Commercial Longline Vessel Standardized Catch Rates of Tilefish in the US South Atlantic, 1993-2010

Kevin McCarthy

National Marine Fisheries Service, Southeast Fisheries Science Center
Sustainable Fisheries Division, 75 Virginia Beach Drive, Miami, FL, 33149-1099
Kevin.J.McCarthy@noaa.gov
Sustainable Fisheries Division Contribution SFD-2011-002

Introduction

Handline, electric reel (bandit rig), and longline landings and fishing effort of commercial vessels operating in the Gulf of Mexico and U.S. South Atlantic have been reported to the National Marine Fisheries Service (NMFS) through the Coastal Fisheries Logbook Program (CFLP) conducted by the NMFS Southeast Fisheries Science Center. The program collects landings and effort data by fishing trip from vessels that are federally permitted to fish in a number of fisheries managed by the Gulf of Mexico and South Atlantic Fishery Management Councils. The coastal logbook program began in 1992 in the US South Atlantic with the objective of a complete census of coastal fisheries permitted vessel activity, however in Florida a 20% sample of vessels was selected to report. Beginning in 1993, reporting in Florida was increased to include all vessels permitted for federally managed coastal fisheries.

The CFLP available catch per unit effort (CPUE) data were used to construct a standardized abundance index for tilefish. The index was constructed using data reported from commercial bottom longline trips in the US South Atlantic. Tilefish data were sufficient to construct an index of abundance including the years 1993-2010.

Methods

Available Data

For each fishing trip, the coastal logbook database included a unique trip identifier, the landing date, fishing gear deployed, areas fished (Figure 1), number of days at sea, number of crew, gear specific fishing effort, species caught and weight of the landings. Fishing effort data available for longline gear included number of sets and number of hooks fished per set. Multiple areas fished and multiple gears fished may be recorded for a single fishing trip. In such cases, assigning catch and effort to specific locations or gears was not possible; therefore, only trips which reported one area (i.e., subregion, as defined below) and one gear fished were included in these analyses.

Data were further restricted to include only those trips with landings and effort data reported within 45 days of the completion of the trip. Reporting delays beyond 45 days (some reporting delays were longer than one year) likely resulted in less reliable effort data. Landings data, however, may have been reliable even with lengthy reporting delays if trip ticket reports were referenced by the reporting fisher.

Clear outliers in the data, e.g. values falling outside the 99.5 percentile of the data, were excluded from the analyses. These included longline data from trips reporting more than 24 sets per day, more than 3,500 hooks per set, fewer than 25 hooks per set, or longline lengths more than 20 miles or less than 1 mile. Data from trips with reported crews of more than 5 or trips of more than 14 days at sea were also excluded from the analyses. Approximately 67 percent of longline trips were retained for analyses following all data filtering.

Management measures, specifically closed seasons, required that additional data be excluded from the analyses. Closed seasons occurred yearly beginning in 2006 due to quota restrictions and data reported during closed seasons were excluded from the analyses. No minimum size was in effect for the commercial tilefish fishery during the period 1993-2010 and therefore had no effect on the analysis. Trip limit restrictions, however, were in effect beginning in 1994. Targeting of trips may have been affected if a trip limit was met. Coastal logbook data are trip-based, therefore, effort cannot be unambiguously apportioned if targeting changed during a trip. Effects of trip limits were examined by identifying those trips that met or exceeded the trip limit (5,000 pounds gutted weight from 1994-2005; 4,000 pounds from 2006-2010). For those trips that met or exceeded the trip limit, the proportion of tilefish to all other species landed was determined. It was assumed that targeting did not change during a trip if a small proportion of other species were landed from the trip. In such cases, the trip was retained for the analysis.

Tilefish trips were identified using a data subsetting technique (modified from Stephens and MacCall, 2004) intended to restrict the data set to trips with fishing effort in presumptive tilefish habitat. Such an approach was necessary because fishing location was not reported to the CFLP at a spatial scale adequate to identify targeting based upon the habitat where the fishing occurred. The modified Stephens and MacCall method was an objective approach in which a logistic regression was applied to estimate the probability that tilefish could have been encountered given the presence or absence of other species reported from the trip. As a function of the species reported from a trip, a score was assigned to the trip and that score was converted into the probability of observing tilefish. Trips with scores above a critical value were included in the CPUE analysis. That critical value was set at the score that minimized the number of predictions of tilefish occurring when the species was actually absent (false positives) while also minimizing incorrect predictions of tilefish absence when the species was actually present (false negatives). Figure 2 provides species-specific regression coefficients. The magnitude of the coefficients indicates the predictive impact of each species.

