Received: 9 October 2013
IOTC-2013-WPTT15-23

Standarized CPUE for juveniles yellowfin, skipjack and bigeye tuna from the european purse Seine fleet in the Indian ocean from 1981 to 2011

by

M. Soto[1], A. Delgado de Molina[2], E. Chassot[3]

Abstract

In this document three abundance indices are obtained for the juveniles of tropical tunas (yellowfin (≤10Kg), skipjack and bigeye (≤10Kg) of European purse seine fishery in the Indian ocean from 1981 to 2011 using generalized linear models. Catch and effort data come from detailed daily logbooks. Catch rates are modelled using the delta lognormal model. The method estimates a combined cpue of the three species from aggregated catches, and the proportion of catches for each species, so the final individual abundance indices are calculated multiplying both estimators for each species. Explanatory factors used in the analysis are: year, zone, quarter, holding capacity, country and starting date of the vessel. Year is the most explanatory factor of variability in cpue and, depending on the species, the fishing area and the quarter are significant. Vessel characteristics have a significant explanatory effect in observed aggregated catch rates.

Introduction

Since the last two decades, the increasing use of drifting fishing aggregative devices (FADs) by the purse seine fleets operating in the Indian Ocean has changed the length distributions of the tunas tropical landings. In contrast to non-associated school sets which target large fish (mainly yellowfin, *Thunnus albacares*), FADs fishing operations concern skipjack (*Katsuwonnus pelamis*) and juveniles of yellowfin and bigeye tunas (*Thunnus obesus*). With this consideration in mind, the aim of this paper is to develop a standardization procedure of CPUEs for FADs fishing operations. Since, purse seine fishermen may target alternatively associated schools and FADs schools, the presence of a high amount of zero-catch per fishing day may be expected in the data set. As explained in the Method section, in such a situation, delta-lognormal method is an appropriate tool for standardizing CPUEs (Lo et al, 1992, Stefansson, 1996)

Material and Methods

Standarized catch rates of juveniles of yellowfin, skipjack and bigeye were estimated simultaneously for the three species using the generalized linear model assuming a delta-lognormal error distribution (Soto *et al.*, 2009). The analysis has been carried out

[1] Instituto Español de Oceanografía. Corazón de María 8. 28002 Madrid (ESPAÑA). E_mail:maria.soto@md.ieo.es.
[2] Instituto Español de Oceanografía. Centro Oceanográfico de Canarias. Apdo. de Correos 1373. 38080 Santa Cruz de Tenerife. Islas Canarias (ESPAÑA). E_mail: tunidos@ieo.rcanaria.es.
[3]Institut de Recherche pour le D´eveloppement, UMR 212 EME (IRD/Ifremer/Universit´e de Montpellier 2), Avenue Jean Monnet, BP 171, 34 203 S`ete Cedex, FRANCE.

with catch and effort data from logbooks, once the specific composition of catches has been corrected (Anon, 1984, Pallarés y Petit, 2001...) and from detailed fleet data. Catch and effort data are obtained by set, while fleet data contain information about age of vessel, physical characteristics (length, holding capacity, GTR, ..) and vessel history. French, Spanish and NEI fleet data have been analyzed together. In this analysis, the NEI fleet was associated to the Spanish purse seine fleet following results of discriminant analysis (Soto et al., 2002). The period considered goes from 1981 to 2011, years where detailed logbooks are available.

It was considered a minimum threshold of effort by vessel of 120 fishing days per year. This threshold was selected after to analyze the yields as a function of fishing time of vessels and to observe that there was no correlation between both, neither between fleets nor between the whole of the fleets, and also, that the variability, higher for vessels with short fishing periods, was tending to stabilize from this threshold. Later, a selection of vessels operating in the fishery for more than 15 years was done with the intention of analyzing data from vesels that would contribute to obtain trends more representative of real abundance.

Once the selection of representative vessels was done, there were established categories according to the holding capacity, measured in m^3, trying to balance all the categories with a representative number of observations. This characteristic defines well the vessel capacity as the probability of bias and imprecisions are very little. Vessel categories observed are the following:

Category	Holding capacity
1	< 750 m^3
2	750 - 1249 m^3
3	> 1250 m^3

Considering the possible interaction of the fleet and the category of the vessels as in Soto *et al.* (2003) a mixed variable category-country was defined with the following levels:

Level	Country	Harvest capacity
1	France	< 750 m^3
2	France	750 - 1249 m^3
3	France	> 1550 m^3
4	Spain	< 750 m^3
5	Spain	750 - 1249 m^3
6	Spain	> 1550 m^3

Data of catches and effort were restricted to those obtained from FADs, aggregated by logs per day, because the catches of juveniles of the purse seine fleet during the period considered are obtained almost exclusively from logs. Fishing areas selected for juveniles of yellowfin and bigeye and skipjack catches were East Somalia, Maldivas, Chagos and Canal Mozambique, those under fishing on FADs mode.

