STANDARDIZATION OF THE CATCH PER UNIT EFFORT FOR SWORDFISH (*XIPHIAS GLADIUS*) FOR THE SOUTH AFRICAN LONGLINE FISHERY

Denham Parker, Henning Winker, Wendy West & Sven E. Kerwath

2017-06-21 19:22:56

Summary

South Africa has two commercial fishing sectors which either target, or catch as bycatch, tuna and tuna-like

Keywords

Swordfish, standardized cpue, longline

Introduction

Commercial fishing for large pelagic species in South Africa dates back to the 1960s (*Welsh, 1968 and 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 linefishing. Pelagic longline vessels are the only vessels that target swordfish, with some bycatch being caught by mid-water trawl vessels. South African swordfish do not display basking behaviour (*West et al., 2012*) and hence does not support a harpoon fishery; and gillnet fishing in South Africa is prohibited. 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 and 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 swordfish- and tuna-directed vessels. On average, 15 South African vessels are active in a year and target swordfish in 20-30m length vessels. Additionally, foreign flagged vessels catch swordfish as bycatch. South Africa's swordfish catches reached a peak in 2002 at 1 187 t, and have been on the decline with average catches of 372 t for the period 2009-2014. 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. 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. Consequently, all tunas and billfish stocks with the exception of the southern bluefin tuna (*Thunnus maccoyii*), are artificially divided into Atlantic and Indian Ocean stocks along this boundary, regardless of their true stock structure and distribution. Since questions remain about the origin of South African caught swordfish, it remains uncertain if the artificial split in reporting stock indices indeed reflects a biological meaningful separation of stocks.

In 2013 South Africa presented a standardized catch rate of swordfish (in number) caught by the South Africa longline fleet in the South Atlantic Ocean between 1998 and 2012 (SCRS/2013/159).The Group acknowledged the effort made and recommended further improvement regarding the model formulation and the predictions for extracting the year effect on the standardized index. The Group decided not to include this series in the stock assessmentm modelling process.

Here we present improved standardised catch-per-unit-effort (CPUE) indices that were obtained with a generalised additive mixed model (GAMM) of swordfish catch and effort data from the South African pelagic longline fleet operating in the South Atlantic Ocean between 2004 and 2015. Catch and effort data were subsetted to the ICCAT area of the South Atlantic Ocean (<20°E). The GAMM was fitted using a tweedie distribution and included year, month, latitude, longitude, fishing tactic (targeting) as fixed factors and had a random vessel effect. Targeting was determined by clustering PCA loadings of the root-root transformed, normalized catch composition.

Materials and Methods

Catch and effort data preparation

Model framework

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 and vessel as random effect. In an attempt 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 of 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. 1996*) to the catch composition matrix. To obtain the input data matrix for CLARA, we transformed the CPUEi,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 CPUEi,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 (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 for Catell's Scree test in association with the Kaiser-Guttman rule (Eigenvalue > 1), called Optimal Coordinate test, which available in the R package 'nFactors' (*Raîche 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. 3) and correspondingly two clusters were selected as optimal for the CLARA clustering.

The CPUE records were fitted by assuming Tweedie distribution (*Candy 2004, 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 (*Candy 2004, 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 modelling 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. 5). This resulted in a power parameter $p = 1.3$ with an associated dispersion parameter of $\phi = 20.5$ for the full GAMM. The full GAMM evaluated for swordfish was:

$$CPUE_i = exp(\beta_0 + Year + s_1(Month) + s_2(Long, Lat) + FT + \alpha_V)$$

where *s1*() denotes cyclic cubic smoothing function for *Month*, *s2*() 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, 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 = 28) 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 and 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 2015*). 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* = 7), representative of the high catch quarter and *FT* was fixed to the fishing tactic the produced highest average catch rates (*FT* = 2). The expected yearly mean *CPUEy* 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:

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

and

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

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

References

Figures

Figure 1: Annual effort for the combined South African longline fleets. 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.