Index Development

Longline catch rate was calculated as weight of tilefish per hook fished (hours fished were not consistently reported for longline gear to the CFLP and could not be reliably included in the analysis):

CPUE = pounds of tilefish/(number of sets*number of hooks per set)

Seven factors were considered as possible influences on the proportion of trips that landed tilefish and on the catch rate of tilefish. An additional factor, number of hooks fished, was examined for its affect on the proportion of positive trips. In order to develop a well balanced sample design it was necessary to define categories within some of the factors examined:

Factor	Levels	Value
Year	18	1993-2010
Season	4	Jan-Mar, Apr-Jun, Jul-Sep, Oct-Dec
Subregion	6	Stat areas 2480-2680, 2779-2780, 2878-2880, 2979-3180, 3271-3280, 3374-3677 see Figure 1
Longline length (ll_length)*	4	1-3, 3.1-4.5, 4.6-6.9, 7-20 miles
Days at sea (seadays)*	4	1, 2, 3, 4-5, 6-7, 8-14 days
Crew (crew1)*	2	1-2, 3-5 crew members
Distance between hooks (Hook Distance, hk_dist1)*	4	<26, 26-32, 32.1-43.9, 44+ feet
Hooks fished (hks_fished)*[1]	N/A	Continuous

* Names in parentheses appear in some figures and tables.
[1] Hooks fished was examined only for the proportion positive analyses.

The delta lognormal model approach (Lo et al. 1992) was used to construct a standardized index of abundance. This method combines separate general linear model (GLM) analyses of the proportion of successful trips (trips that landed tilefish) and the catch rates on successful trips to construct a single standardized CPUE index. Parameterization of each model was accomplished using a GLM analysis (GENMOD; Version 8.02 of the SAS System for Windows © 2000. SAS Institute Inc., Cary, NC, USA).

For each GLM analysis of proportion positive trips, a type-3 model was fit, a binomial error distribution was assumed, and the logit link was selected. The response variable was proportion successful trips. During the analysis of catch rates on successful trips, a type-3 model assuming lognormal error distribution was examined. The linking function selected was "normal", and the response variable was log(CPUE). The response variable of longline data was calculated as: log(CPUE)=ln(pounds of tilefish/hooks fished). All 2-way interactions among significant main effects were examined. Higher order interaction terms were not examined.

A forward stepwise regression procedure was used to determine the set of fixed factors and interaction terms that explained a significant portion of the observed variability. Each potential factor was added to the null model sequentially and the resulting reduction in deviance per degree of freedom was examined. The factor that caused the greatest reduction in deviance per degree of freedom was added to the base model if the factor was significant based upon a Chi-Square test ($p<0.05$), and the reduction in deviance per degree of freedom was $\geq$1%. This model then became the base model, and the process was repeated, adding factors and interactions individually until no factor or interaction met the criteria for incorporation into the final model.

Once a set of fixed factors was identified, the influence of the YEAR*FACTOR interactions were examined. YEAR*FACTOR interaction terms were included in the model as random effects. Selection of the final mixed model was based on the Akaike's Information Criterion (AIC), Schwarz's Bayesian Criterion (BIC), and a chi-square test of the difference between the –2 log likelihood statistics between successive model formulations (Littell et al. 1996).

The final delta-lognormal model was fit using a SAS macro, GLIMMIX (Russ Wolfinger, SAS Institute). To facilitate visual comparison, a relative index and relative nominal CPUE series were calculated by dividing each value in the series by the mean cpue of the series.

Results and Discussion

The final models for the binomial on proportion positive trips (PPT) and the lognormal on CPUE of successful trips for each species were:

PPT = Subregion + Hook Distance + Hooks Fished + Year + Longline Length + Subergion*LL Length + Hooks Fished*Year + Subregion*Hooks Fished

LOG(CPUE) = Subregion + Year + Hook Distance + Subregion*Year + Year*Hook Distance

In the proportion positive analysis, the GENMOD algorithm failed to converge when the interactions Year*Longline Length, Year*Hook Distance, Year*Subregion, and Subregion*Hook Distance were tested for inclusion in the final binomial model. That lack of convergence may have been due to inadequately populated data matrices resulting from the relatively small yearly sample sizes. Those interaction terms were dropped from the analysis. The linear regression statistics for fixed effects and the analyses of the mixed model formulations of the final models are summarized in Table 1. Although the interaction terms Hook Distance*Hooks Fished and Hook Distance*Longline Length met the criteria for inclusion in the binomial model (significant chi-square test, reduction in deviance per degree of freedom $\geq$1%) those terms were excluded from the final model. Log likelihood, AIC, and BIC statistics indicated a decrease in fit of the model to the data when those interaction terms were included (Table 1C).