As it is not possible to allocate effort by set between species, catches were aggregated by day and then by month to avoid the excess of null observations in catches and also in the number of sets; then, a combined nominal monthly CPUE was defined as:

$$CPUE = \frac{YFT_1 + SKJ + BET_1}{nset}$$

where YFT_1 are the catches of juveniles of yellowfin (<10 Kg), *SKJ* are the catches of skipjack, BET_1 are the catches of juveniles of bigeye (<10 Kg) in tons and *nset* is the nominal effort of the European purse seine fleet measured in number of sets by day aggregated by month. For each of the three species a specific nominal *CPUE* was defined as

$$CPUE_{sp} = CPUE \cdot p_{sp}$$

where the specie is $sp = YFT_1, SKJ, BET_1$, and the proportion of catches of each specie over total catches is

$$p_{sp} = \frac{sp}{YFT_1 + SKJ + BET_1 + other}$$

The estandarization procedure used was the generalized linear models (GLM) (McCullagh and Nelder, 1989). The combined *CPUE* was estimated assuming that $CPUE+k$ follows a lognormal distribution based on the observation of the normal QQ-plots and the results of Kolmogorov test, where

$$k = 0.01 \cdot mean\,CPUE$$

The proportion of catches for each species, p_{sp}, was modelled independently from the combined *CPUE* assuming a binomial error distribution.

In Figure 1 we can observe histograms of the independent variables used in the Delta model: (a) the distribution of the combined CPUE in logarithmic scale for the lognormal model; and (b), (c), and (d) with the distribution of proportion of observed zero and positive catches on FADs per day for YFT_1, *SJK* and BET_1, respectively.

The independent factors related with abundance considered were: *year*, *fishing area* and *quarter*. Regarding factors related with the vessels, it was considered a combined variable of category of holding capacity and country (*category-fleet*). The age of vessel used in previous studies of *CPUE* (Soto, 2002) was represented by the *operating date* of the vessel in the analysis. Nevertheless, this factor usually causes interactions with the factor year, is not very significant in the models, and masked the abundance effect.

Three abundance indices were obtained from GLM analysis. By one side, a combined positive *CPUE* was estimated from year LSMeans of the lognormal model. By the other side, estimated proportions of catches were estimated for each species from year LSMeans of the binomial model. The specific index for each species was finally calculated as the product of *year* LSMeans of lognormal model and binomial models. Variance of the indices were calculated using the Delta method (Casella, G., 2002), based on the Taylor development of the function

$$g(\mu, p_{sp}) = \mu \cdot p_{sp},$$

where $sp = YFT_1, SKJ, BET_1$, μ is the estimator of combined *CPUE* from the lognormal model and p_{sp} the estimator of proportion of catches of each specie, assuming that both estimators are independent and there are no covariace terms different from zero.

Analysis and model formulations for the delta model were done using the R statistical software package (R Development Core Team, 2013). In general, model evaluation and diagnosis was carried out through residual analysis (McCullagh and Nelder, 1989).

Diagnostic plots are presented for each delta model component: partial residuals for all components, including partial against *year* for each species, and QQ-plots and histograms of Chi-squared residuals for the lognormal component.

A stepwise regression procedure was used to determine the set of systematic factors that significantly explained the observed variability in each model. A Chi-squared test was used to evaluate the statistical significance of an additional factor (McCullagh and Nelder, 1989). Furthermore, the corresponding percentage of deviance explained by each factor relative to the maximum model was estimated to obtain a profile of the most important explanatory factors in the model. A statistically significant variable (*p-value*<0.05) may, in some instances, be omitted from the model if the amount of variation explained by the variable is small in relation to the complexity that it adds (Stefánsson, 1996). The final models included the *year* plus a selection of other explanatory factors that explained more the 5% of the deviance percentage in the models.

3. Results

Positive catches

For the lognormal component, all the factors included are statistically significant. *Quarter* is the most significant factor to explain the variability observed, even more than factor *year*. It explains about 34% of the deviance, while the *year* factor explains the 29%. The *fishing area* is also an important factor related with the positive total *CPUE*, explaining almost the 18% of the variability in the deviance of the model. Regarding the factors related with the fleet, the combined factor *holding capacity-fleet* explains the 15% of the variability, while the *operating date* only explains the 3.4% of the deviance in the model, unless it is statistically significant.

Proportion of positive catches

In the binomial models for the proportion of positives catches of the three species, the factor *year* is the most important, explaining more than 86% in the case of YFT_1 and almost the 80% in the case of *SKJ* and BET_1. *Fishing area* explains the 8% and 13% of the deviances for YFT_1 and *SKJ*, respectively, and *quarter* explains the 17% of the deviance for BET_1. Vessel factors are not significant in the binomial models to explain the proportion of positive catches.

Vessel characteristics are only informative to explain the variability of the combined *CPUE* but not the proportion of positive catches of each species. As in previous analysis (Soto, 2008), the *staritng date* of the vessel does not improve the results to show evidence that the age of vessel influences the *CPUE*. The effect of the vessel on the *CPUE* is more explicit through the category and fleet characteristics than the *starting date*, as they allows to differentiate vessels by a less biased criteria in order to relate them with observed catches.

Selected model

The results of deviance analysis are shown in Table 1. For the lognormal model *quarter*, *year*, *fishing area*, and *category-fleet* are the main explanatory factors and selected for the final model. The stepwise regression includes also the *starting* date of the vessel, but it was decided not to include it in the final model as it explains less than 5% of the deviance in the model. Also, the *starting date* contains non linear effects on the *CPUE* that are difficult to quantify making catchability non constant: old vessels have improved technology and also have more trained crew than other new vessels (Ref.). For the proportion of catches of YFT_1 and *SKJ*, *year*, and *fishing area* are the main explanatory factors selected in the final model; for the proportion of catches of BET1, *year* and *quarter*. Furthermore, the selected factors with the stepwise regression explain more than 5% of the deviance, so they are the same as the factors selected in the deviance analysis.