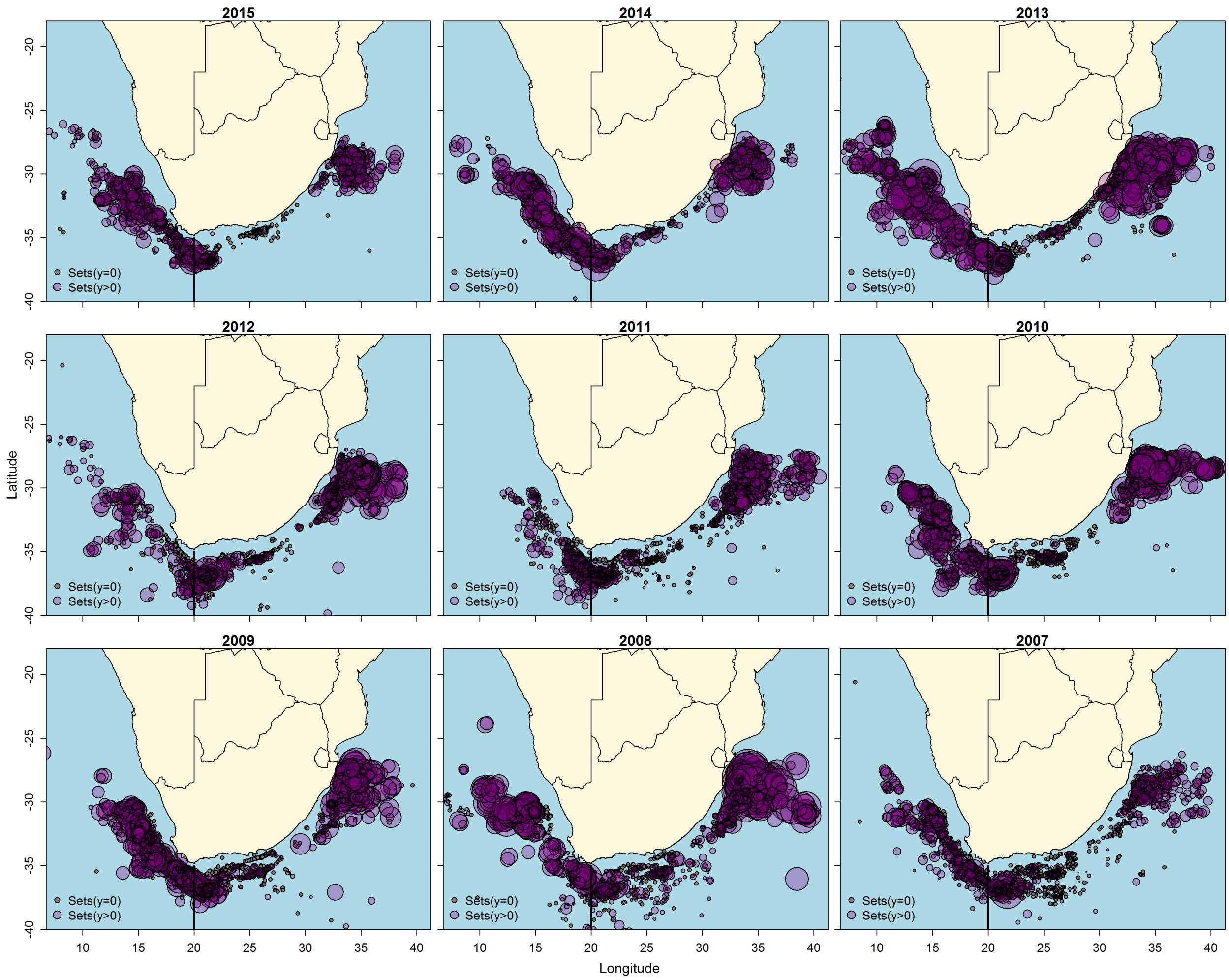

2015
2014
2013
2012
2011
2010
2009
2008
2007
Sets(y=0)
Sets(y>0)
Latitude
Longitude
-20
-25
-30
-35
-40
10
15
20
25
30
35
40

Figure 1:

