Cattell's scree test. Methodology, 9: 23–29.

Shono, H., 2008. Application of the Tweedie distribution to zero-catch data in CPUE analysis. *Fisheries Research*, 93(1-2), pp.154-162.

Smith, C. D., and Glazer, J. 2007. New standardized South Atlantic Albacore CPUE for the South African baitboat fishery, 1999-2005. Col. Vol. Sci. Pap. ICCAT., 60: 481–491.

Stefánsson, G. (1996). Analysis of groundfish survey abundance data: combining the GLM and delta approaches. *ICES Journal of Marine Science*, 53: 577–588

Struyf, A., Hubert, M., and Rousseeuw, P. J. 1997. Integrating robust clustering techniques in S-PLUS. Computational Statistics & Data Analysis, 26: 17–37.

Tascheri, R., Saavedra-Nievas, J.C.C., and Roa-Ureta, R. 2010. Statistical models to standardize catch rates in the multi-species trawl fishery for Patagonian grenadier (*Macruronus magellanicus*) off Southern Chile. Fisheries Research, 105: 200–214. http://linkinghub.elsevier.com/retrieve/pii/S016578361000127X.

Thorson, J. T., and Minto, C. 2014. Mixed effects: a unifying framework for statistical modelling in fisheries biology. *ICES Journal of Marine Science,* 70: 602–614. http://icesjms.oxfordjournals.org/cgi/doi/10.1093/icesjms/fsu213.

Welsh, J. G. 1968. A new approach to research on tuna in South African waters. Fisheries Bulletin of South Africa, 5: 32–34.

West, W. 2016. Genetic stock structure and estimation of abundance of swordfish (*Xiphias gladius*) in South Africa. MSc dissertation, University of Cape Town. 147 pp. Winker, H., Kerwath, S. E., and Attwood, C. G. 2013. Comparison of two approaches to standardize catch-per-unit-effort for targeting behaviour in a multispecies hand-line fishery. Fisheries Research,139:118–131.Elsevier B.V. http://linkinghub.elsevier.com/retrieve/pii/S0165783612003311.

Winker, H., Kerwath, S. E., and Attwood, C. G. 2014. Proof of concept for a novel procedure to standardize multi- species catch and effort data. *Fisheries Research*, 155: 149–159.

Wood, S. N. 2006. Generalized Additive Models: an introduction with R.