Relative nominal CPUE, number of trips, proportion positive trips, and relative abundance indices are provided in Table 2 for the tilefish model. The delta-lognormal abundance index constructed for tilefish, with 95% confidence intervals, is shown in Figure 3.

Plots of the proportion of positive trips per year, nominal cpue, frequency distributions of the proportion of positive trips, frequency distributions of log(CPUE) for positive catch, cumulative normalized residuals, and plots of chi-square residuals by each main effect for the binomial and lognormal models are shown in Figures 4-7. The observed proportion positive tilefish trips was high; greater than 0.85 for most of the time series and often greater than 0.9 (Figure 4A). Plots of chi-square residuals of the binomial data reveal some observations as outliers. In addition, the plots suggest possible non-random patterns in the residuals of the factors subregion, distance between hooks (hk_dist1), and longline length (ll_length) (Figures 5 B, C, and E). Non-random patterns in residual plots suggest that linear models are not appropriate for such data. Residual plots of the lognormal data, however, do appear to be random (Figure 7). Those diagnostic plots indicate that the fit of the data to the lognormal model was acceptable. The frequency distribution of log(CPUE) data were slightly skewed from the expected normal distribution (Figure 6), however. Those variations from the expected fit of the lognormal data were likely not sufficient to invalidate the analyses.

Tilefish standardized catch rates for longline vessels had no clear trend over the initial 12 years of the time series. During the period 2004-2010, however, yearly mean cpue increased almost fourfold. Confidence intervals around the yearly mean cpue were smaller during the 12 years of relatively constant cpue. During the final six years of the series, however, confidence intervals were larger during the period of cpue increase. Coefficients of variation (CV) were in the range 0.18-0.27 over the entire time series. Additional investigation is necessary to determine whether tilefish population increase during the final six years of the time series or a change in catchability beginning in 2005 best explain the observed increase in cpue.

Literature Cited

Littell, R.C., G.A. Milliken, W.W. Stroup, and R.D Wolfinger. 1996. SAS® System for Mixed Models, Cary NC, USA:SAS Institute Inc., 1996. 663 pp.

Lo, N.C., L.D. Jackson, J.L. Squire. 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Can. J. Fish. Aquat. Sci. 49: 2515-2526.

Stephens, A. and A. McCall. 2004. A multispecies approach to subsetting logbook data for purposes of estimating CPUE. Fisheries Research 70:299-310.

Table 1. Linear regression statistics for the GLM models on proportion positive trips (**A**) and catch rates on positive trips (**B**) for tilefish in the South Atlantic for vessels reporting longline gear landings 1993-2010. Analysis of the mixed model formulations of the proportion positive trip model (**C**) and positive trip model (**D**). The likelihood ratio was used to test the difference of –2 REM log likelihood between two nested models. The final model is indicated with gray shading. See text for factor (effect) definitions.

A.

Type 3 Tests of Fixed Effects

Effect	*Num DF*	*Den DF*	*Chi-Square*	*F Value*	*Pr > ChiSq*	*Pr > F*
year	17	2251	39.98	2.35	0.0013	0.0014
subregion	5	2251	143.06	28.61	<.0001	<.0001
hook distance	3	2251	60.02	20.01	<.0001	<.0001
hooks fished	1	17	11.14	11.14	0.0008	0.0039
longline length	3	2251	19.01	6.34	0.0003	0.0003
*subregion*ll length*	15	2251	28.19	1.88	0.0204	0.0210
*hooks fished*subregion*	5	2251	18.24	3.65	0.0027	0.0027

B.

Type 3 Tests of Fixed Effects

Effect	*Num DF*	*Den DF*	*Chi-Square*	*F Value*	*Pr > ChiSq*	*Pr > F*
year	17	50	116.58	6.86	<.0001	<.0001
subregion	5	71	174.41	34.88	<.0001	<.0001
hook dist	3	50	40.22	13.41	<.0001	<.0001

Table 1 (continued).

C.

Proportion Positive Trips	*-2 REM Log likelihood*	*Akaike's Information Criterion*	*Schwartz's Bayesian Criterion*	*Likelihood Ratio Test*	*P*
year + subregion + hk_dist1 + hks_fished + ll_length + subregion*ll_length	16561.8	16563.8	16569.5	-	-
year + subregion + hk_dist1 + hks_fished + ll_length + subregion*ll_length + year*hks_fished	15791.0	15795.0	15796.8	770.8	<0.0001
year + subregion + hk_dist1 + hks_fished + ll_length + subregion*ll_length + year*hks_fished + subregion*hks_fished	15600.9	15604.9	15606.6	190.1	<0.0001
year + subregion + hk_dist1 + hks_fished + ll_length + subregion*ll_length + year*hks_fished + subregion*hks_fished + hk_dist1*hks_fished	16047.5	16049.5	16050.4	-446.6	N/A

D.