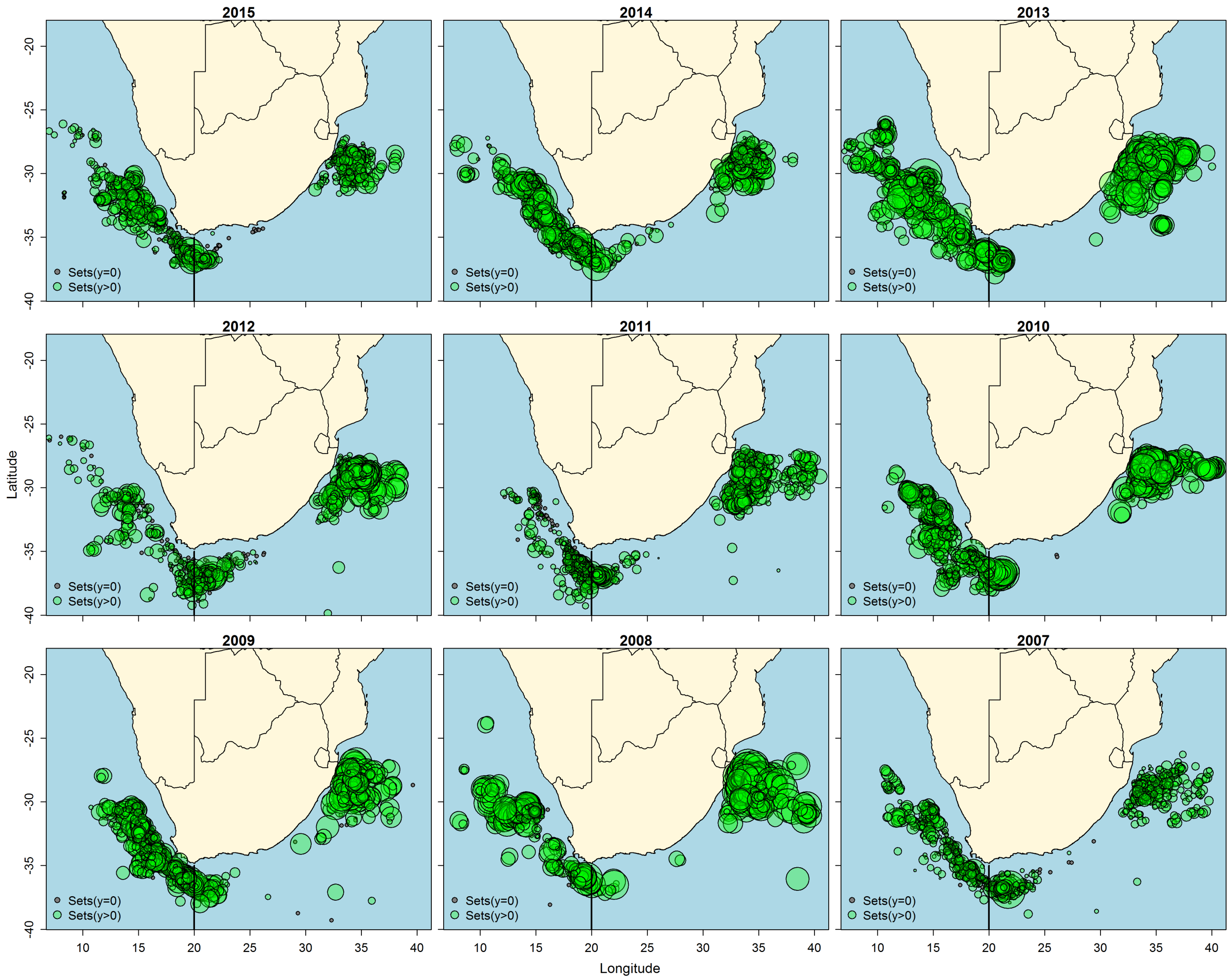


Figure 2:

Figure 2: Annual effort for the South African swordfish directed 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.