CPUE

Observed and standarized scaled cpue series by specie are shown in Figure 5. The juveniles of *YFT* and *BET* series have similar patterns and nominal values are within the confidence intervals of the standarized ones. There are no clear trends during the period considered in both series, and nominal *CPUEs* are very similar to the standardized *CPUE*s. The *SKJ* shows a different behaviour, and it can be observed different time trends since the beginning of the fishery. The nominal *CPUE* is above the estandarize one for almost all the years, and outside the confidence intervals in some cases. The whole period (1982-2011) can be divided in two similar cycles, before and after 1997, where the standardized *CPUE* reaches its minimum value, almost identical than the most recent value considered, 2011. These cycles begin with an increasing average trend in the *CPUE* of 6/7 years and are followed by a decreasing average trend of 10 years.

The lowest standardized *CPUE* values are 1991 for YFT_1, 1997 for *SKJ* and 1993 for BET_1. The three series shows a decreasing trend for the last year considered, 2011.

The three series have been scaled to their maximum value in order to allow patterns in the series to be more easily seen and compared. Figure 4 shows the scaled *CPUE* for the three species together and in Table 3 the corresponding CVs for the nominal and standarized scaled cpue series for each species are showed. It can be seen that variability of standarized scaled series is sensibly lower than nominal series for all the species. Fitting diagnoses are show in Figure 2 for lognormal model and Figure 6 for the binomial models. The residuals follows a relatively linear expected pattern for aggregated catches in the QQ-plot (Figure 2b) and partial residuals of single factors in the lognormal model shows the variability that can be explained by each single factor in the model (Figure 2c). The residuals plot for the proportion of catches shows no trend for the three species (Figure 3).

Comparing the three series of tropical tunas in the Indian Ocean it can be seen that skipjack has been decreasing until 2001 more clearly than yellowfin and bigeye.

Table 4 shows relative values between cpue in the last year, 2011, compared with the first year, 1982, $CPUE_{2011}/CPUE_{1982}$, and also with the average standardized cpue value in the complete period, $CPUE_{2011}/CPUE_{1982\text{-}2011}$. The first one is a measure of the actual state of the *CPUE* with respect to the beginning of the fishery. The second one gives an idea of the state of the actual *CPUE* regarding the whole historic period of the fishery. *YFT*1 cpue is clearly higher in 2011 than in the starting date and than the average, just the opposite of bigeye and skipjack, that show a lower values in 2011 than in 1982 and than the average.

4. Discussion

The delta method has been widely used to construct abundance indices for tuna species. In this study, the delta approach has provided simultaneously three indices for juveniles tropical tunas. The CVs of the indices show the higher variability in the bigeye index.

No strong trends appear in the series of standarized cpues, but it seems that juveniles of bigeye and yellowfin are more similar, with an initial increasing period from 1982 to 1988, also present in the skipjack serie at the beginning of the fishery. Differences between skipjack and the other two species appear from nineteens, where the development of FADs and major changes in technology took place. Also, almost 70% of the catches on FADs are skipjack, and trends for this specie should be more clear for than for the juveniles of yellowfin and bigeye (Soto, 2008). From 1988, where yellowfin and bigeye reach their maximum, there is a decade of decreasing followed by a recovery to come back to similar maximum levels in 1997. From 1988 to 1997, skipjack shows an average decreasing, until a minimum value; then the pattern follows a similar cycle increasing until a maximum in 2001 and decreasing until 2011. The three indices decrease in the last year, 2011.

The source of variability that comes from the fleet is represented by the factors *harvest capacity-fleet* and *starting date*. The factor *harvest capacity-fleet* represents the effect of vessel class and it is only significant to explain the variability of aggregated catch and not for the proportion of individual catches, i.e. there is no evidence of differences between proportions of individual catches between vessel classes. Also, the *starting date* of the vessel has been removed from the final model because the proportion of explained variability of global catch rates is very little (3.4%) and it is not statistically significant for the proportion of individual catches.

It appears the effect of the vessel is independent of the proportions of catches of specie and these binomial variables are only related with the abundance factors. Fleet factors have been removed from the binomial final models of proportion as they are not statistically significant in the binomial models.

In general, the standarization procedure showed that vessel characteristics (country, harvest capacity and age of vessel) have a relative minor explanatory effect on the catch rate of juveniles of tropical tuna in the purse seine fishery.

The goal of the standarization procedure is to eliminate the annual variability in the data that is not attributable to the changes in abundance (Maunder y Punt, 2004). This result is in part achieved as it can be seen in Table 3, where the CVs of nominal cpues are higher than the standarized ones.

References

CASELLA, G., Berger, R. L. (2002), Statistical Inference, 2nd ed.

MAUNDER, M. N. and Punt, A. E., 2004. Standarizing catch and effort data : a review of recent approaches. Fisheries Research 70, pp : 141-159.

McCULLAGH, P. and Nelder, J.A., 1989. Generalized linear models, second ed. Ed. Chapman & Hall.

PALLARÉS, P., Ch. Petit. 1998. Tropical tunas: New sampling and data processing strategy for estimating the composition of catches by species and sizes. Col. Vol. Sci. Pap. ICCAT, 48(2): 230-246 (1998).

R Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

SOTO, M., Pallarés, P., Gaertner, D., Delgado de Molina, A. , Fonteneau, A, Ariz, J., 2002. Standardization of tropical purse seine fishing effort by generalized linear model (GLM). IOTC 2002-WPTT-26.

SOTO, M. , E. Chassot and F. Marsac, 2008. Standarized catch rates for yellowfin (*Thunnus albacares*) and skipjack (*Katsuwonus pelamis*) for the European purse seine fleet of the Indian Ocean, 1984-2007. IOTC 2008-WPPT-26.