Catch Rates on Positive Trips	*-2 REM Log likelihood*	*Akaike's Information Criterion*	*Schwartz's Bayesian Criterion*	*Likelihood Ratio Test*	*P*
year + subregion + hk_dist1	7013.9	7015.9	7021.8	-	-
year + subregion + hk_dist1 + year*subregion	6963.4	6967.4	6972.4	50.5	<0.0001
year + subregion + hk_dist1 + year*subregion + year*hk_dist1	6945.9	6951.9	6959.6	17.5	<0.0001

Table 2. Longline relative nominal CPUE, number of trips, proportion positive trips, and standardized abundance index for tilefish (1993-2010) in the South Atlantic.

YEAR	Normalized Nominal CPUE	Trips	Proportion Successful Trips	Standardized Index	Lower 95% CI (Index)	Upper 95% CI (Index)	CV (Index)
1993	0.713455	223	0.955157	0.623027	0.418142	0.928305	0.201383
1994	0.644691	311	0.900322	0.623359	0.432052	0.899374	0.184838
1995	0.723717	260	0.888462	0.765415	0.516622	1.134019	0.198467
1996	0.49463	192	0.885417	0.452756	0.284306	0.721013	0.235836
1997	0.444536	208	0.875	0.609743	0.404116	0.919999	0.207862
1998	0.622591	158	0.797468	0.712064	0.447211	1.133771	0.235749
1999	1.229995	198	0.818182	0.636855	0.370461	1.094808	0.275943
2000	0.986446	228	0.877193	0.676573	0.410235	1.115827	0.25412
2001	0.663059	191	0.905759	0.635274	0.414456	0.973742	0.216002
2002	0.458428	137	0.854015	0.519392	0.31795	0.84846	0.249123
2003	0.551598	132	0.863636	0.662369	0.426945	1.027609	0.222257
2004	0.409837	113	0.787611	0.589789	0.354317	0.981749	0.258977
2005	0.985622	66	0.848485	1.060471	0.631062	1.782073	0.263964
2006	0.836367	105	0.733333	1.324642	0.87967	1.994699	0.206838
2007	2.391053	145	0.917241	1.907125	1.232525	2.950955	0.220892
2008	1.649657	105	0.67619	1.996987	1.219977	3.26888	0.250192
2009	1.090824	154	0.441558	1.849201	1.112741	3.073083	0.258114
2010	3.103493	138	0.985507	2.354958	1.600827	3.464351	0.194812

Figure 1. Coastal Logbook defined fishing areas.

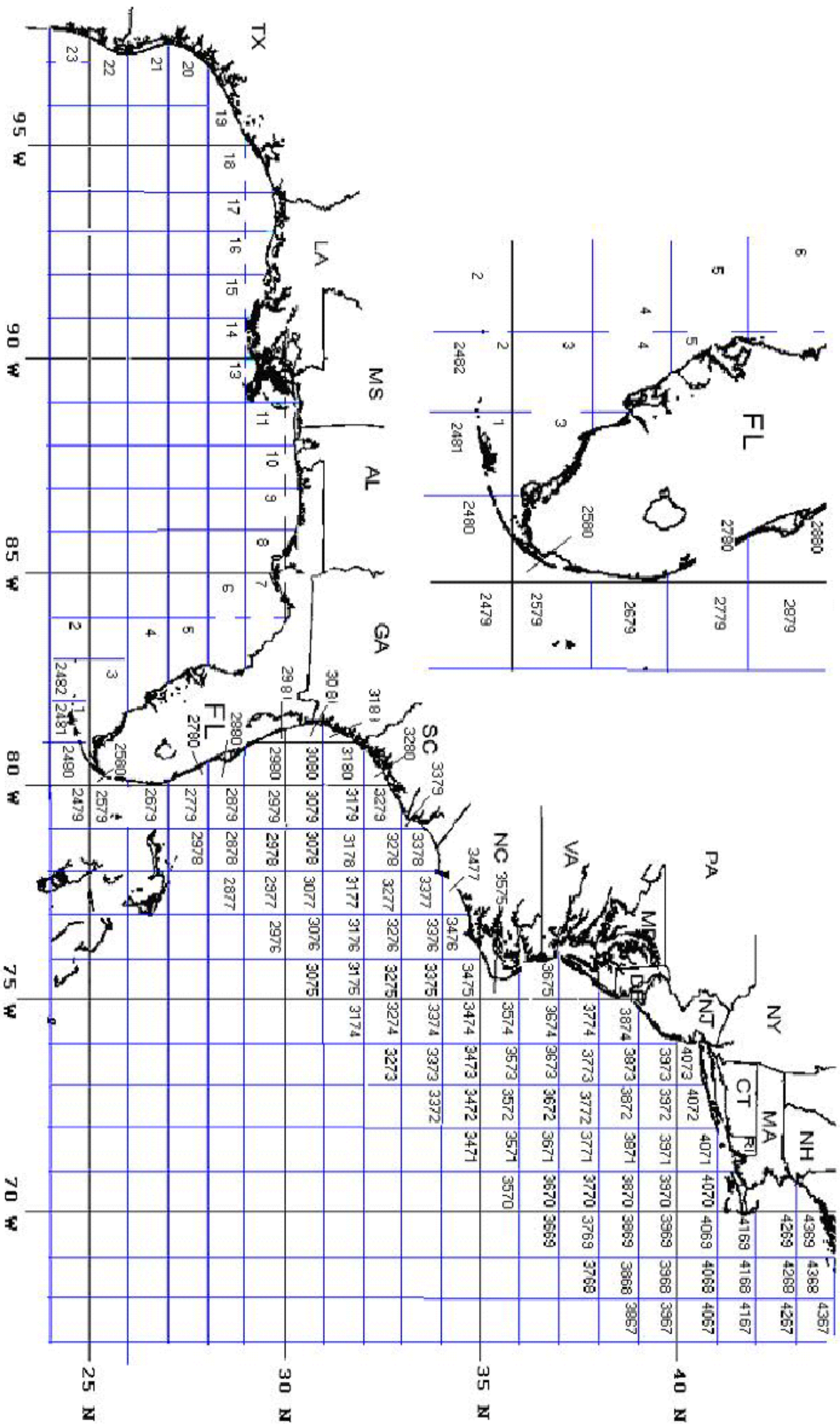


Figure 2. Regression coefficients from the Stephens & MacCall analyses. Positive coefficients signify species that had positive associations with the target species. The magnitude of the coefficients indicates the predictive impact of each species. The value for "non co-occurring" is the regression intercept and denotes the probability a trip was fishing in the target species' habitat, but did not report any of the listed species. Species included were reported on at least one percent of longline trips in the South Atlantic.

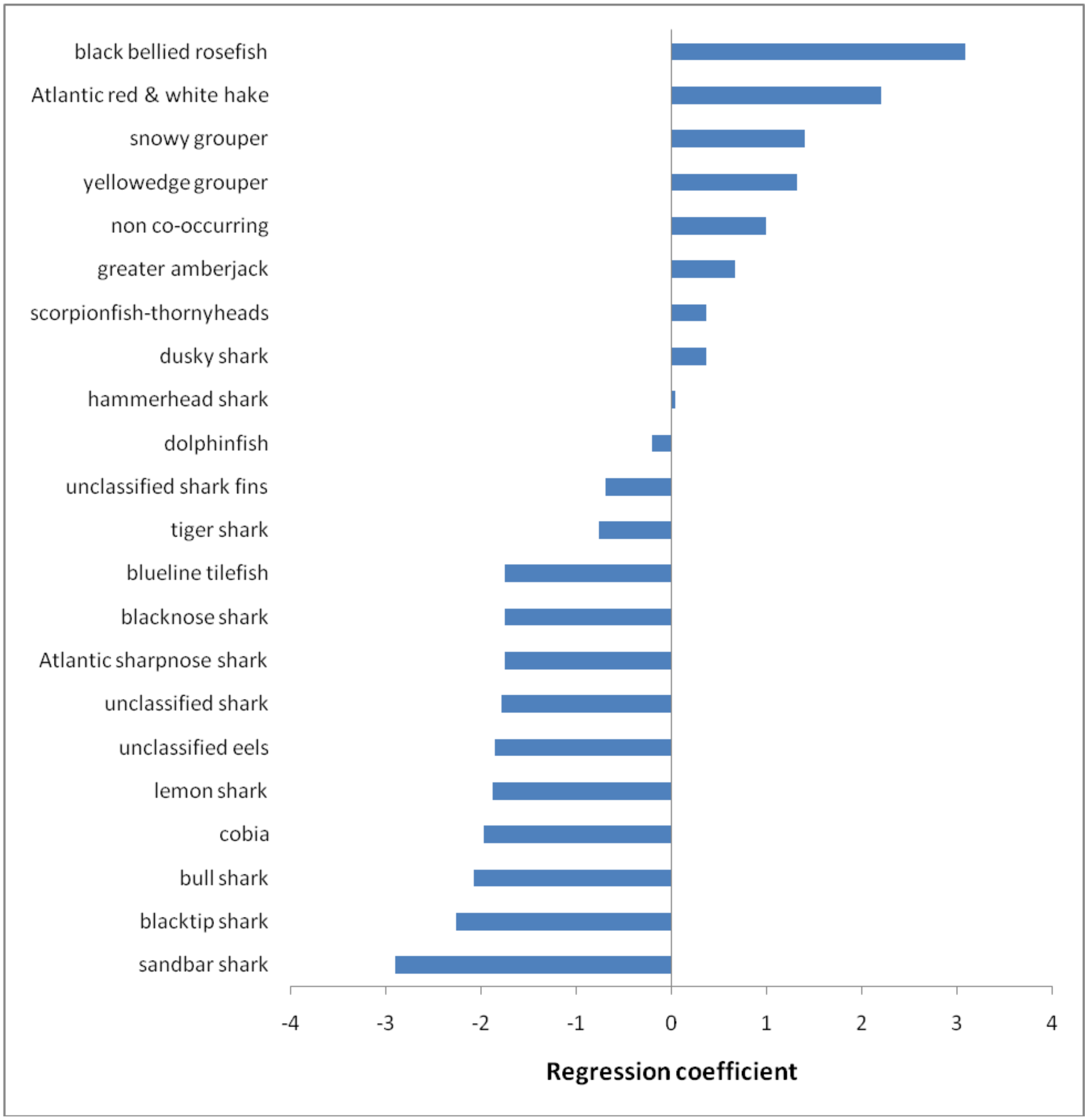


Figure 3. Tilefish nominal CPUE (solid circles), standardized CPUE (open diamonds) and upper and lower 95% confidence limits of the standardized CPUE estimates (dashed lines) for commercial vessels fishing longline gear in the South Atlantic.

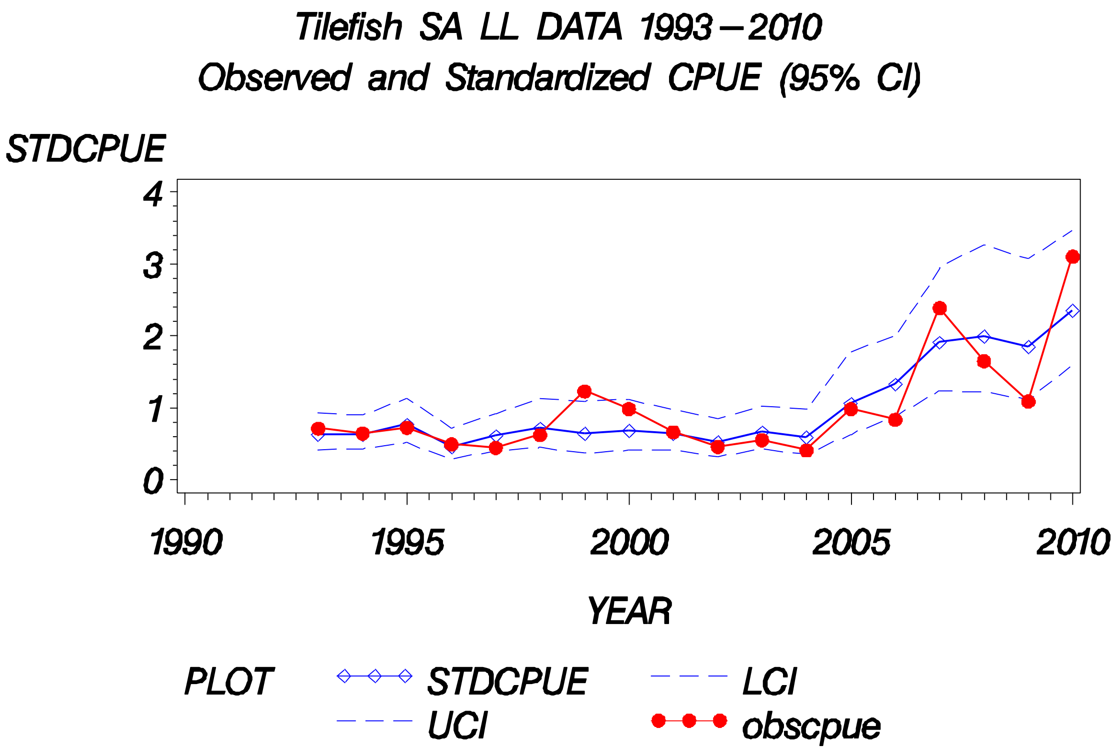
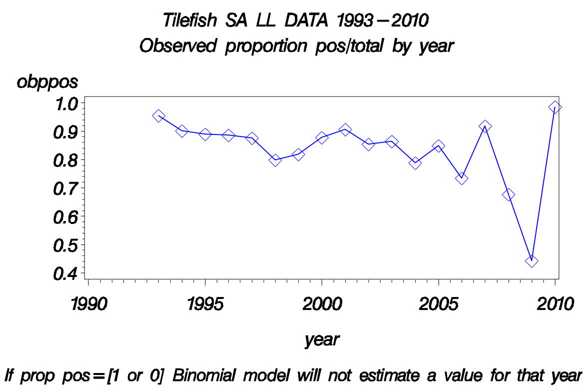


Figure 4. Annual trend in **A.** the proportion of positive trips and **B.** nominal CPUE of the South Atlantic 1993-2010 tilefish commercial longline gear data.

A.



B.

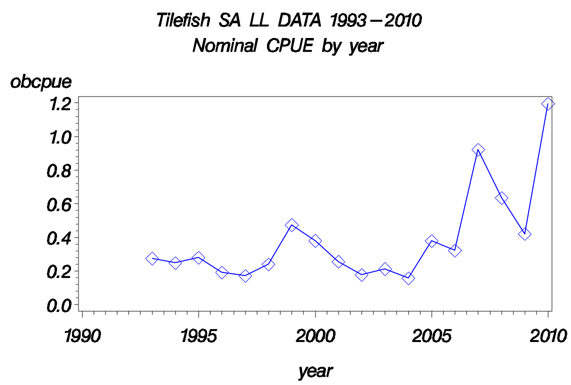
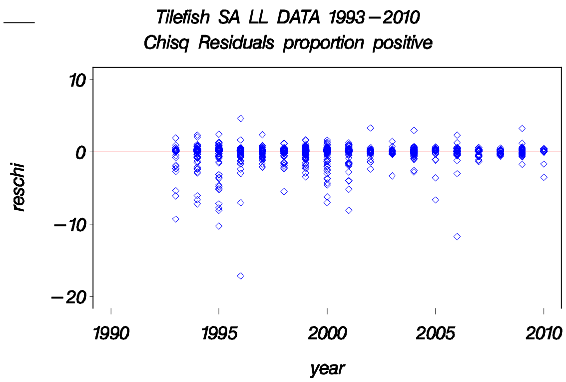
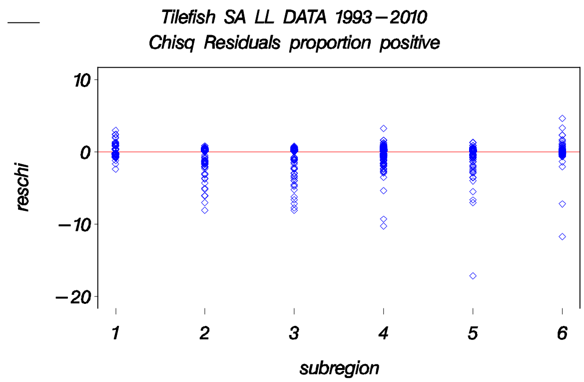


Figure 5. Diagnostic plots for the binomial component of the South Atlantic 1993-2010 tilefish commercial longline gear model: **A**. the Chi-Square residuals by year; **B**. the Chi-Square residuals by subregion; **C**. the Chi-Square residuals by hook distance; **D**. Chi-square residuals by hours fished; and **E**. Chi-square residuals by longline length.

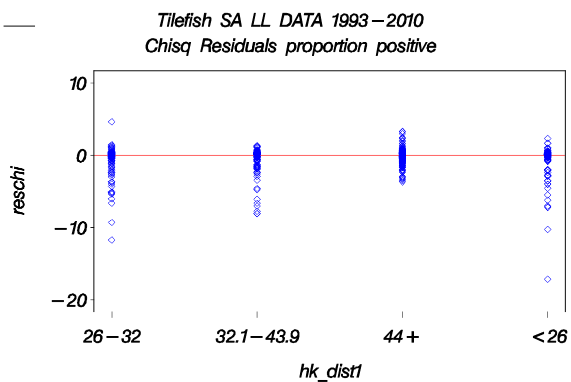
A.



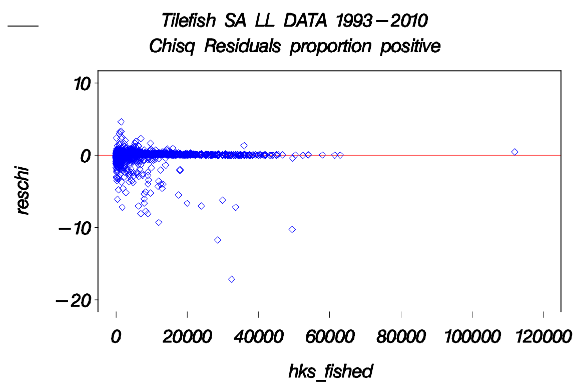
B.



C.



D.



E.

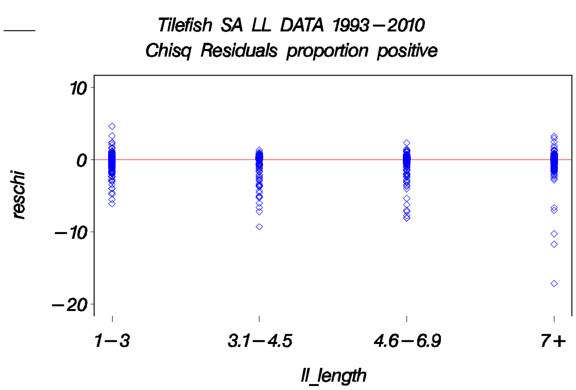
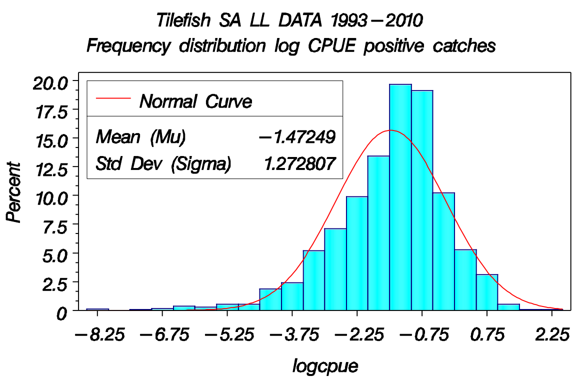


Figure 6. Diagnostic plots for the lognormal component of the South Atlantic 1993-2010 tilefish commercial longline gear model: **A.** the frequency distribution of log(CPUE) on positive trips, **B.** the cumulative normalized residuals (QQ-Plot) from the lognormal model. The red line is the expected normal distribution.

A.



B.

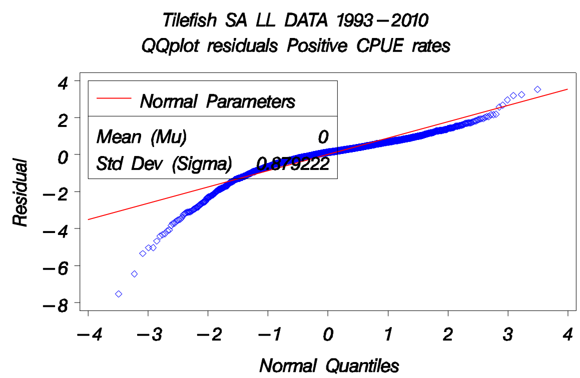
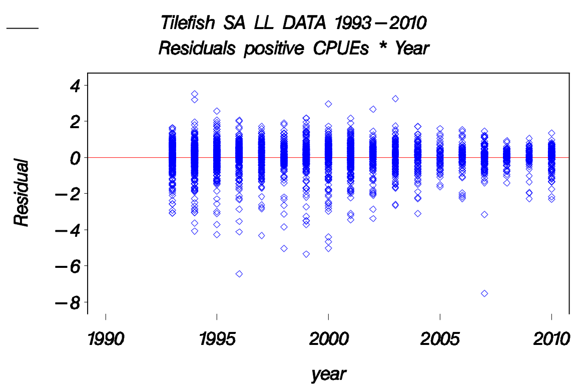
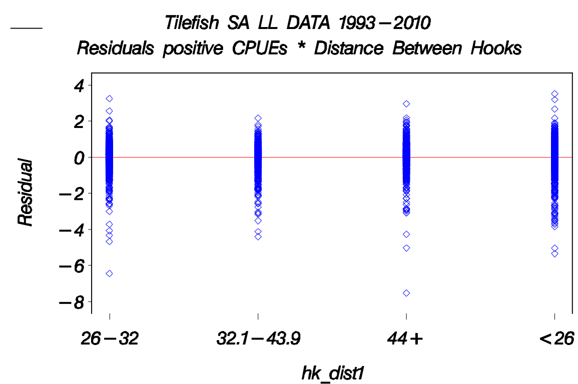


Figure 7. Diagnostic plots for the lognormal component of the South Atlantic 1993-2010 tilefish commercial longline gear model: **A**. the Chi-Square residuals by year; **B**. the Chi-Square residuals by hook distance; and **C**. the Chi-Square residuals by subregion.

A.



B.



C.