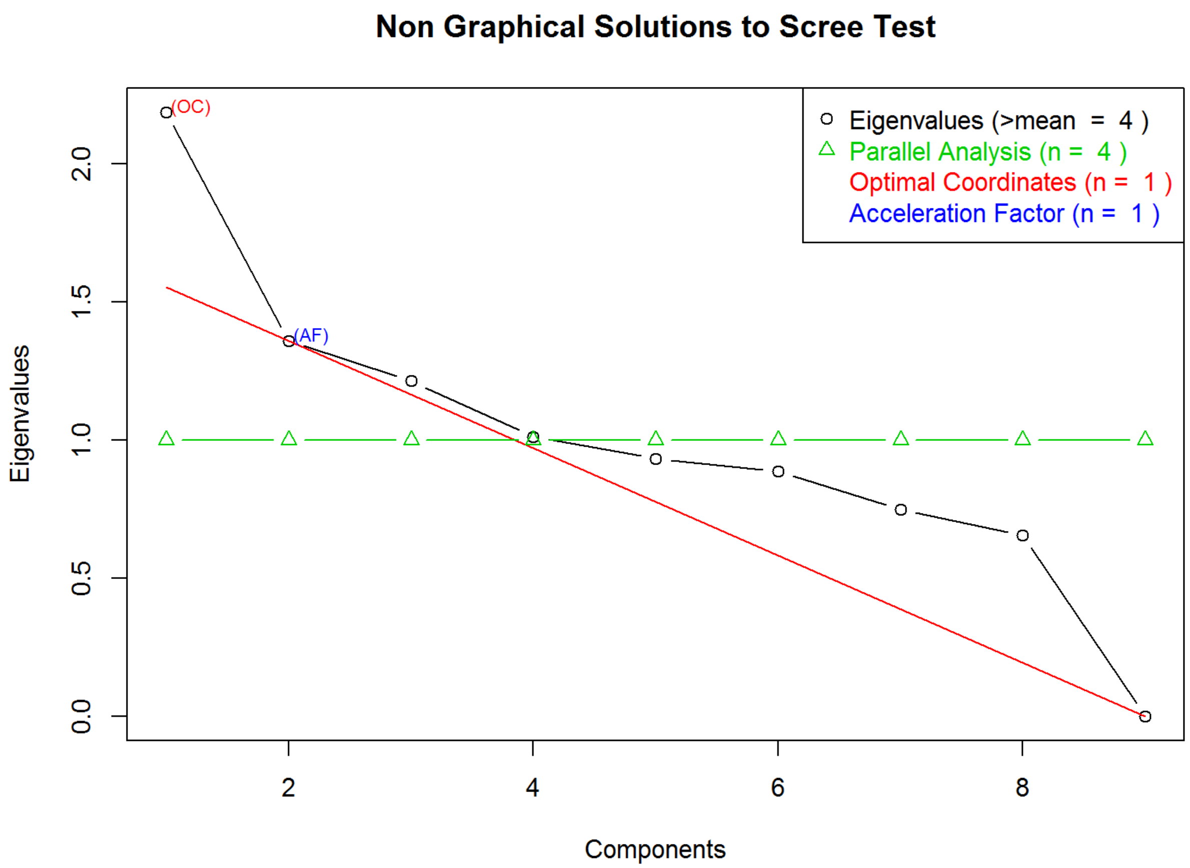


Figure 3:

Figure 3: A non-graphical solution to the Scree test to determine the optimal number of clusters in the multivariate analysis to assess the influence of fishing tactic on CPUE estimation.

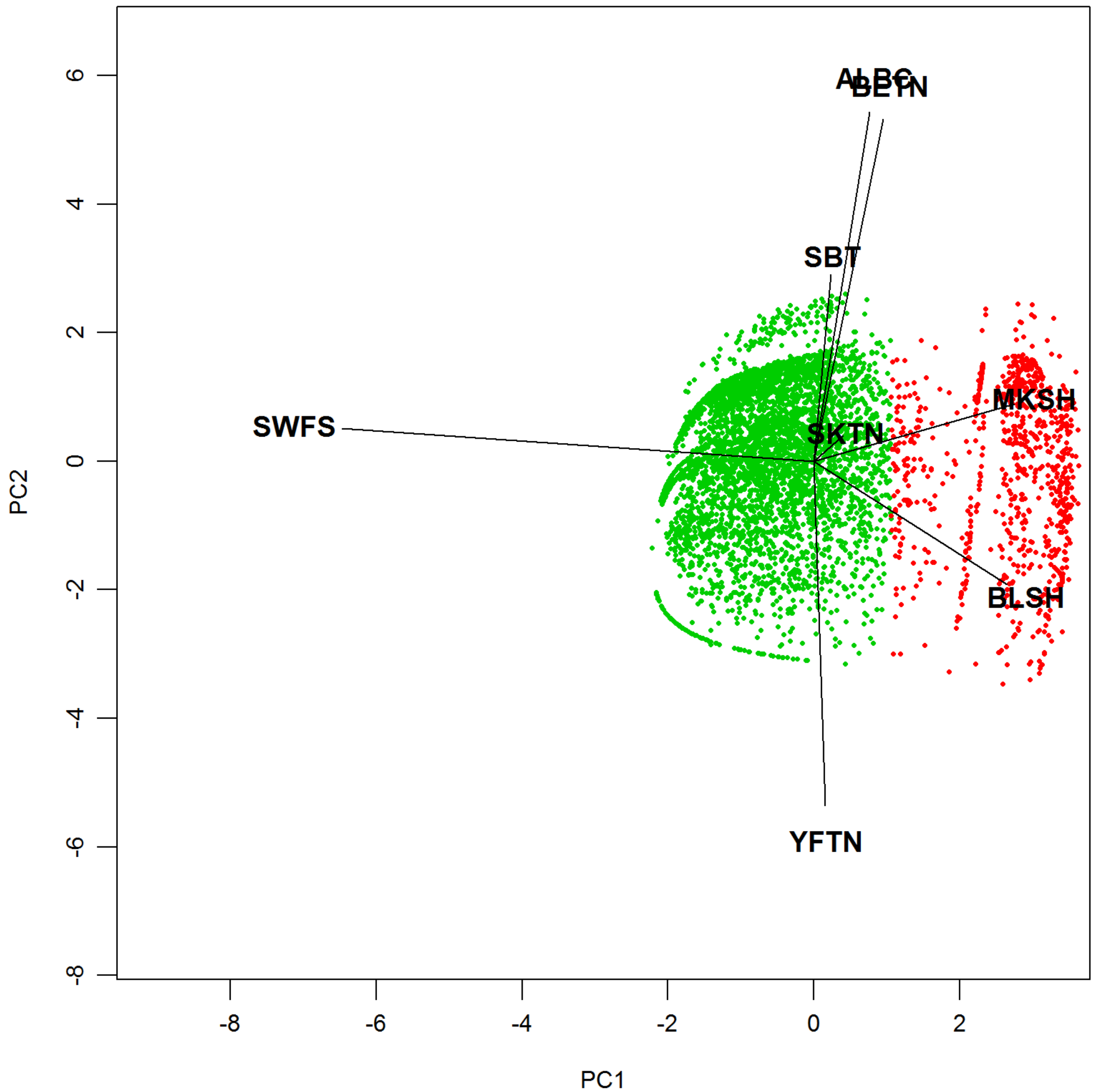


Figure 4:

Figure 4: 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. FT 1: Cluster one dominated is by shark (blue and shortfin mako) catches. FT 2: Cluster two is dominated by swordfish and tuna catches.

