Food and Agriculture Organization of the United Nations

Indian Ocean Tuna Commission
Commission des Thons de l'Ocean Indien
iotc ctoi

[IOTC-2026-WPB24-20]

CPUE STANDARDISATION OF SWORDFISH CAUGHT BY LONGLINE IN SOUTH AFRICAN WATERS

WENDY M. WEST, W[1] AND CHARLENE DA SILVA, C[1,2]

Purpose

The paper presents an updated standardised CPUE index for swordfish caught by the South African pelagic longline fleet and describes the modelling framework used to account for changes in targeting behaviour, seasonality, spatial distribution and vessel effects. The paper invites discussion on the resulting abundance trend and the suitability of the index for future Indian Ocean swordfish stock assessments.

Discussions

The paper is to request review and comments on the CPUE standardisation methodology, including the selection of the Tweedie GAMM framework, treatment of fishing tactics, spatial averaging approach, and interpretation of the resulting abundance index

Feedback

The paper is to request review and comments on the CPUE standardisation methodology, including the selection of the Tweedie GAMM framework, treatment of fishing tactics, spatial averaging approach, and interpretation of the resulting abundance index

Conclusion

The paper is to seek confirmation that the analysis provides an appropriate and scientifically defensible standardised CPUE index for use in the assessment and monitoring of swordfish in the IOTC area

Recommendation

The paper is to seek endorsement of the updated standardised CPUE index for consideration during future swordfish stock assessment and scientific advice processes within the IOTC.

Decisions

The paper is to seek endorsement of the updated standardised CPUE index for consideration during future swordfish stock assessment and scientific advice processes within the IOTC.

Abstract

A standardised CPUE index for swordfish (Xiphias gladius) is presented, derived from 10 826 sets recorded by the South African pelagic longline fleet operating along the east coast of South Africa between 2004 and 2025. A tweedie generalised additive mixed model (GAMM) was used to account for zero inflation in the catch data. Covariates included year, month, fishing tactic, and a two-dimensional spatial smoother representing the spatial distribution of fishing effort, with vessel included as a random effect to account for vessel-specific differences in catchability. Model selection via the corrected Akaike Information Criterion (AICc), with all covariates contributing meaningfully to explaining variation in catch rates. Marginal means with spatial averaging were used to derive the standardised year index, isolating the year effect and spatial effect from the influence of seasonal and operational covariates. Residual diagnostics confirmed adequate model fit. The standardised CPUE index showed a stable trend up to 2017 and indicates an increase in abundance since. Swordfish CPUE had a definitive seasonal trend, with catch rates higher in winter (July - October) than in the rest of the year.

CPUE STANDARDISATION OF SWORDFISH CAUGHT BY LONGLINE IN SOUTH AFRICAN WATERS

Wendy M. West, W[1] and Charlene da Silva, C[1,2]

SUMMARY

A standardised CPUE index for swordfish (Xiphias gladius) is presented, derived from 10 826 sets recorded by the South African pelagic longline fleet operating along the east coast of South Africa between 2004 and 2025. A tweedie generalised additive mixed model (GAMM) was used to account for zero inflation in the catch data. Covariates included year, month, fishing tactic, and a two-dimensional spatial smoother representing the spatial distribution of fishing effort, with vessel included as a random effect to account for vessel-specific differences in catchability. Model selection via the corrected Akaike Information Criterion (AICc), with all covariates contributing meaningfully to explaining variation in catch rates. Marginal means with spatial averaging were used to derive the standardised year index, isolating the year effect and spatial effect from the influence of seasonal and operational covariates. Residual diagnostics confirmed adequate model fit. The standardised CPUE index showed a stable trend up to 2017 and indicates an increase in abundance since. Swordfish CPUE had a definitive seasonal trend, with catch rates higher in winter (July - October) than in the rest of the year.

KEYWORDS

CPUE, swordfish, longline, abundance, tweedie, GAMM

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, because 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 swordfish- 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 (**Figure 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 stock origin of South African caught swordfish remains uncertain. Population genetics suggest an admixture zone between 17°E and 30°E, and the conventional reporting boundary at 20°E may not reflect a biologically meaningful stock separation (West, 2016). Pending further investigation, the present analysis was restricted to data east of 29°E to minimise potential mixing with the Atlantic stock.

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 400 t in the ICCAT area of competence.

In this study, we standardise swordfish (*Xiphias gladius*) CPUE from the South African pelagic longline fishery, expressed as kg per 1000 hooks. Prior to fitting the final standardisation model, a model family comparison was conducted to evaluate error distributions that are suitable for accommodating the high proportion of zero catches that is characteristic of pelagic longline data.

1. Data and methods – catch and effort data preparation

Catch and effort data for the period 2004-2025 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, (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 or erroneous catch position (198 records removed), sets in which < 500 hooks were deployed (154 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) (30 593 records removed). The final dataset contained 10 826 sets.