SOTO, M., Pallarés, P., Delgado de Molina, A., Gaertner, D. 2009. Standardized CPUE for juvenile yellowfin, skipjack and bigeye tuna from the european purse seine fleet in the Atlantic Ocean from 1991 to 2006. *Collective Volume of Scientific Papers ICCAT* **64**, 4 (2009) 1044-1053.

STEFÁNSSON, G. 1996. análisis of groundfish survey abundante data: combining the GLM and delta approaches. ICES journal of Marine Science, 53:577-588.

Table 1: Deviance table for the lognormal model and the proportion of catches of each species. Explanatory factors are emboldened.

Model formulation	Df	Change in deviance	Residual deviance	p-value	Percentage of total deviance
Positive CPUE					
1	1		4086.6		
Factor					
+**year**	29	171.7	3914.8	<0.001	29.2%
+**quarter**	3	202.4	3712.4	<0.001	34.4%
+**area**	3	106.6	3605.8	<0.001	18.1%
+**harvest capacity-fleet**	4	87.6	3518.2	<0.001	14.9%
+starting date	1	20.5	3497.7	<0.001	3.5%
Proportion of positive YFT1					
1	1		220.2		
Factor					
+**year**	29	78.9	141.4	<0.001	86.5%
+quarter	3	3.8	137.6	0.285	4.1%
+**area**	3	7.4	130.1	0.059	8.1%
+harvest capacity-fleet	4	0.4	129.8	0.986	0.4%
+starting date	12	0.8	129.0	0.999	0.8%
Proportion of positive SKJ					
1	1		567.5		
Factor					
+**year**	29	85.8	481.5	<0.001	79.6%
+quarter	3	2.2	479.3	0.531	2.0%
+**area**	3	14.4	464.9	0.023	13.4%
+harvest capacity-fleet	4	1.1	463.8	0.893	1.0%
+starting date	12	4.3	459.5	0.978	4.0%
Proportion of positive BET1					
1	1		170.2		
Factor					
+**year**	29	47.2	123.0	0.018	77.3%
+**quarter**	3	10.4	112.5	0.015	17.1%
+area	3	2.8	109.8	0.427	4.5%
+harvest capacity-fleet	4	0.2	109.6	0.996	0.3%
+starting date	12	0.5	109.1	1.000	0.8%

Table 2: Relative standarized *CPUE* and standard deviation for YFT_1, *SKJ* and BET_1.

	YFT1		SKJ		BET1	
	CPUE	SD	CPUE	SD	CPUE	SD
1982	0,62	1,49	0,52	0,46	0,57	3,27
1983	0,51	0,51	0,48	0,17	0,58	1,11
1984	0,64	0,29	0,56	0,09	0,70	0,63
1985	0,62	0,24	0,58	0,08	0,64	0,53
1986	0,77	0,27	0,75	0,09	0,78	0,59
1987	0,68	0,21	0,65	0,07	0,70	0,46
1988	1,00	0,29	0,92	0,10	1,00	0,64
1989	0,61	0,16	0,56	0,05	0,63	0,36
1990	0,70	0,19	0,64	0,06	0,73	0,41
1991	0,29	0,20	0,66	0,07	0,33	0,43
1992	0,40	0,18	0,67	0,06	0,25	0,39
1993	0,34	0,17	0,66	0,06	0,23	0,38
1994	0,52	0,19	0,72	0,07	0,42	0,41
1995	0,46	0,17	0,55	0,06	0,52	0,36
1996	0,64	0,17	0,55	0,06	0,55	0,37
1997	0,89	0,16	0,43	0,05	0,81	0,34
1998	0,51	0,15	0,50	0,05	0,39	0,34
1999	0,84	0,20	0,62	0,06	0,83	0,43
2000	0,66	0,17	0,83	0,06	0,46	0,37
2001	0,52	0,15	0,70	0,05	0,65	0,32
2002	0,78	0,19	1,00	0,07	0,97	0,43
2003	0,50	0,17	0,71	0,06	0,45	0,38
2004	0,53	0,17	0,70	0,06	0,56	0,38
2005	0,46	0,15	0,67	0,05	0,38	0,33
2006	0,54	0,15	0,64	0,05	0,47	0,33
2007	0,30	0,12	0,42	0,04	0,42	0,26
2008	0,39	0,15	0,47	0,05	0,57	0,33
2009	0,58	0,23	0,62	0,08	0,76	0,50
2010	0,78	0,26	0,62	0,08	0,60	0,55