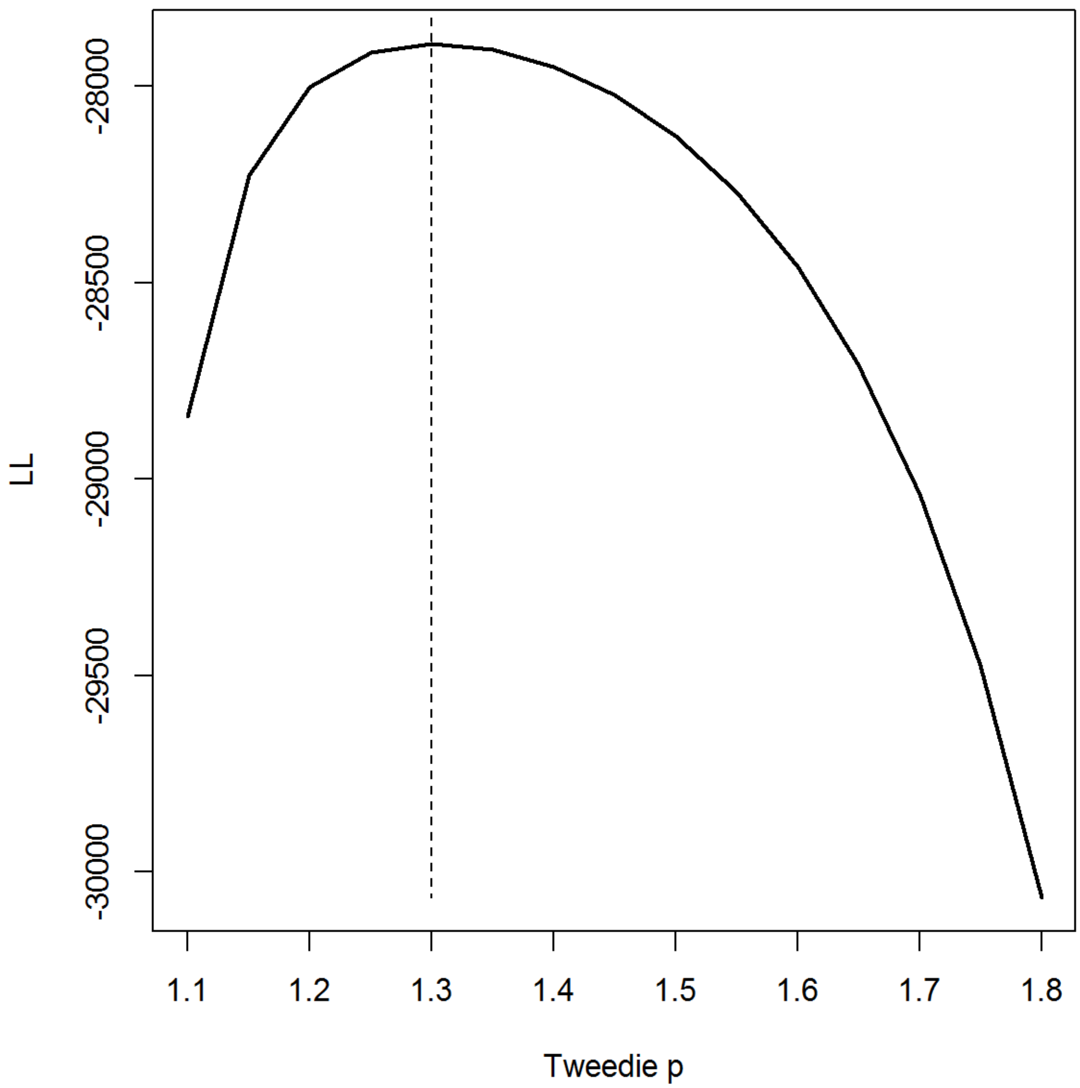


Figure 5:

Figure: 5 Log-likelihood profile for over the grid of power parameters values ($1 < p < 2$) of the tweedie distribution. The vertical dashed line denote the optimized p used in the final standardization GAMM.