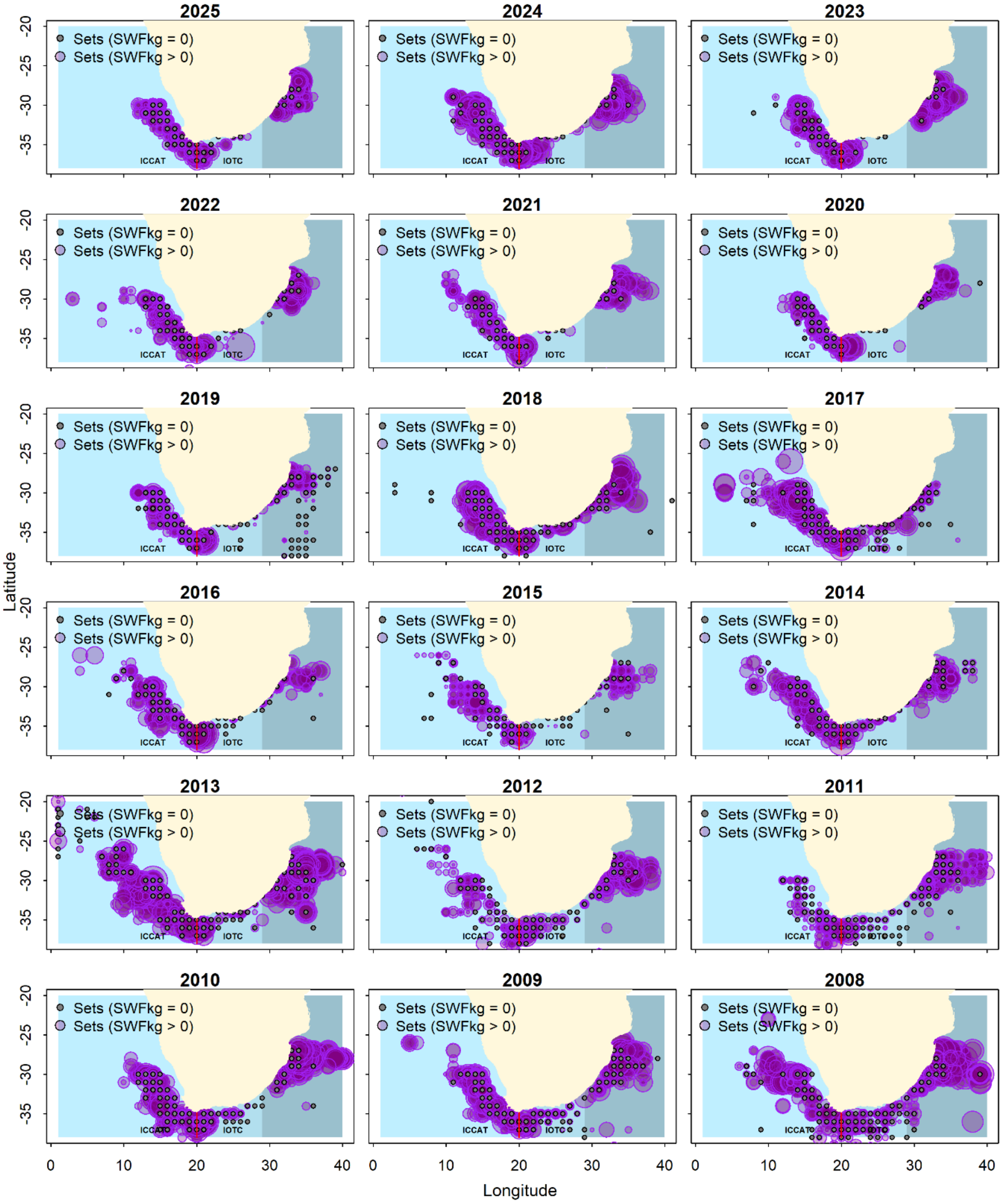

Figure 1. Annual effort distribution for the South African longline fleet. Longline sets that did not encounter a swordfish are the smallest circles, and the circle diameter increases proportional to the weight of swordfish caught per set. The red line indicates the ICCAT/IOTC boundary, and the region east of 29º E is a darker shade.

2. Model framework

2.1. *Fishing Tactic*

Fishing tactics were identified from the species composition of individual longline sets using a two-stage multivariate approach combining Principal Component Analysis (PCA) with cluster analysis (He et al., 1997; Carvalho et al., 2010; Winker et al., 2013).

2.1.1.Principal Component Analysis

The data matrix comprising the $CPUEi{,}j$ records for all reported species was extracted from the dataset and we transformed the $CPUEi{,}j$ matrix of record *i* and species *j* into its Principal Components (PCs) using Principal Component Analysis (PCA): the CPUE records were normalised into relative proportions by weight to eliminate the influence of catch volume, fourth-root transformed and PCA-transformed.

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).

2.1.2. Clustering

Clustering of the catch composition data was conducted by applying a non-hierarchical clustering technique known as CLARA (Clustering Large Applications) (Struyf et al., 1997) to the catch composition matrix. CLARA was applied to the retained PC scores to partition sets into distinct fishing tactics. For the clustering analysis, all CPUE was modelled as catch in metric tons per species per vessel per day. The R package 'cluster' was used to perform the CLARA analysis. 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).

2.2. Distribution model comparison

Model family comparison was conducted prior to variable selection to identify the error distribution that best described the observed data. The swordfish CPUE data in this study are characterised by a high proportion of zero catches combined with a continuous, right-skewed distribution of positive catch rates (**Figure 2**). Selecting an inappropriate family can produce biased parameter estimates, unreliable confidence intervals, and a misleading standardised index.

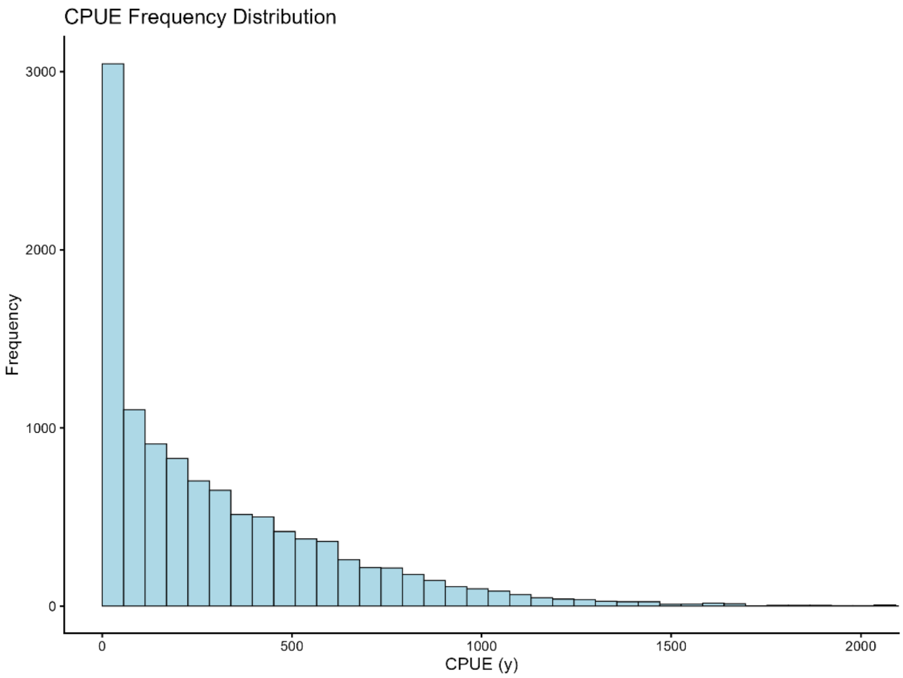

Figure 2. Nominal CPUE (swordfish kg.1000 hooks^{-1}) frequency distribution.