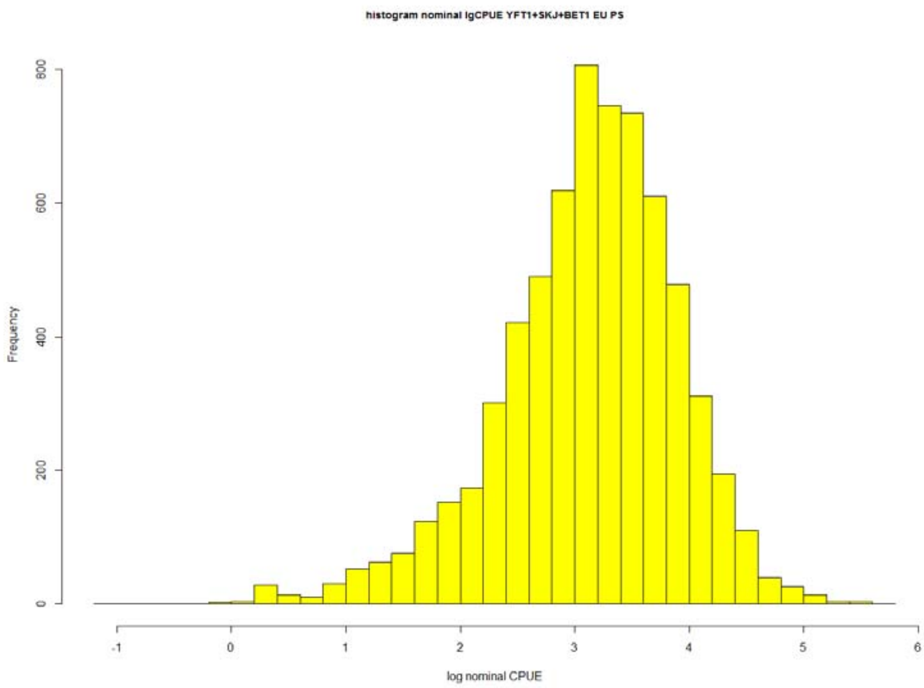
Table 3: Variation coefficients of nominal and standarized CPUEs for YFT_1, *SKJ* and BET_1.

	CV nominal CPUE			CV standarized CPUE		
	YFT1	SKJ	BET1	YFT1	SKJ	BET1
1982	4,58	16,30	2,38	2,38	0,89	5,71
1983	5,12	17,21	2,62	1,00	0,35	1,92
1984	6,70	20,72	2,85	0,45	0,17	0,90
1985	5,03	19,76	2,68	0,39	0,14	0,83
1986	5,52	19,56	2,81	0,35	0,12	0,77
1987	5,60	20,02	2,92	0,31	0,11	0,67
1988	5,39	18,57	2,79	0,29	0,11	0,64
1989	6,22	22,12	2,97	0,27	0,10	0,56
1990	5,51	19,81	2,69	0,27	0,10	0,56
1991	5,50	18,14	3,31	0,68	0,10	1,31
1992	5,02	17,54	2,99	0,45	0,09	1,58
1993	5,09	18,14	5,53	0,50	0,09	1,63
1994	4,82	18,86	2,90	0,36	0,09	0,97
1995	9,55	19,87	3,49	0,36	0,10	0,70
1996	5,71	18,09	2,59	0,26	0,10	0,67
1997	5,42	27,51	2,82	0,18	0,11	0,42
1998	6,00	18,33	6,24	0,30	0,10	0,87
1999	5,19	17,96	3,21	0,24	0,10	0,52
2000	5,27	19,31	3,26	0,26	0,07	0,81
2001	5,83	26,35	2,87	0,28	0,07	0,50
2002	7,14	21,53	2,49	0,25	0,07	0,44
2003	7,87	27,37	7,04	0,35	0,08	0,86
2004	6,68	22,76	3,35	0,33	0,08	0,68
2005	5,79	20,84	3,24	0,33	0,08	0,87
2006	7,11	28,08	3,21	0,28	0,08	0,69
2007	7,15	25,50	2,81	0,39	0,10	0,62
2008	6,33	22,16	2,85	0,39	0,11	0,58
2009	7,38	23,05	3,07	0,39	0,12	0,66
2010	6,77	20,03	3,31	0,33	0,14	0,92
2011	9,70	27,12	4,07	0,31	0,15	0,97

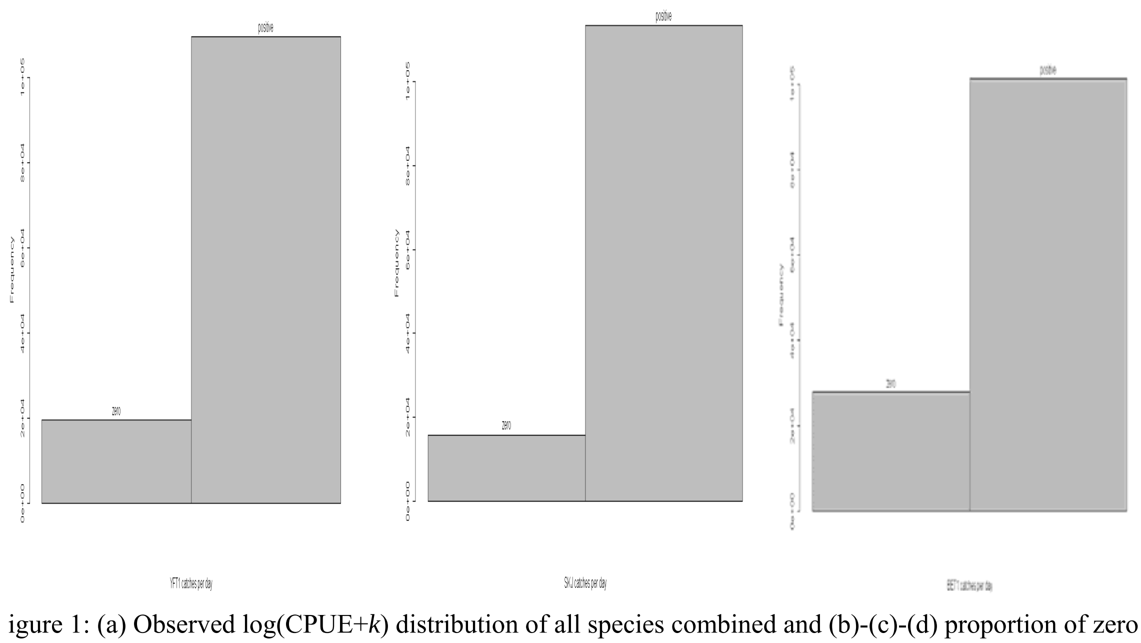
Table 4:Ratios of CPUE in the last year relative to the beginning of the fishery and to the average of the whole period.

	YFT1	SKJ	BET
$CPUE_{2011}/CPUE_{1982}$	1,07	0,89	0,81
$CPUE_{2011}/CPUE_{1982-2011}$	1,13	0,73	0,80

(a)



(b) (c) (d)



igure 1: (a) Observed log(CPUE+*k*) distribution of all species combined and (b)-(c)-(d) proportion of zero and positives purse seine catches per day on FADs for YFT_1, *SKJ* and BET_1.

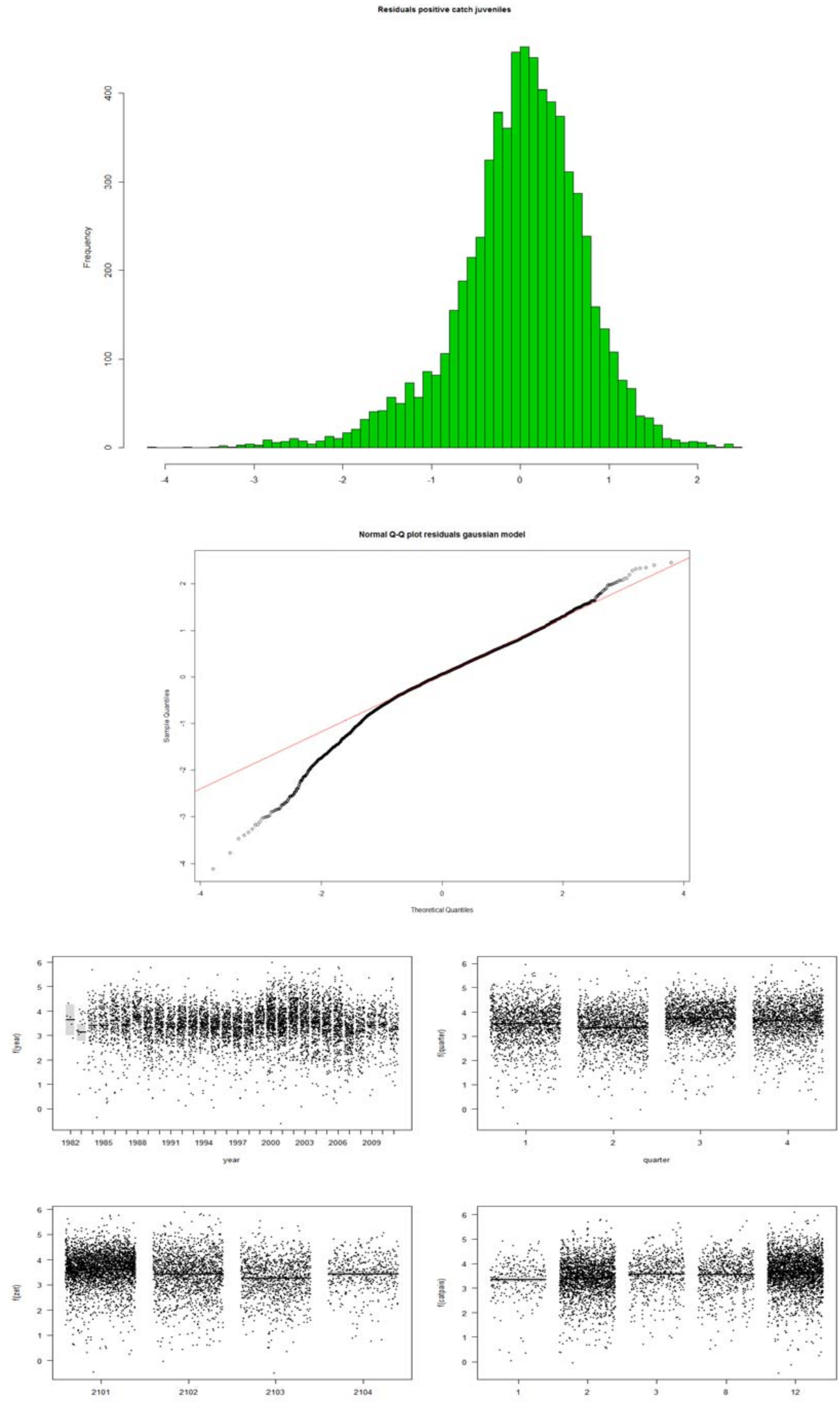


Figure 2:Histogram and Q-Q plot of Chi-squared residuals and partial residuals of lognormal model for the combined cpue.

YFT_1

SKJ

BET_1

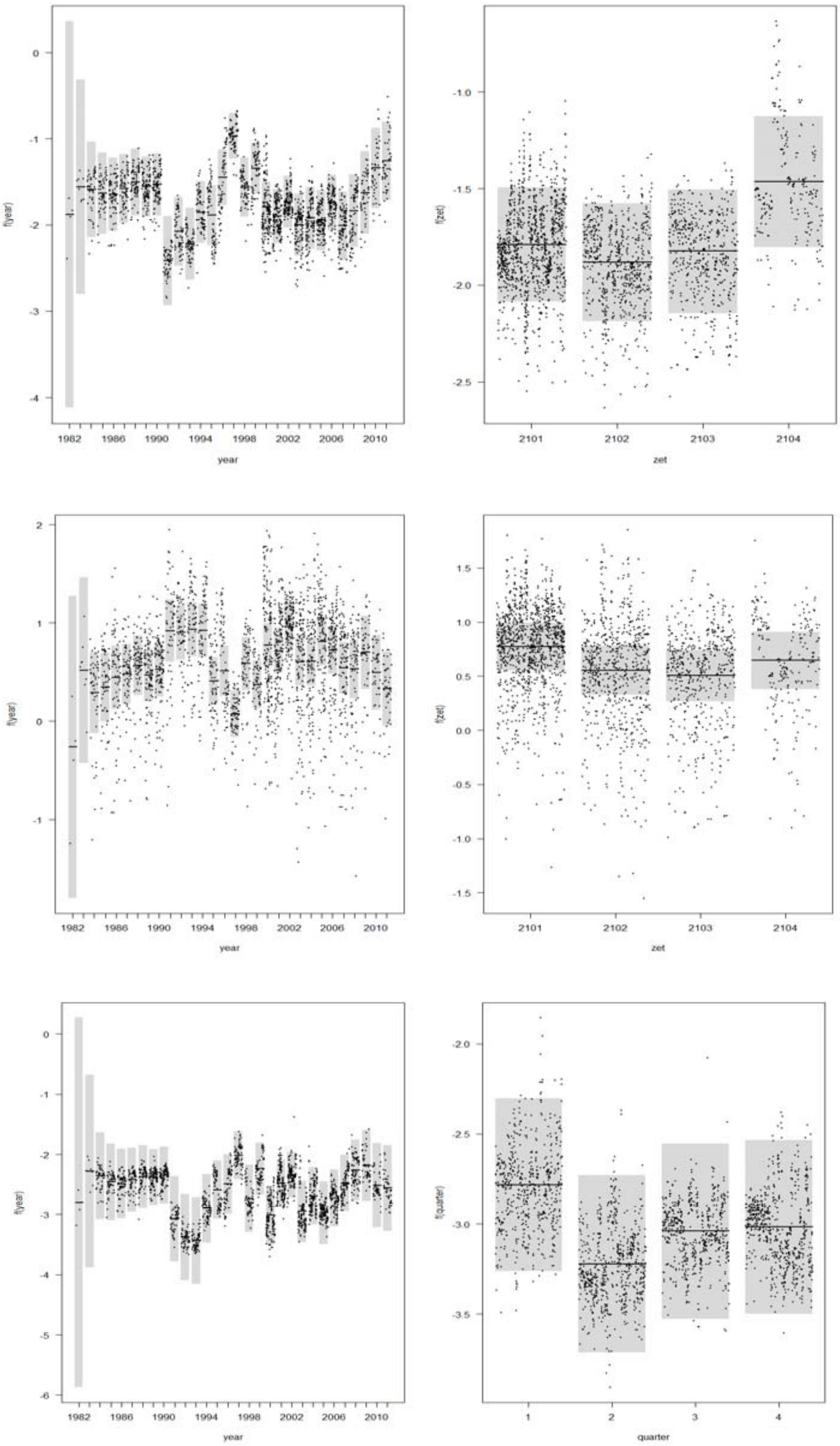


Figure 3: Partial residuals for the final models selected for the propoirtion of catches of yellowfin, skipjack and bigeye.

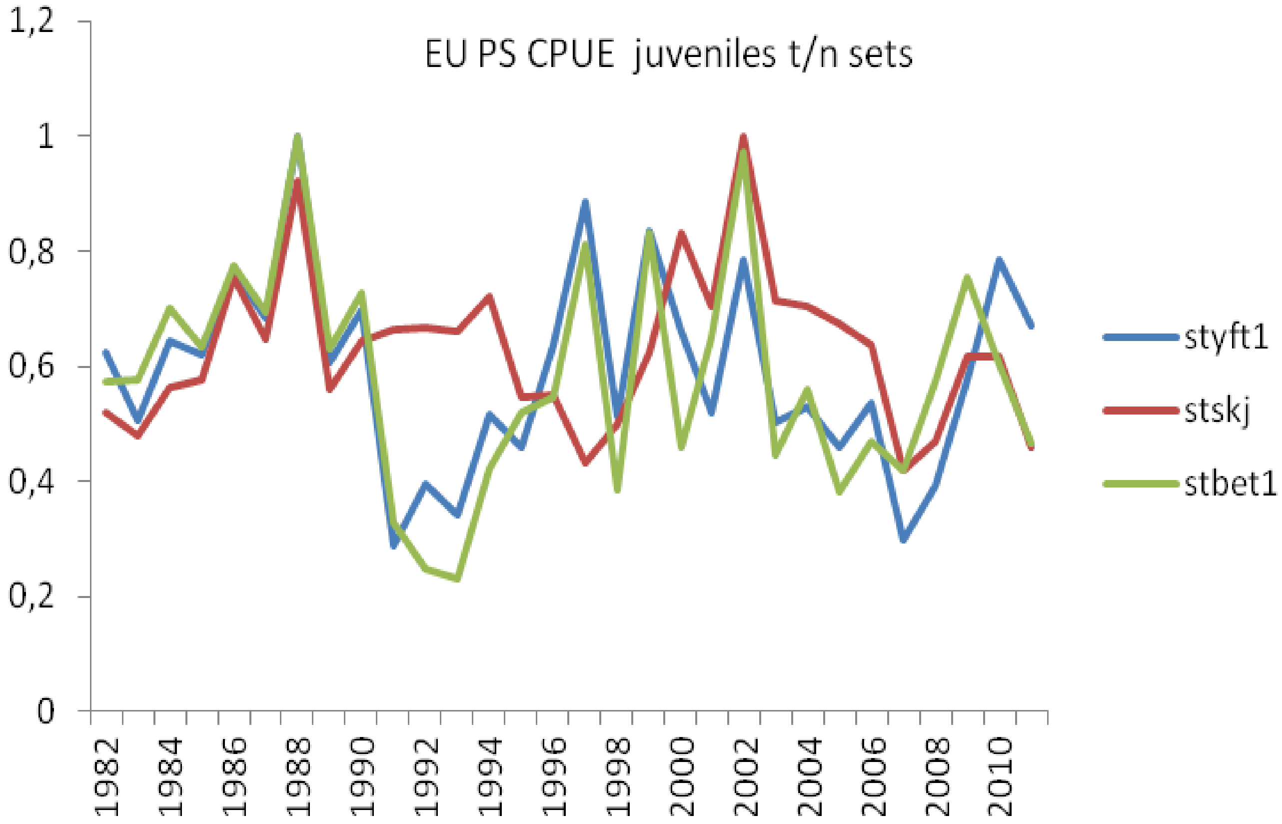


Figure 4: Standarize EU Purse Seine CPUE for juveniles of yellowfin and bigeye and skipjack in the Indian Ocean.

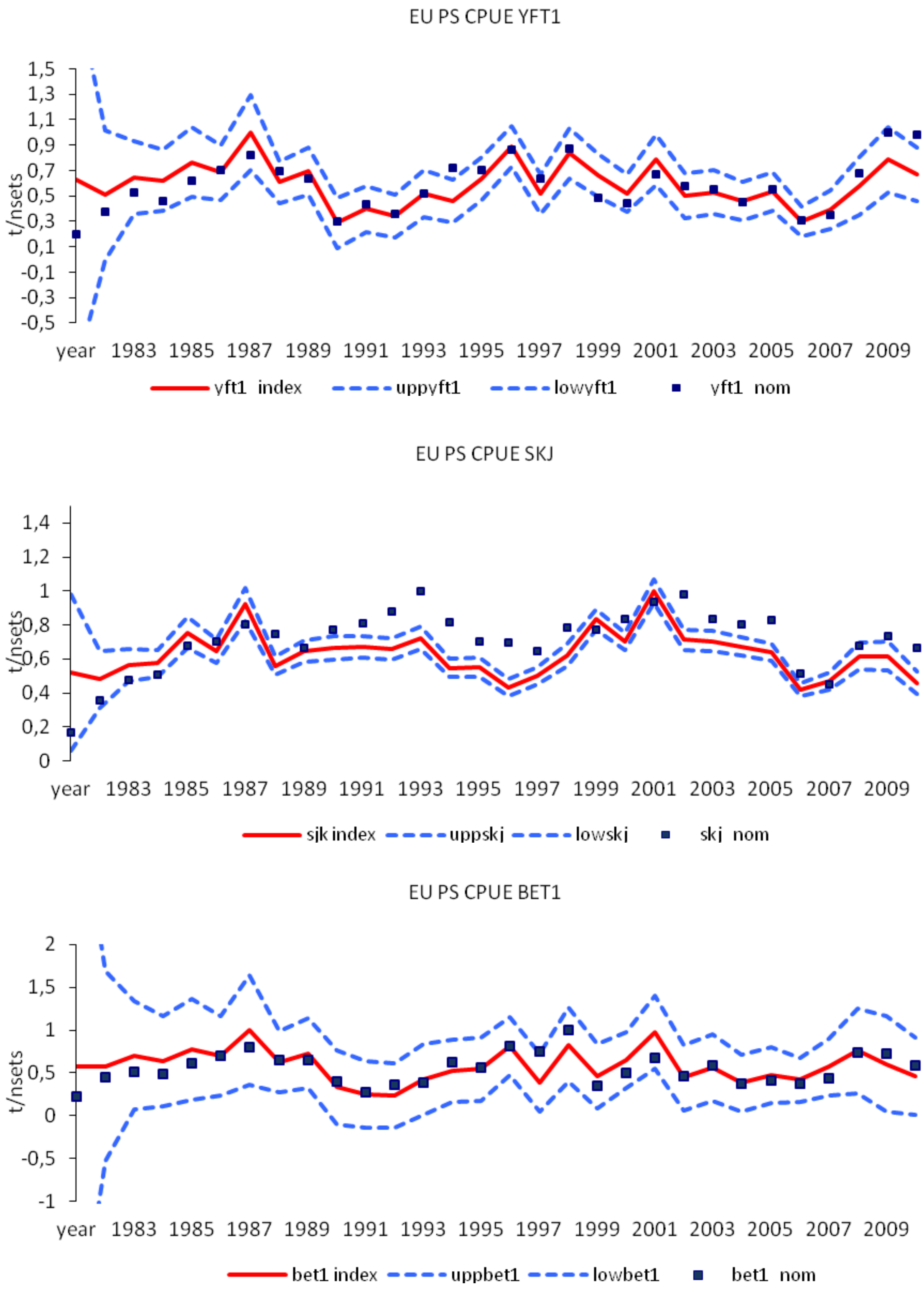


Figure 5: Standarize relative EU Purse Seine CPUE for juveniles of yellowfin and bigeye and skipjack in the Indian Ocean. Confidence intervals and nominal values are also plotted for each specie.