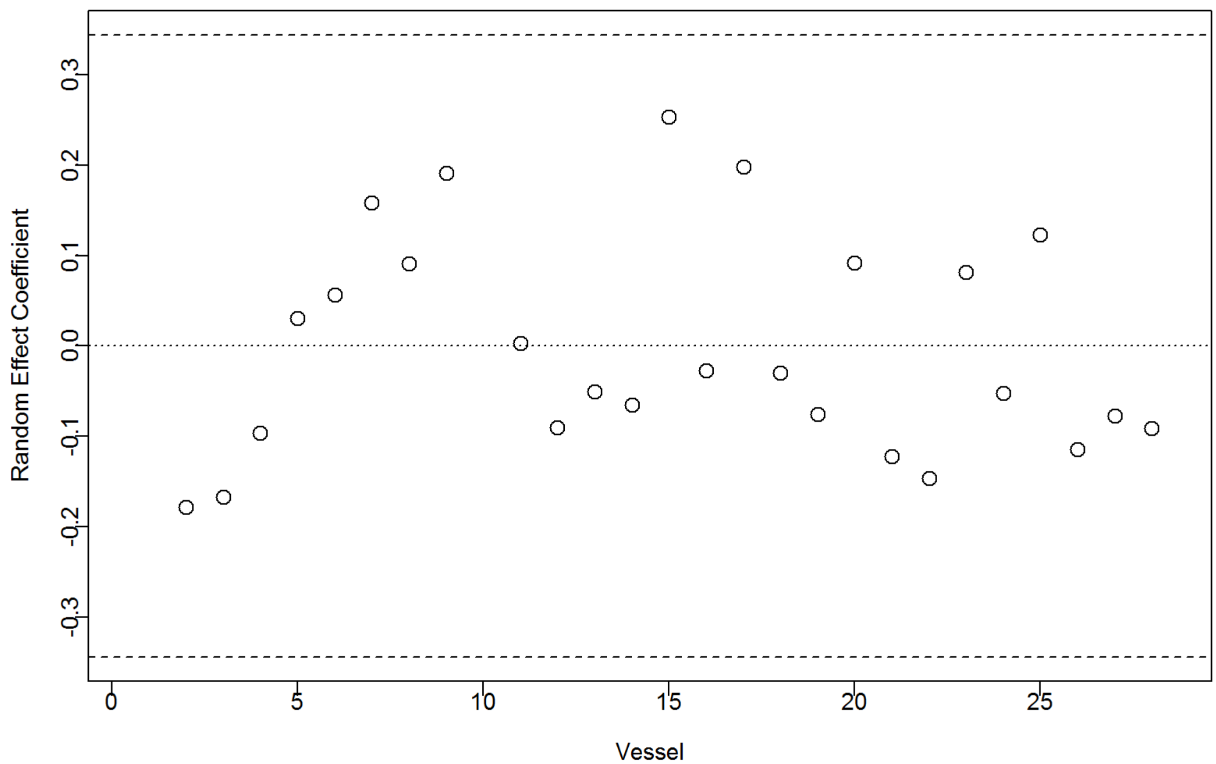
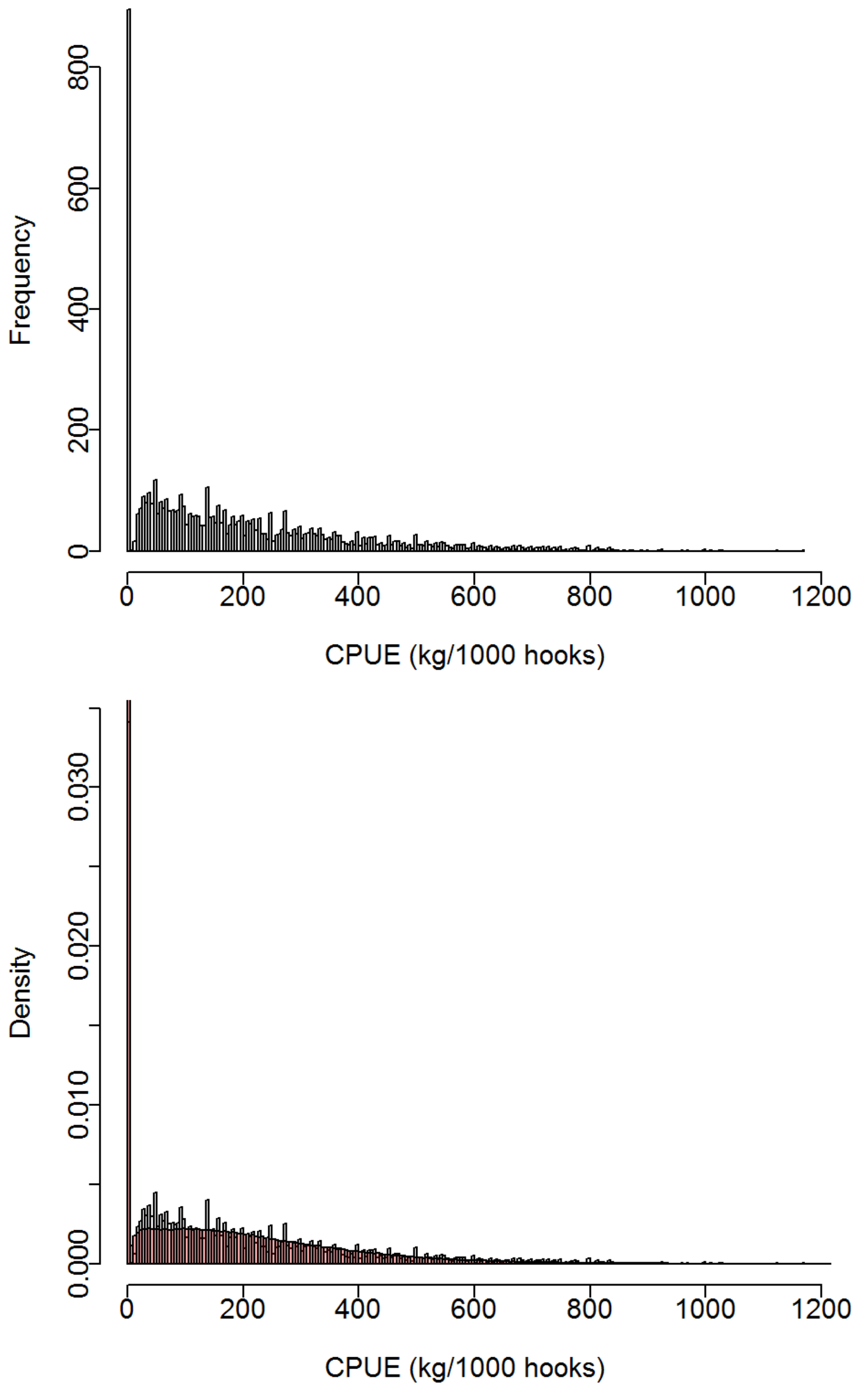