Three candidate distributions most commonly applied to zero-inflated CPUE data were evaluated for their suitability in describing the swordfish CPUE data in this study. The delta-lognormal and delta-gamma models (two-part hurdle, separately estimating the probability of capture, and the expected catch rate given presence), and the Tweedie distribution (single-model, a compound Poisson-gamma structure). All models were fitted using the R package 'sdmTMB' (Anderson et al. 2022), which supports all three candidate distributions evaluated in this study; and it incorporates latitude and longitude as a random spatial field through a stochastic partial differential equation (SPDE) approximation to a Gaussian Markov Random Field (GMRF) that accounts for spatial autocorrelation in the response variable. All candidate models included the categorical covariates Year, Month and FT, and Vessel as a random intercept effect. The three candidate models were compared using the Akaike Information Criterion (AIC), and the model with the lowest AIC was considered the best-supported distribution. The differences in AIC (dAIC) used to assess the relative support for each candidate: models with dAIC < 2 were considered to have equivalent support, those with dAIC between 2 and 7 to have moderate support, and those with dAIC > 10 to have negligible support relative to the best model (Burnham & Anderson 2002, Richards 2008).

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 $1 < p < 2$ (**Figure 3**). This resulted in a power parameter $p = 1.4$ with an associated dispersion parameter of $\phi = 16.64$ for the full GAMM, reflecting the high degree of overdispersion characteristic of swordfish longline CPUE data, in which catch rates are highly variable across sets due to the patchy spatial distribution of swordfish and the influence of unobserved set-level factors on catchability

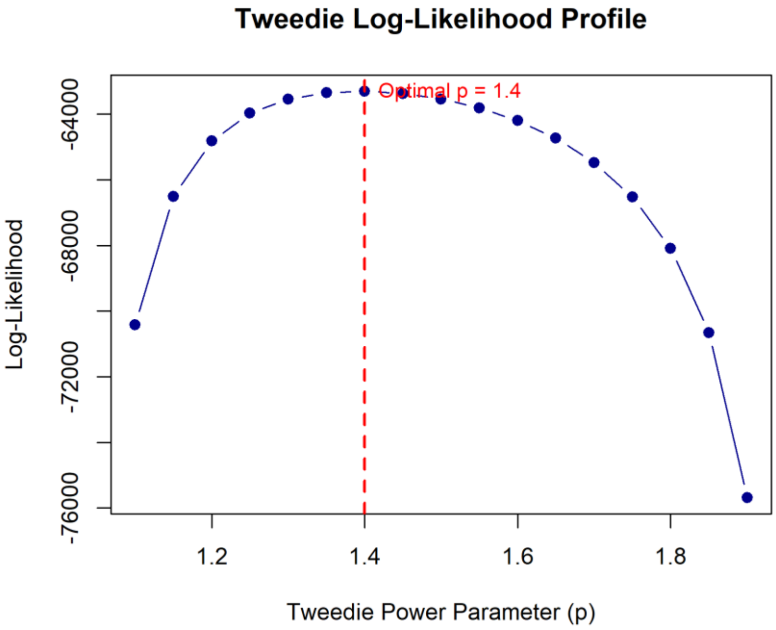

Figure 3. Log-likelihood profile for over the grid of power parameters values ($1 < p < 2$) of the tweedie distribution. The vertical dashed line denotes the optimized p used in the final standardization GAMM.

2.3. Candidate model

Swordfish CPUE was standardized using Generalized Additive Mixed Models (GAMMs), which included the covariates *year*, *month*, 1 × 1 degree latitude *(Lat*) and longitude (*Long*) coordinates, Fishing Tactic *FT* and *vessel* as random effect. CPUE was modelled as catch in kilogram per species per 1000 hooks. All GAMMs were fitted using the 'mgcv' and 'nlme' libraries described in Wood (2017). The full GAMM expressed swordfish CPUE as:

$$CPUE = \exp(\beta 0 + Year + s1(Month) + s2(Long, Lat) + FT + aV \qquad (1)$$

where:

- $s1()$ denotes cyclic cubic smoothing function for Month.
- $s2()$ a thin plate smoothing function for the two-dimensional covariate of Latitude and Longitude.
- FT is the vector of cluster numbers treated as categorical variable for 'Fishing Tactic'
- αV is the random effect for Vessel V.

The inclusion of individual vessels as random effects term provides an efficient way to combine CPUE recorded from various vessels (n = 62) 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).

2.3. Model selection

All candidate models included year as a fixed categorical effect, and the vessel random intercept (to account for between-vessel differences in catchability) (Maunder & Punt 2004).

Candidate models were constructed by removing each additional covariate, month, fishing tactic and spatial field, sequentially from the full model, following the principle that each variable should demonstrably worsen model fit (Burnham & Anderson 2002). This sequential approach is preferred over fitting all possible variable combinations simultaneously (i.e. backward stepwise selection) because it produces an interpretable model selection pathway and reduces the risk of overfitting that arises when many candidate models are compared without biological justification for their structure (Zuur et al. 2009). The candidate model set was therefore defined as follows, with each model representing a specific biological hypothesis about the drivers of swordfish catch rates:

Model 1 (full model):	Year	+ month	+ fishing tactic	+ spatial field	+ vessel
Model 2: (no spatial):	Year	+ month	+ fishing tactic		+ vessel
Model 3: (no FT):	Year	+ month		+ spatial field	+ vessel
Model 4: (no month):	Year		+ fishing tactic	+ spatial field	+ vessel
Model 5 (year only):	Year				+ vessel

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.

2.4. Standardisation of CPUE

Annual CPUE was standardized using the marginal prediction with spatial averaging approach, which involves fixing all covariates other than Year and Lat and Long (here forth spatial grid cells) to a vector of standardized values. The choices made were that Month was fixed to February (Month = 2), representative of the first quarter and FT was fixed to the fishing tactic the produced highest average catch rates (FT =2). For each year, the standardised CPUE index was calculated as the mean of the model-predicted CPUE across all spatial grid cells. This spatial averaging approach holds the distribution of fishing effort constant across years, ensuring that interannual changes in the standardised index reflect changes in albacore availability rather than shifts in where the fleet operated.

The expected yearly mean $CPUE_y$ and standard-error of the expected log($CPUE_y$) for the vector of standardized covariates $\mathbf{X}_0$ were then calculated as average across all *Lat-Long* combinations *a*, such that:

$$E[CPUE_y(\mathbf{X}_0^T\hat{\beta})] = \frac{1}{A}\sum_a^A \exp(\hat{\mu}_{y,a})$$
(2)

and

$$\hat{\sigma}_y(\mathbf{X}_0^T\hat{\beta})] = \sqrt{\frac{1}{A}\sum_a^A \hat{\sigma}_{y,a}^2}$$
(3)

where $\hat{\mu}_{y,a}$ is standardized, model-predicted log(*CPUE*$_{y,a}$) for year *y* and *Lat* and *Long* for grid cell . $\hat{\sigma}_{y,a}$ 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 of **X** is transposed.

Predicted values from the GAMM were back transformed to obtain yearly mean CPUE with 95% confidence intervals. The resulting index was normalised to a geometric mean of one across the time series to facilitate comparison with the nominal CPUE index and with indices derived from alternative model structures in the sensitivity analysis.

3. Results and conclusions

3.1. Fishing Tactic

The number of clusters was set to nPC + 1 and following evaluation of the scree plot (**Figure 4**) and optimal coordinates test, four (4) clusters were selected as the optimum solution. Four distinct fishing tactics (FT) were present in the data, each representing a different pattern of species composition across sets.

The PCA biplot displays each set as a point in the space defined by PC1 (x-axis) and PC2 (y-axis), coloured by its assigned fishing tactic (**Figure 5**). Species loadings are overlaid as vectors, where the direction of each vector indicates the species' contribution to each PC axis and the length reflects the strength of that contribution. Species with short vectors contribute little to the first two PCs and may be better represented by higher-order components.

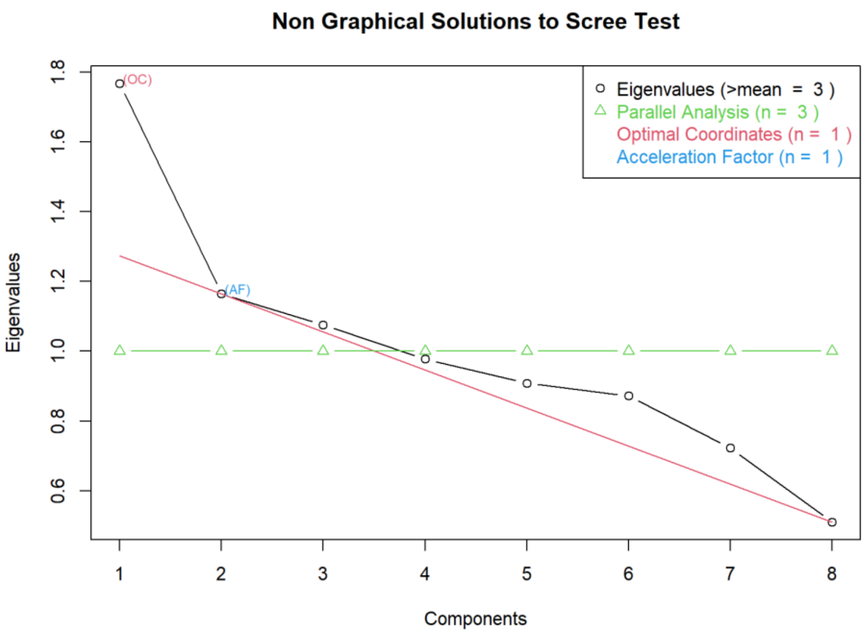

Figure 4. Scree plot of eigenvalues from the Principal Component Analysis (PCA) Kaiser criterion (eigenvalue > 1) and the Optimal Coordinates.

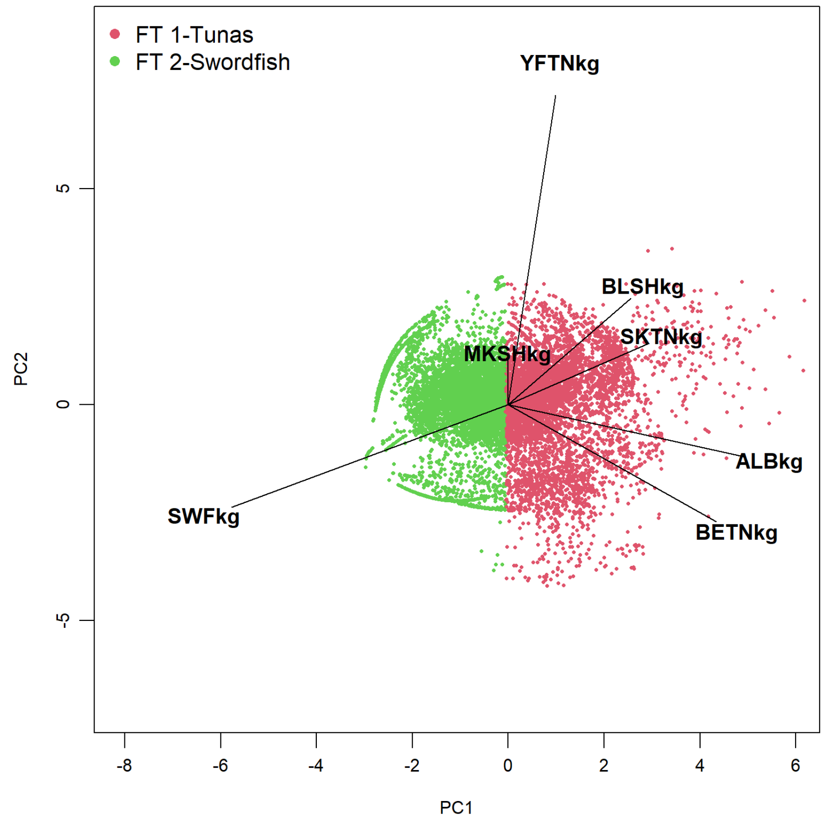

Figure 5. A graphical representation of the two clusters that characterise the different fishing tactics projected

over the first two Principal Components (PCs). Fishing Tactic 1 (FT1): Cluster one (red) is predominantly tunas (yellowfin, bigeye, albacore) and shark (blue and shortfin mako) catch. Fishing Tactic 2 (FT2): Cluster two (green) is swordfish.

3.2. Distribution model comparison

Although the delta-gamma model provided the best fit to the data based on AIC, the Tweedie model also received substantial support, with a ΔAIC of only 3.65 relative to the top-ranked model (**Table 1**). The previous standardisation conducted by da Silva et al. (2017; 2025) employed a Tweedie distribution, providing a valuable basis for comparison with the updated analysis. Given the relatively small difference in model support and the importance of maintaining methodological consistency between assessments, the Tweedie model was selected as the final distribution for index standardisation. This approach facilitates direct comparison of abundance trends through time while retaining a model structure that adequately describes the characteristics of the CPUE data.

Table 1. Comparison of statistical distribution families fitted to swordfish longline CPUE data, ranked by Akaike Information Criterion (AIC). All models were fitted with identical fixed and random effect structures. Δ AIC is the difference in AIC relative to the best-fitting model; models with Δ AIC < 2 are considered equivalent, those with 2 < Δ AIC < 7 have moderate support, and Δ AIC > 10 to have negligible support relative to the best model. df = degrees of freedom.

Model	Distribution	df	AIC	Δ AIC
Delta-Gamma	Binomial + Gamma	75	126941.6	0.0000

Tweedie	Tweedie	39	126945.2	3.65
Delta-Lognormal	Binomial + Lognormal	75	127540.4	598.82

3.3. Model selection

Model selection via AICc identified the full model, incorporating year, a cyclic seasonal smoother, a two-dimensional spatial smoother, fishing tactic as a fixed effect, and vessel as a random intercept effect, as the most parsimonious description of swordfish CPUE (**Table 2**). The full model achieved the lowest combined AIC and BIC scores, outperforming all reduced candidate models in which individual covariates were systematically removed. The margin by which the full model outperformed each reduced model (dAIC > 2 in all cases) indicates that each covariate contributed meaningfully to explaining variation in swordfish CPUE, and that no variable could be removed without a meaningful loss of explanatory power. The analysis of deviance for the step-wise regression procedure showed that all the covariates considered were highly significant ($p < 0.001$) (**Table 2**). Total deviance explained by the model was 56.19%

The selection of the full model confirms that swordfish catch rates in this longline fishery were influenced by a combination of operational and spatiotemporal factors. Month captured seasonal variability in swordfish availability or fleet behaviour, area reflected broad-scale spatial differences in habitat quality or environmental conditions, and fishing target accounted for differences in catchability associated with directed versus incidental swordfish effort. Vessel was retained as a random effect throughout, accounting for unobserved vessel-specific differences in fishing power and reporting behaviour that are unrelated to underlying swordfish abundance.

The final model took the form:

$$CPUE = \exp(\beta 0 + Year + s1(Month) + s2(Long, Lat) + FT + aV \quad (4)$$

Table 2. Results from the GAMM applied to swordfish (*Xiphias gladius*) indicating the deviance explained by

parameters selected for the final model.

Model	Formula	DF	AIC	BIC	Deviance	Deviance explained	% deviance explained	P value
Model 1 (full)	Year +month+ fishing tactic + spatial field + vessel	87	126809.8	127620.5	185530.6	237926.7	56.19	<0.001
Model 2 (no spatial):		72	127289.8	127987.1	193518.2	229939.1	54.3	<0.001

Model 3 (no FT):	88	127741.3	128549.6	199828.8	223628.5	52.81	<0.001
Model 4 (no month):	80	128351.9	129110.5	209090.3	214367	50.62	<0.001
Model 5 (null):	66	130160.5	130806.7	241422.8	182034.5	42.99	<0.001

3.4. Standardisation of CPUE

Diagnostic evaluation of the tweedie model indicated a satisfactory fit to the observed swordfish CPUE data (**Figure 6**). The distribution tracks the reference line of the QQ plot reasonably well, suggesting that the model describes average catches with only minor deviations in the upper tail. The residuals versus linear predictor plot showed no strong systematic trend, indicating that the model adequately described the mean structure of the data. However, the spread of residuals increased slightly with increasing fitted values, reflecting the heteroscedasticity characteristic of catch-rate data and accommodated by the Tweedie distribution. A distinct lower boundary of negative residuals was evident, which is typical of Tweedie models fitted to data containing many low or zero observations. The histogram of deviance residuals was centred close to zero and approximately symmetric, although slightly right-skewed due to a small number of unusually large catches. The observed versus fitted values plot demonstrated that the model captured the overall relationship between observed and predicted CPUE, with increasing variability at higher fitted values. Several extreme observations were underestimated by the model, but these represented a small proportion of the dataset.

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). Overall, the diagnostic plots suggest that the Tweedie GAMM provided a reasonable description of the CPUE data and was suitable for deriving a standardised abundance index.

Residual plots showed no systematic patterns across years, months, vessels, or fishing target categories, confirming that the covariates included in the full model adequately captured the major sources of variation in swordfish CPUE and that no substantial model structure was left unexplained (**Figure 7**).

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). In past analyses (Kerwath et al., 2012) there was no significant improvement in explanatory power by including vessel type (based upon gross registered tonnage (GRT), length, use of live bait and sonar information) 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. The inclusion of the random vessel effect produced the most parsimonious error structure.

In accordance with the previous analyses (da Silva et al., 2017; Parker et al 2017; Parker and Kerwath, 2020, Parker, 2022 and da Silva et al. 2025), swordfish CPUE from the South African pelagic longline fishery displayed a definitive seasonal trend, with higher catch rates in austral winter (June - October) than the rest of the year (**Figure 8.a**). This may be linked to suitable environmental conditions such as stable structured thermal fronts and eddies that attract prey aggregations and large pelagic predators in the South African EEZ during the winter period. The elevated swordfish CPUE associated with Tactic 2 (swordfish) is as expected (**Figure 8.b**).

The standardised swordfish CPUE index derived from the current analysis (2004–2025) was compared with the previously published index of da Silva et al. (2017), which covered the period 2004–2016 (**Figure 9**). To facilitate comparison of temporal trends, each index was independently normalised by its own geometric mean, such that the value of one represents the average relative abundance over the respective time series.

The two indices exhibited a high degree of agreement throughout the overlapping period (2004–2016), displaying similar interannual patterns in relative abundance. Both indices identified periods of above-average abundance in 2008 and 2011 and a pronounced decline to a common minimum in 2014. The previous index generally produced slightly higher relative abundance estimates than the current analysis during most years between 2004 and 2016, although differences were small and the overall temporal trends were consistent.

Extending the time series to 2025 revealed a sustained increase in the standardised CPUE index following 2020 (**Figure 10, Table 3**). Relative abundance increased markedly from 2021 onwards, reaching its highest level in 2025 at approximately 1.6 times the geometric mean. Confidence intervals widened in the latter years of the series, indicating greater uncertainty in recent index estimates. Nevertheless, the extended time series suggests that swordfish relative abundance has increased substantially since the low observed in 2014 and has remained above the long-term average since 2021.

Changes in the temporal or spatial structure of the data such as shifts in fishing effort distribution or species composition, the inclusion or reparameterisation of covariates, and the effect of the updated dataset altering the balance between zero and positive catch rates could be explored with a spatiotemporal modelling approach such as sdmTMB (Anderson et al. 2022), which would strengthen confidence in the robustness of the standardised swordfish CPUE index going forward.

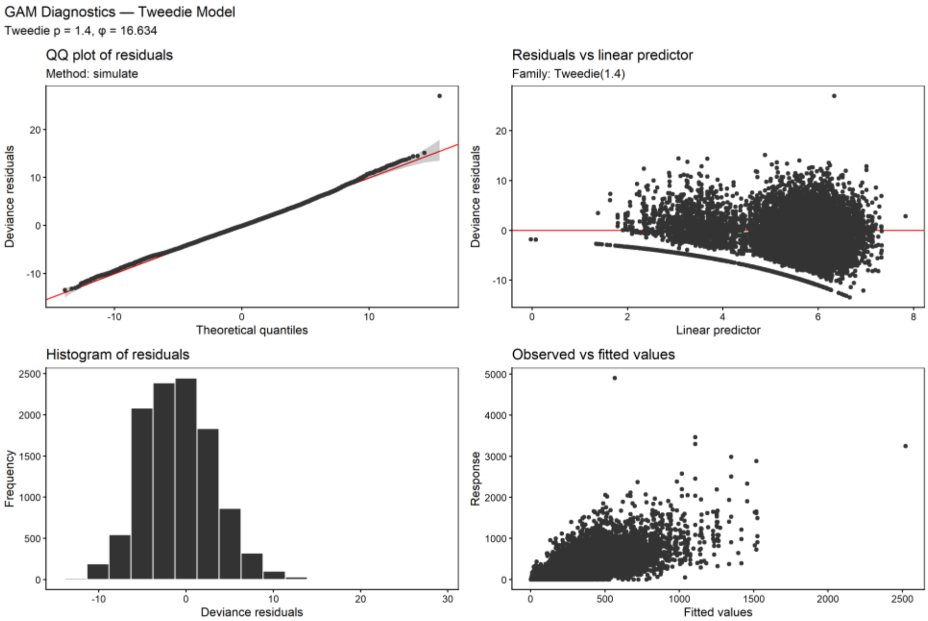

Figure 6. Diagnostic evaluation of the tweedie standardisation model.

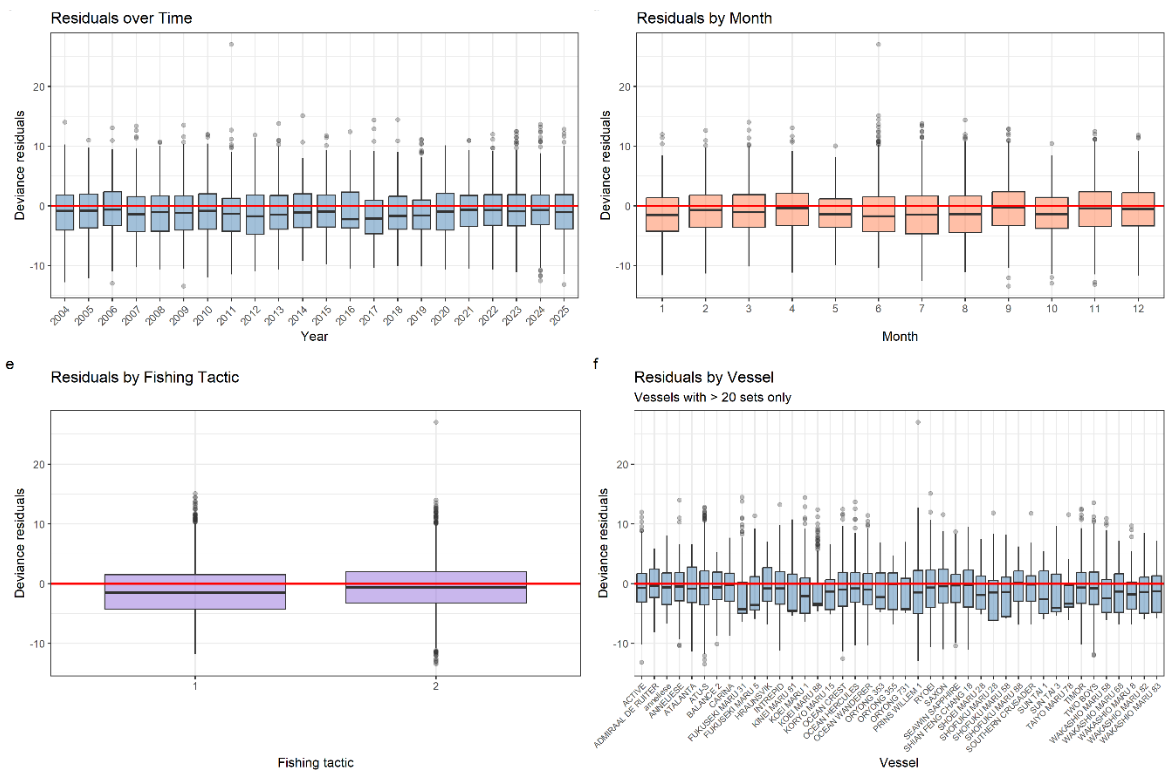

Figure 7. Residual plots of the tweedie standardisation showed no systematic patterns across years, months, vessels, or fishing target categories.

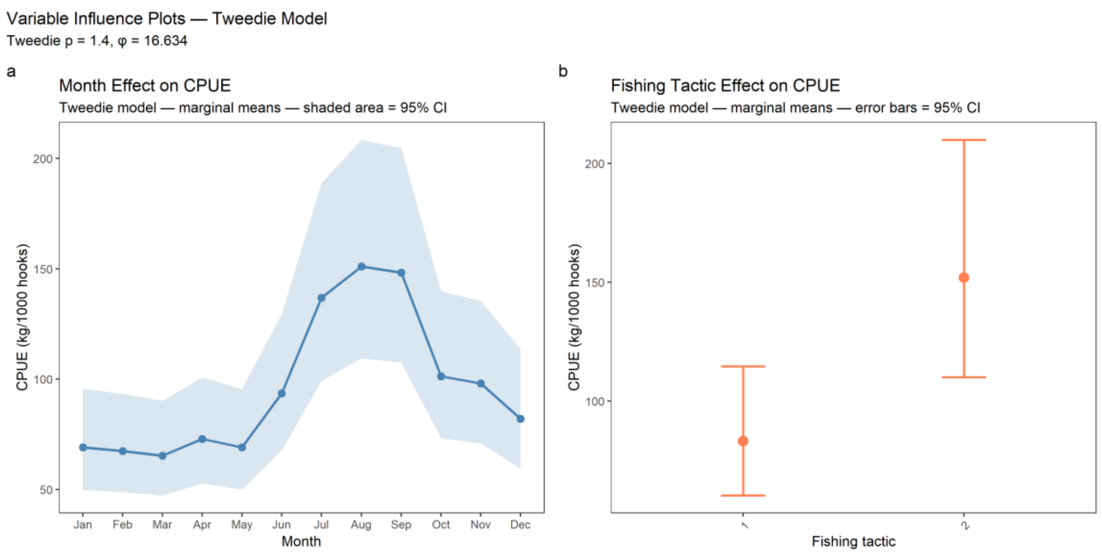

Figure 8. The influence of the fixed effects Month and Fishing Tactic on the CPUE of swordfish when

modelled using the tweedie model applied to the South African longline data.

Table 3. Normalised nominal and standardised CPUE values, including coefficient of variation (CV, %) and confidence intervals (LCI, UCI) for swordfish (*Xiphias gladius*) for the period 2004 – 2025.

Year	Nominal CPUE	Standardised CPUE	CV (%)	LCI	UCI
2004	1.31	1.03	17.06	0.74	1.44
2005	1.22	1.08	17.05	0.77	1.51
2006	1.80	0.96	16.97	0.69	1.34
2007	0.96	0.81	16.88	0.58	1.13
2008	0.96	1.18	16.67	0.85	1.64
2009	0.82	0.79	16.82	0.57	1.11
2010	1.37	1.02	16.75	0.73	1.41
2011	1.04	1.05	16.65	0.76	1.46
2012	0.86	0.92	16.67	0.66	1.27
2013	0.83	0.86	16.70	0.62	1.19
2014	0.67	0.61	17.17	0.43	0.85
2015	1.22	0.82	17.04	0.59	1.14
2016	0.73	0.82	18.29	0.57	1.17
2017	0.22	1.15	19.68	0.78	1.69
2018	0.84	0.86	17.58	0.61	1.22
2019	0.79	0.90	17.02	0.64	1.25
2020	1.09	0.92	17.23	0.66	1.29
2021	1.09	1.09	16.83	0.78	1.52

2022	1.36	1.39	16.72	1.00	1.92
2023	1.57	1.47	16.68	1.06	2.04
2024	1.17	1.26	16.74	0.90	1.74
2025	1.75	1.58	16.76	1.13	2.19

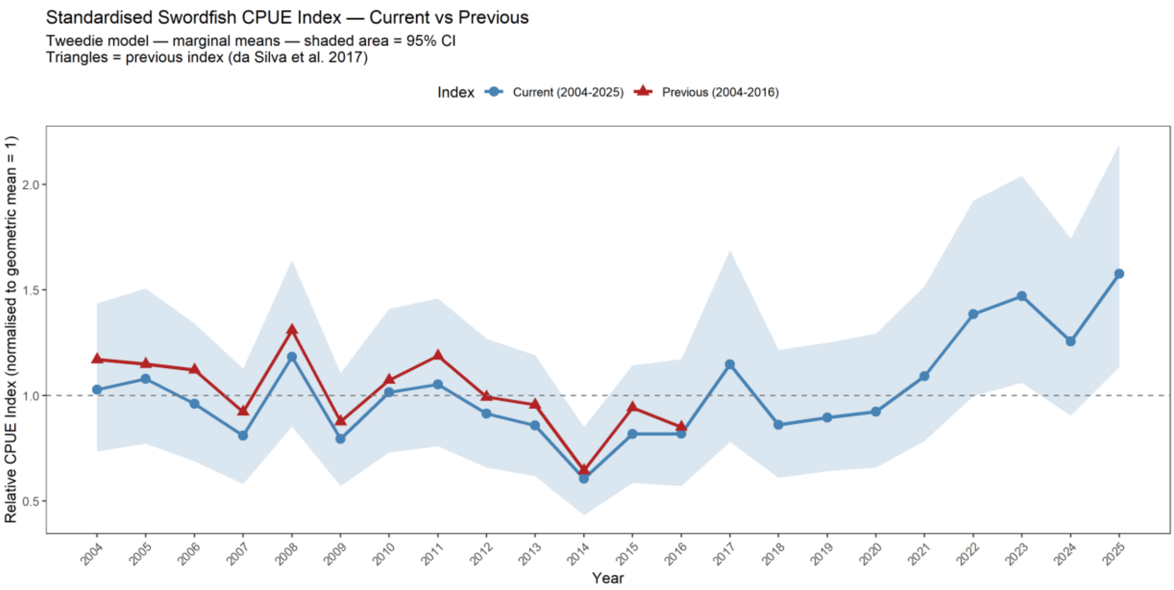

Figure 9. Comparison of the previously published standardised swordfish CPUE index (2004–2016; da Silva et al. 2017) and the updated standardised CPUE index (2004–2025) derived from the present analysis using a Tweedie GAMM.

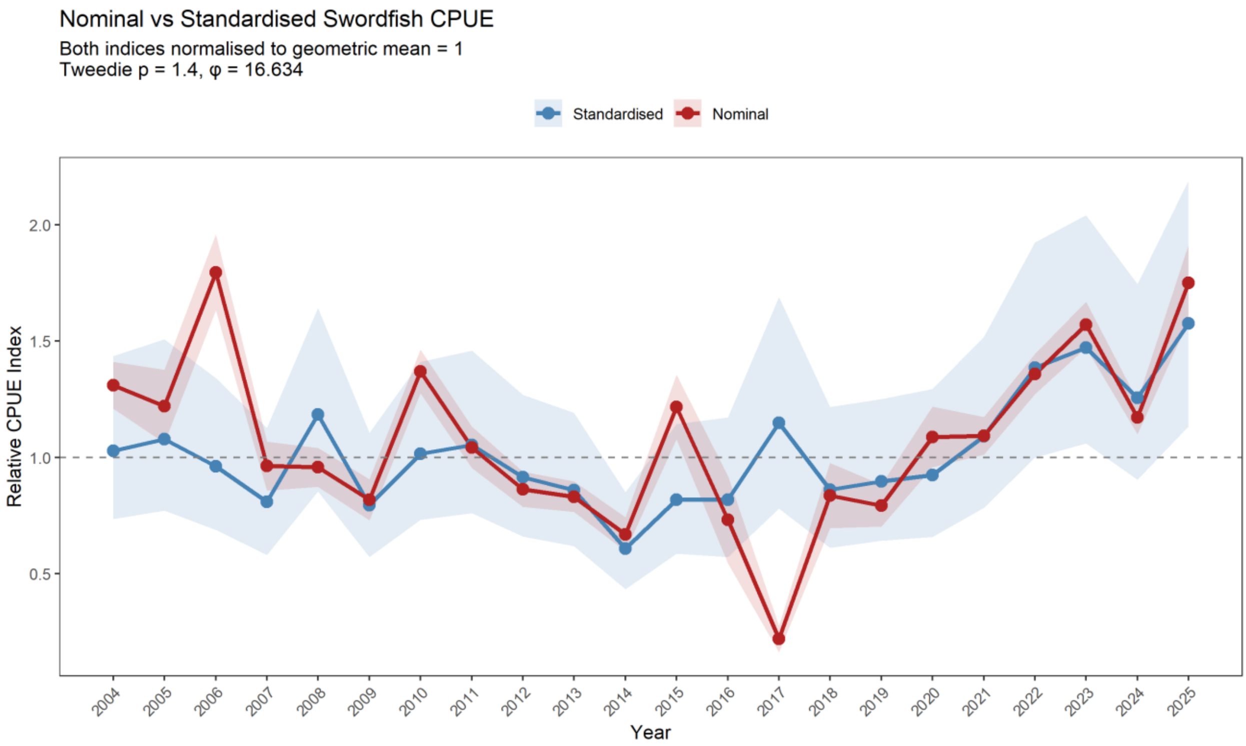

Figure 10. Standardised and nominal swordfish CPUE for the South African pelagic longline fishery for the period 2004 to 2025, normalised by the geometric mean. The 95% confidence intervals are denoted by shaded areas.

References

Anderson, S.C., Ward, E.J., English, P.A., Barnett, L.A.K., and Thorson, J.T. 2022. sdmTMB: an R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. bioRxiv 2022.03.24.485545. https://doi.org/10.1101/2022.03.24.485545

Bigelow, K.A., Boggs, C.H., and He, X. 1999. Environmental effects on swordfish and blue shark catch rates in the US North Pacific longline fishery. Fisheries Oceanography 8, 3: 178–198.

Burnham, K.P. and Anderson, D.R. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd edition. Springer, New York.

Carruthers, T.R., Ahrens, R.N.M., McAllister, M.K., and Walters, C.J. 2011. Integrating imperfect monitoring data into the estimation of historical dynamics and MSY for two species of Pacific tuna. Fisheries Research 108,1: 119–133.

Carvalho, F. C., Murie, D. J., Hazin, F. H. V., Hazin, H. G., Leite-Mourato, B., Travassos, P., and Burgess, G. H. 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.

da Silva, C., West, W.M. and Kerwath, S.E. CPUE Standardisation of swordfish (*Xiphias gladius*) caught in South African waters. IOTC-2025-WPB23-22.

Gelman, A. and Hill, J. 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Cambridge.

Gilman, E., Boggs, C., and Brothers, N. 2006. Performance assessment of an underwater setting chute to mitigate seabird bycatch in the Hawaii pelagic longline tuna fishery. Ocean and Coastal Management 48,3–4: 235–260.

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.

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.

Lenth, R.V. 2024. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.10.1. https://CRAN.R-project.org/package=emmeans

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.

Lindgren, F., Rue, H., and Lindström, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society Series B 73,4: 423–498.

Lo, N.C., Jacobson, L.D., and Squire, J.L. 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Canadian Journal of Fisheries and Aquatic Sciences 49,12: 2515–2526.

Maunder, M.N. and Punt, A.E. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70,2–3: 141–159.

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

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.

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

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.

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.

Poisson, F., Gaertner, J.C., Fauvel, C., Durbec, J.P., and Taquet, M. 2010. Effects of lunar cycle and fishing operations on longline-caught pelagic fish: fishing performance, capture depth, and bait consumption. Fisheries Research 106,3: 372–382.

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.

Richards, S.A. 2008. Dealing with overdispersed count data in applied ecology. *Journal of Applied Ecology* 45,1: 218–227.

Robins, C.M., Wang, Y.G., and Die, D. 1998. The impact of global positioning systems and plotters on fishing power in the northern prawn fishery, Australia. Canadian Journal of Fisheries and Aquatic Sciences 55,7: 1645–1651.

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

Smith, E.P. 1988. Estimation of bias in a nonlinear model. *Biometrics* 44, 2: 583–590.

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.

Shelton, A.O., Thorson, J.T., Ward, E.J., and Feist, B.E. 2014. Spatial semiparametric models improve estimates of species abundance and distribution. Canadian Journal of Fisheries and Aquatic Sciences 71,11: 1655–1666.

Smith, E.P. 1988. Estimation of bias in a nonlinear model. Biometrics 44,2: 583–590.

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. 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*(3), pp.200-214.

Thorson, J.T., Shelton, A.O., Ward, E.J., and Skaug, H.J. 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfish. ICES Journal of Marine Science 72,5: 1297–1310.

Ward, P. and Myers, R.A. 2005. Inferring the depth distribution of catchability for pelagic fishes and correcting for variations in the depth of longline fishing. Canadian Journal of Fisheries and Aquatic Sciences 62,5: 1130–1142. Sonnet 4.6

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. 2017. *Generalized Additive Models: An Introduction with R*. 2nd edition. Chapman and Hall/CRC, Boca Raton.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., and Smith, G.M. 2009. *Mixed Effects Models and Extensions in Ecology with R*. Springer, New York.