Figure 6:

Figure 6: Random effects coeffients (dots) illustrating the deviation from the mean of zero across the 28 vessels retained for the analysis. Dashed lines denote the 95% confidence interval of the mean ??v.

Figure 7: CPUE frequency, and density, distributions for the South African swordfish directed longline fishery. These distributions support the use of the Tweedie distribution form in the GAMM.



Frequency
0
200
400
600
800
0
200
400
600
800
1000
1200
CPUE (kg/1000 hooks)
Density
0.000
0.010
0.020
0.030
0
200
400
600
800
1000
1200
CPUE (kg/1000 hooks)

Figure 7:

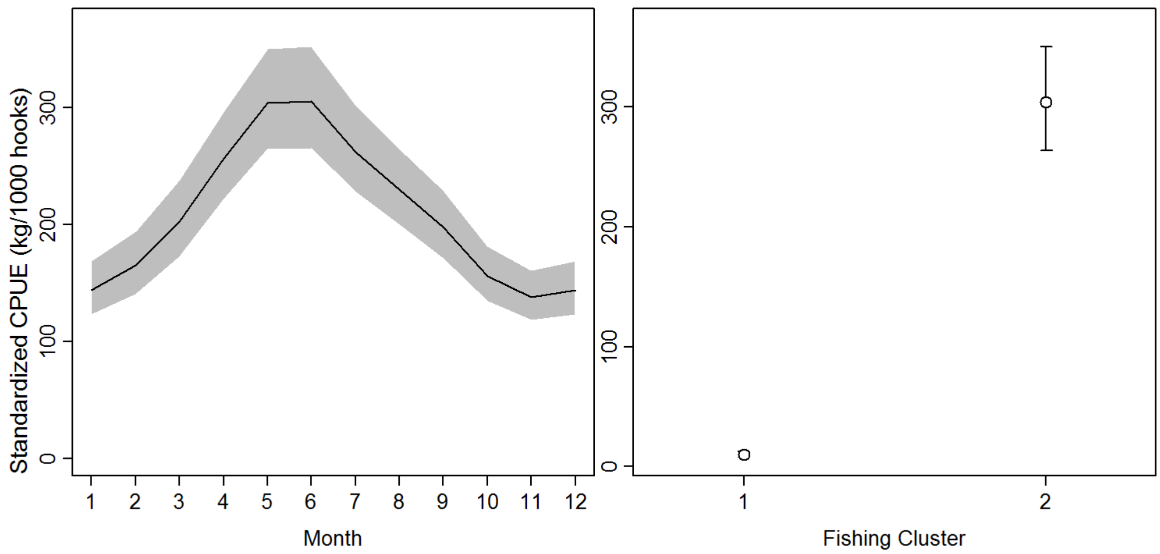


Figure 8:

Figure 8: The influence of the fixed effects *Month* and *Fishing Cluster* on the CPUE of swordfish when modelled using the GAMM applied to the South African swordfish directed longline data.

Figure 9: (a) Standardized CPUE for the swordfish directed longline fishery of South Africa for the time period 2004 to 2015. The 95% confidence intervals for the nominal CPUE are denoted by grey shaded areas and (b) comparison of nominal and the various standardized CPUE models.

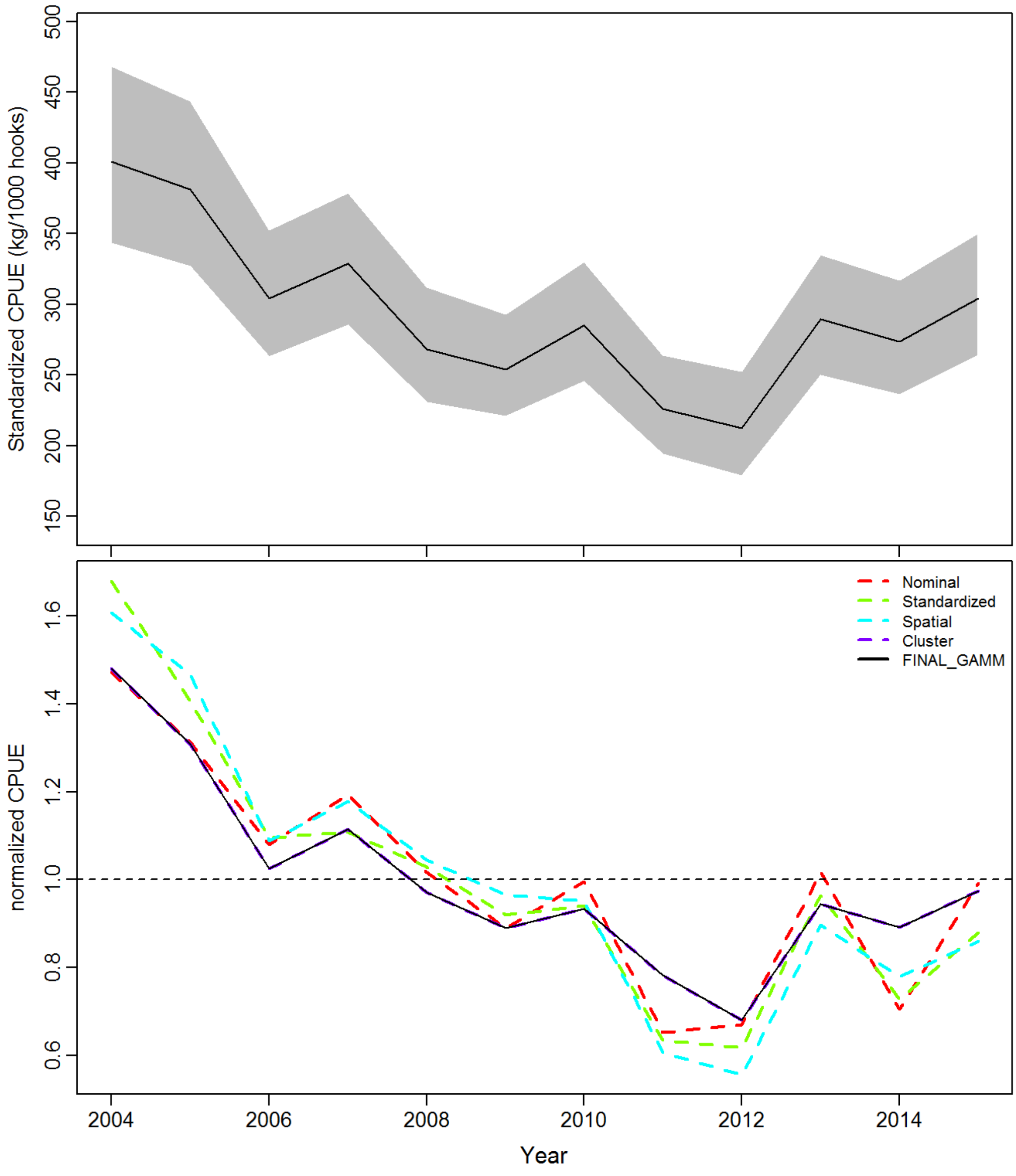

Standardized CPUE (kg/1000 hooks)
150
200
250
300
350
400
450
500
normalized CPUE
0.6
0.8
1.0
1.2
1.4
1.6
Nominal
Standardized
Spatial
Cluster
FINAL_GAMM
2004
2006
2008
2010
2012
2014
Year

Figure 9: