Commercial Indices of Abundance for Greater Amberjack in the Gulf of Mexico

Steven Saul

SEDAR33-AW18

10 July 2013

Please cite as:

Saul, S. 2013. Commercial Indices of Abundance for Greater Amberjack in the Gulf of Mexico. SEDAR33-AW18. SEDAR, North Charleston, SC. 17 pp.

Commercial Indices of Abundance for Greater Amberjack in the Gulf of Mexico

Steven Saul
National Marine Fisheries Service, Southeast Fisheries Science Center
Sustainable Fisheries Division, 75 Virginia Beach Drive, Miami, FL 33149
Steve.Saul@noaa.gov

Sustainable Fisheries Division Contribution Number: SFD-2013-018

Introduction

The National Marine Fisheries Service (NMFS) collects information on catch and fishing effort from the commercial fishing industry in the Southeastern Region through the Southeast Fisheries Science Center's Coastal Logbook Program. Individuals who carry commercial federal fishing permits are required to provide information on their landings and fishing effort for each trip that they take. The program began in 1990 however a complete census of the commercial fishing activity in the Southeastern Region was not obtained until 1993 when all federally permitted vessels were required to report. Using the catch and effort data provided by this program, two indices of abundance were calculated for greater amberjack (*Seriola dumerili*) in the Gulf of Mexico: one for the handline fishery and one for the longline fishery. Indices of abundance were standardized using the Delta lognormal approach and indices were created for years 1993 through 2012. The same methodologies used during the last assessment to standardize these indices were implemented during this study and past and present indices were compared for continuity.

Methodology

For each fishing trip, the coastal logbook database includes a unique trip identifier, the landing date, fishing gear deployed, areas fished, the number of days at sea, the number of crew, gear specific fishing effort, species caught and weight of the landings. Fishing effort data available for longline 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 always possible therefore only trips which reported one area category and one gear fished were included in these analyses. In addition, only trips that took place exclusively in the Gulf of Mexico were included in the analysis. Outliers were removed from the data by looking at the following variables and removing trips where values in at least one of these variables fell above or below the 99.5 percentile: trip length, number of lines for handline or number of sets for longline, number of hooks per line, number of crew, and the hours fished per day.

The coastal logbook program reports the area where it fished in blocks created by lines of the same integer latitude and longitude. For greater amberjack, however, the number of trips taken within each of the areas was not enough for the analysis, so neighboring areas were grouped into larger areas as follows:

- Areas 17-22 = west LA and TX

- Areas 12-16 = LA
- Areas 6-11 = NW Florida and AL
- Areas 4 and 5 = west FL
- Areas 2 and 3 = SW Florida

As was done in the last assessment, Area 1, which corresponds predominantly to the Florida Keys and the Dry Tortugas, was dropped because greater amberjack do not frequently occur here. These larger areas are referred to here on out as the variable "new_area."

The Stephens and MacCall (2004) multispecies approach to subsetting logbook data was used to select trips that theoretically fished in habitat where greater amberjack could have been caught. This approach uses the species composition of each trip in a logistic regression of species presence/absence to infer if effort on that trip occurred in similar habitat to greater amberjack. If effort on a trip was determined to occur in similar habitat to greater amberjack, then that trip was used in the analysis (Stephens and MacCall 2004). In addition, any trips that may have caught exclusively greater amberjack were kept in the dataset and included in the analysis.

Starting in 1998, a closed season was implemented for greater amberjack during which the fishery was closed March through May. As a result, for all years in the dataset (1993 through 2012), trips that took place in the months of March, April, and May were removed. These closed season months were also removed for years prior to 1998 because including these months in the model for only some years could bias the index due to seasonal differences in abundance of greater amberjack. Electric reel (bandit) and manual handline were combined and will be referred to as handline. The reason they were combined is because they are often reported together on the same trip, or one may be reported in place of the other, and as a result, it is not possible to apportion fishing effort. Trips that claimed to report fishing both handline and longline gear were not included because it was not possible to apportion fishing effort. For handline gear, only trips that fished less than ten hooks per line were included in the analysis. For longline gear, only trips that reported at least 10 sets per day or trip duration of only one day were included in the analysis. The subset of data from these two gears in these ways was done in order to maintain consistency with what was done during the last assessment. Fishing effort for the handline fishery was defined as hook days, while effort for the longline fishery was defined as every 100 hooks. According to the previous assessment and update assessment reports, this is what was done during the last assessment.

In order to remove the influence of extraneous factors on trends in abundance, the indices were statistically standardized using a delta lognormal approach. The Delta lognormal modeling approach combines separate generalized linear model (GLM) analyses of the proportion of successful trips (trips that landed red snapper) and the catch rates on successful trips to construct a single standardized CPUE index (Lo et al. 1992, Hinton and Maunder 2004, Maunder and Punt 2004). Parameterization of each model was accomplished using a stepwise approach and Akaike's information criteria (AIC). For each GLM procedure 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 the proportion of successful trips across strata. 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 calculated as the natural log of CPUE. The catch per unit effort was calculated on an individual trip basis and was equal to the number of fish caught on a given trip divided by the effort, where effort was the product of the number of anglers in the group that was interviewed and the total hours fished.

A stepwise approach was used to quantify the relative importance of the explanatory factors. First a GLM model was fit to the null model (only the intercept) and the AIC, deviance and degrees of freedom were calculated. Next, a suite of models was tested where each potential explanatory factor was added to the null model. Again, the AIC, deviance, and degrees of freedom were calculated. The model with the factor that had the lowest AIC became the new base model and the process was repeated adding factors individually until either the AIC was no longer further reduced or the all he factors were added to the model. In addition to screening using AIC, factors were also screened and not added to the model if the reduction in deviance per degree of freedom was less than one percent. This screening was implemented in order to fit a more parsimonious model, given the fact that factors which reduce the deviance by so little exert little influence on the index trend. Two-way interactions among significant main effects were examined and significant interaction effects were included in the model. The final delta-lognormal model was fit using a SAS macro, GLIMMIX (Little et al. 1996, Russ Wolfinger, SAS Institute). To facilitate visual comparison, a relative standardized index and relative nominal CPUE series were calculated by dividing each value in the series by the mean value of the entire time-series.

Results and Discussion

For the handline fishery, many exploratory attempts were made to replicate the index that was estimated during the last assessment for the years that it would overlap with this current assessment (Figure 1). This was done by following the methodology described during the previous assessment and update assessment, as well as the computer code from the last assessment however these efforts were not successful. One difference was that the list of species that was selected by the Stephens and MacCall (2004) procedure was different from that which was selected during the previous assessment and the last update assessment. However, did not describe the difference because even when the same list of species that was selected last time is hard coded and used with the new years of data, the index could not be replicated. Different calculations of fishing effort were also explored, however this did not match what was done last time. Finally, an index of all trips was constructed without applying the Stephens and MacCall (2004) approach (or any other approach), and this still did not match what was done last time. However, when the final cleaned dataset from last time is used to calculate the indices, the index from last time could be replicated. Thus, it appears that for some reason that could not be identified, the trips that were used for the analysis last time are different from those used in the analysis this time (Figure 1). Thus, it was recommended by the indices working group during the data workshop that the handline index produced for this document, using the Stephens and MacCall (2004) approach be used in this assessment. For the longline fishery, we did not experience this problem, and the index estimated for the previous update assessment matches the index that was estimated for this current assessment during the years that they overlap (Figure 2). As a result, this longline index was recommended by the indices working group during the data workshop for inclusion.

Sample size tables by year for the handline and longline fleets show the number of total trips, positive trips (those that caught an amberjack), and the proportion of positive to total trips that were made in the cleaned data set both before and after the Stephens and MacCall (2004) approach was applied (Tables 1 and 2). The tables also provide the proportion of total trips that were retained by the Stephens and MacCall (2004) approach. Tabular results showing the list of species selected by the Stephens and MacCall (2004) approach and their corresponding estimated coefficients for the handline and longline fleets can be found in Tables 3 and 4. Results of the fit to the logistic regression for the Stephens and MacCall (2004) approach can be found in Figure 3 for handline and figure 6 for longline. These figures show that for both the handline and longline portions of the data, the model fit the observed number of trips well, and found a minimum probability threshold.

Tables 5 and 6 provide the final deviance tables and final AIC values for the binomial (Table 5) and lognormal (Table 6) components of the handline model. For the handline fleet, the final binomial model found year, new_area, and fishing effort to be statistically significant and reduce more than one percent of the deviance, while the final lognormal model found year, new_area, and the interaction between year and new_area to be statistically significant, and reduce more than one percent of the deviance. Tables 7 and 8 provide the final deviance tables and final AIC values for the binomial (Table 7) and lognormal (Table 8) components of the longline model. For the longline fleet, the final binomial model found year and new_area to be statistically significant and reduce more than one percent of the deviance, while the final lognormal model found year, new_area, and the interaction between year and new_area to be statistically significant, and reduce more than one percent of the deviance.

Final estimated abundance index values, and their associated coefficients of variation, and upper and lower confidence intervals can be found in Table 9. Figures 4 and 7 present plots of the nominal and standardized index values by year, with confidence limits, as well as the model fit to the binomial portion of the model for the handline and longline fleets respectively. In general, the indices for the two fleets are divergent with the handline index suggesting an increasing trend in the recent years and the longline index suggesting a decreasing trend in the recent years. Such a disagreement could be the artifact of changes in fisher behavior, such as the longline fleet being forced to fish further offshore in more recent years in response to changing gear based regulations. Residual patters for each fleet's index (Figures 5 and 8) are generally normally distributed and without any biasing trends.

In summary, the data workshop indices of abundance working group recommended that both the handline and longline indices be included in the assessment. Some of the reasons for this are because both indices represent a complete census of the commercial fishing activity geared toward greater amberjack in the Gulf of Mexico, both indices cover the entire Gulf of Mexico, and both indices cover a decent time period – from 1993 to the present.

References

Hinton, Michael G. and M.N. Maunder. 2004. Methods for standardizing CPUE and how to select among them. Col. Vol. Sci. Pap. ICCAT 56(1): 169-177.

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 deltalognormal models. Can. J. Fish. Aquat. Sci. 49: 2515-2526.

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

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

Tables

Table 1: Sample sizes of total and positive trips before and after the Stephens and MacCall (2004) trip selection, and the number of trips retained by this procedure, for greater amberjack from the handline fishery.

Year	totalTrips_all	posTrips_all	ppos_all	totalTrips_subset	posTrips_subset	ppos_subset	Percent_Trips_Retained
1993	5502	778	14.14	1329	777	58.47	24.15
1994	5798	763	13.16	1293	763	59.01	22.30
1995	5931	876	14.77	1536	874	56.90	25.90
1996	6486	1029	15.86	1801	1029	57.13	27.77
1997	6741	876	13.00	1513	876	57.90	22.44
1998	7684	690	8.98	1446	690	47.72	18.82
1999	6474	631	9.75	1018	631	61.98	15.72
2000	6398	614	9.60	915	614	67.10	14.30
2001	6502	596	9.17	1078	596	55.29	16.58
2002	6071	688	11.33	1192	688	57.72	19.63
2003	5907	793	13.42	1214	793	65.32	20.55
2004	5219	736	14.10	1125	736	65.42	21.56
2005	4322	607	14.04	996	607	60.94	23.04
2006	4425	439	9.92	734	437	59.54	16.59
2007	3920	271	6.91	457	271	59.30	11.66
2008	3872	311	8.03	523	311	59.46	13.51
2009	4461	337	7.55	566	337	59.54	12.69
2010	2937	200	6.81	324	200	61.73	11.03
2011	3513	194	5.52	387	194	50.13	11.02
2012	3417	79	2.31	290	79	27.24	8.49

Table 2: Sample sizes of total and positive trips before and after the Stephens and MacCall (2004) trip selection, and the number of trips retained by this procedure, for greater amberjack from the longline fishery.

Year	totalTrips_all	posTrips_all	ppos_all	totalTrips_subset	posTrips_subset	ppos_subset	Percent_Trips_Retained
1993	1007	251	24.93	368	251	68.21	36.54
1994	1219	262	21.49	397	262	65.99	32.57
1995	1193	260	21.79	417	260	62.35	34.95
1996	1235	259	20.97	377	259	68.70	30.53
1997	1389	327	23.54	510	327	64.12	36.72
1998	1211	250	20.64	445	250	56.18	36.75
1999	981	242	24.67	365	242	66.30	37.21
2000	935	234	25.03	382	234	61.26	40.86
2001	917	232	25.30	335	232	69.25	36.53
2002	907	301	33.19	376	301	80.05	41.46
2003	1020	359	35.20	449	359	79.96	44.02
2004	1011	280	27.70	364	280	76.92	36.00
2005	802	249	31.05	301	249	82.72	37.53
2006	920	297	32.28	368	297	80.71	40.00
2007	689	230	33.38	313	230	73.48	45.43
2008	783	293	37.42	358	293	81.84	45.72
2009	342	141	41.23	210	141	67.14	61.40
2010	345	76	22.03	143	76	53.15	41.45
2011	452	49	10.84	181	49	27.07	40.04
2012	439	20	4.56	154	20	12.99	35.08

Table 3: Species selected by the Stephens and MacCall (2004) procedure for the handline data that theoretically co-occur with greater amberjack and the coefficient that was estimated for each species during the logistic regression.

Species Name	**Coefficient**
SNAPPER,YELLOWTAIL	-0.05
PORGY,RED,UNC	-0.12
KING MACKEREL	-0.50
SPANISH MACKEREL	-0.54
GROUPER,RED	-0.56
SNAPPER,RED	-0.56
SNAPPER,LANE	-0.70
GRUNT,WHITE	-0.80
SEA BASSE,ATLANTIC,BLACK,UNC	-1.07
TRIGGERFISH,GRAY	0.22
SNAPPER,MANGROVE	0.26
GROUPER,GAG	0.36
SNAPPER,MUTTON	0.62
COBIA	0.63
GROUPER,BLACK	0.67
SNAPPER,VERMILION	0.85
SCAMP	1.25

Table 4: Species selected by the Stephens and MacCall (2004) procedure for the longline data that theoretically co-occur with greater amberjack and the coefficient that was estimated for each species during the logistic regression.

Species Name	Coefficient
SNAPPER,RED	-0.02
SNAPPER,LANE	-0.04
SHARK,SANDBAR	-0.07
GROUPER,RED	-0.58
HIND,SPECKLED	0.05
SNAPPER,MANGROVE	0.12
SNAPPER,VERMILION	0.17
SCAMP	0.26
TRIGGERFISH,GRAY	0.27
SNAPPER,SILK	0.29
DOLPHINFISH	0.32
PORGY,RED,UNC	0.32
COBIA	0.32
GROUPER,BLACK	0.33
GROUPER,GAG	0.34
TILEFISH	0.41
MARGATE	0.42
GROUPER,SNOWY	0.43
SCORPIONFISH-THORNYHEADS	0.47
GROUPER,WARSAW	0.53
GROUPER,YELLOWEDGE	0.55
TILEFISH,BLUELINE	0.57
SNAPPER,MUTTON	0.92

Table 5: Final deviance table for the handline binomial model measuring presence-absence of greater amberjack on a trip. The table shows the order of the factors as they were added sequentially to the model such that fit diagnostics listed for each factor were the diagnostics from a model that included that factor and all of the factors listed above it in the table.

Factor	Df	Deviance	Resid. Df	Resid. Dev	AIC	perc_dev_reduced	log_likelihood	like_ratio	P_value
Null	1	26817.35	19736.00	26817.35	26819.35	NA	-13408.68	NA	NA
new_area	4	838.08	19732.00	25979.27	25989.27	3.13	-12990.63	836.08	0.00
new_area + effort	7	956.27	19725.00	25023.00	25047.00	3.57	-12516.50	948.27	0.00

Table 6: Final deviance table for the handline lognormal model for greater amberjack. The table shows the order of the factors as they were added sequentially to the model such that fit diagnostics listed for each factor were the diagnostics from a model that included that factor and all of the factors listed above it in the table.

Factor	Df	Deviance	Resid. Df	Resid. Dev	AIC	perc_dev_reduced	log_likelihood	like_ratio	P_value
Null	1	7478.61	4611.00	7478.61	15321.66	NA	-7659.83	NA	NA
new_area	4	415.33	4607.00	7063.28	15066.14	5.55	-7529.07	261.52	0.00
new_area + year	19	456.20	4588.00	6607.08	14796.21	6.10	-7379.11	299.93	0.00
new_area + year + new_area*year	74	325.68	4511.00	6263.46	14703.89	4.35	-7277.95	202.32	0.00

Table 7: Final deviance table for the longline binomial model measuring presence-absence of greater amberjack on a trip. The table shows the order of the factors as they were added sequentially to the model such that fit diagnostics listed for each factor were the diagnostics from a model that included that factor and all of the factors listed above it in the table.

Factor	Df	Deviance	Resid. Df	Resid. Dev	AIC	perc_dev_reduced	log_likelihood	like_ratio	P_value
Null	1	8572.848	6812	8572.848	8574.848	NA	-4286.42	NA	NA
year	19	564.6374	6793	8008.211	8048.211	6.586346	-4005.11	562.6374	2.24E-124
year + new_area	4	30.83797	6786	7942.989	7996.989	0.359717	-3994.49	24.83797	6.24E-07

Table 8: Final deviance table for the longline lognormal model for greater amberjack. The table shows the order of the factors as they were added sequentially to the model such that fit diagnostics listed for each factor were the diagnostics from a model that included that factor and all of the factors listed above it in the table.

Factor	Df	Deviance	Resid. Df	Resid. Dev	AIC	perc_dev_reduced	log_likelihood	like_ratio	P_value
Null	1	7478.61	4611.00	7478.61	15321.66	NA	-7659.83	NA	NA
new_area	4	415.33	4607.00	7063.28	15066.14	5.55	-7529.07	261.52	0.00
new_area + year	19	456.20	4588.00	6607.08	14796.21	6.10	-7379.11	299.93	0.00
new_area + year + new_area*year	74	325.68	4511.00	6263.46	14703.89	4.35	-7277.95	202.32	0.00

Table 9: Estimates of the handline and longline indices, with their upper (UCI) and lower (LCI) confidence intervals, and their corresponding coefficient of variations (CV).

YEAR	HANDLINE GEAR				LONGLINE GEAR			
	CPUE	LCI	UCI	CV	CPUE	LCI	UCI	CV
1993	0.617	0.304	1.253	0.366	0.683	0.490	0.952	0.167
1994	0.569	0.280	1.154	0.365	0.569	0.408	0.795	0.168
1995	0.713	0.352	1.444	0.364	0.715	0.514	0.994	0.166
1996	0.729	0.360	1.477	0.364	0.707	0.509	0.983	0.166
1997	0.618	0.306	1.250	0.363	0.772	0.553	1.079	0.168
1998	0.602	0.300	1.207	0.359	0.582	0.413	0.821	0.173
1999	0.645	0.317	1.313	0.367	0.755	0.541	1.055	0.168
2000	0.891	0.436	1.821	0.369	0.799	0.570	1.120	0.170
2001	0.669	0.331	1.353	0.363	0.900	0.638	1.269	0.173
2002	0.986	0.488	1.992	0.363	1.258	0.900	1.759	0.169
2003	1.278	0.631	2.590	0.364	1.247	0.893	1.741	0.168
2004	1.269	0.627	2.567	0.364	1.438	1.024	2.021	0.171
2005	0.879	0.432	1.790	0.367	1.763	1.244	2.500	0.176
2006	0.947	0.470	1.908	0.361	1.385	0.983	1.952	0.173
2007	0.881	0.437	1.777	0.362	1.104	0.772	1.579	0.180
2008	0.964	0.477	1.948	0.363	1.551	1.094	2.200	0.176
2009	0.913	0.454	1.836	0.360	1.578	1.098	2.266	0.183
2010	1.829	0.897	3.727	0.368	1.354	0.880	2.083	0.218
2011	2.557	1.291	5.062	0.352	0.448	0.256	0.784	0.285
2012	1.444	0.734	2.841	0.348	0.389	0.163	0.927	0.456

Figures

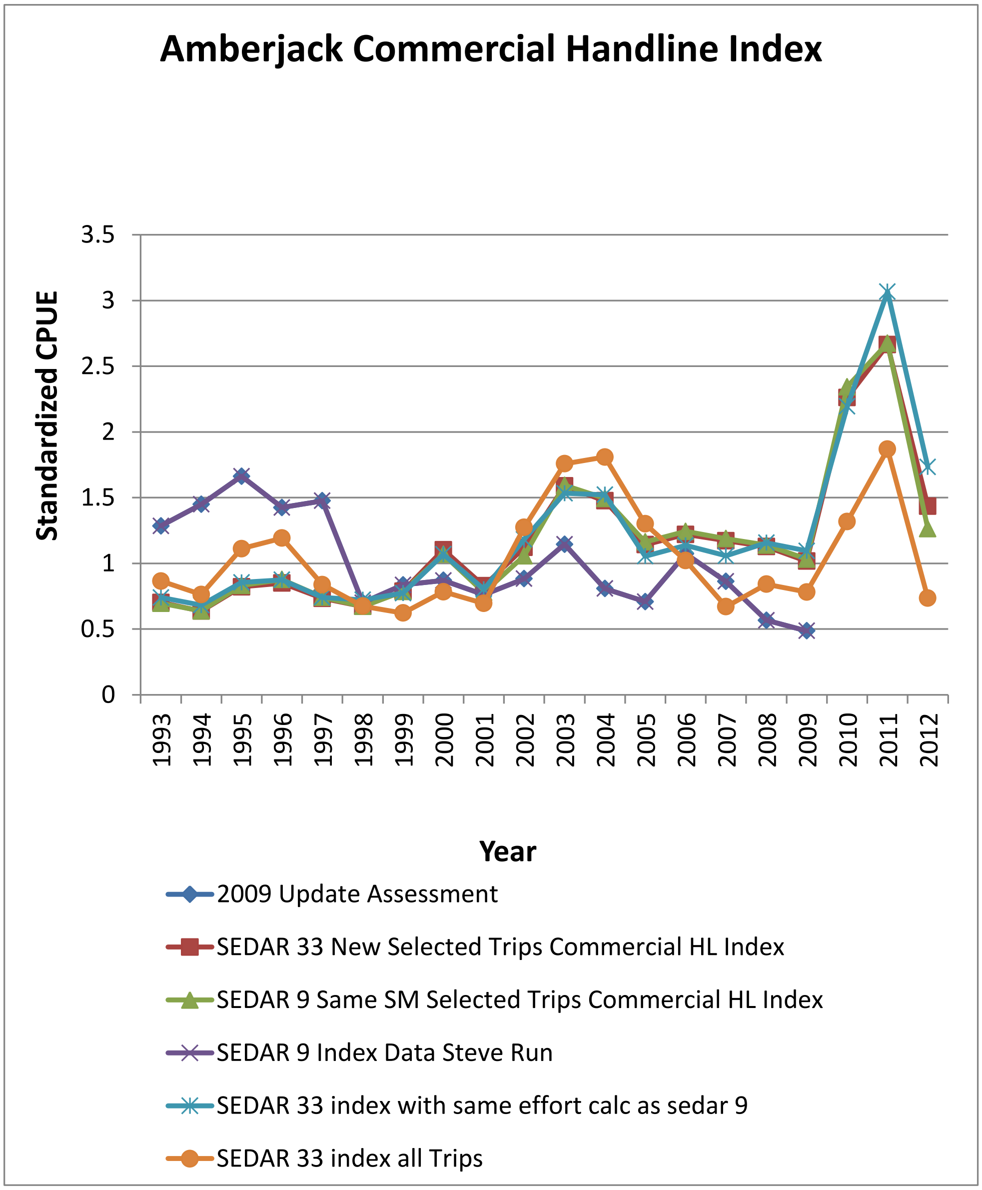

Figure 1: Comparison of the index estimated during the 2009 update assessment, with attempts to replicate that index for this assessment as described in the results section.

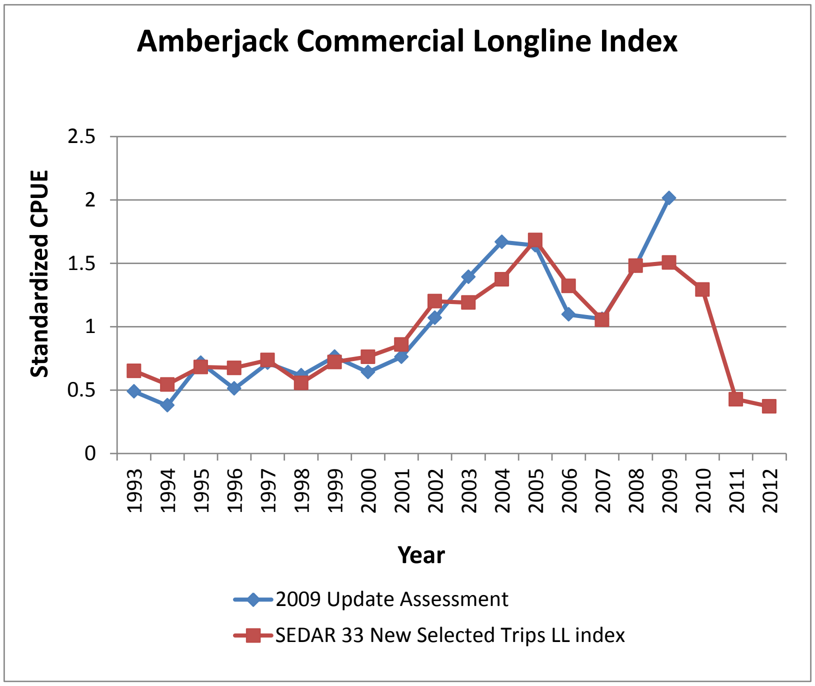

Figure 2: Comparison of the index estimated during the 2009 update assessment with the index calculated for this assessment.

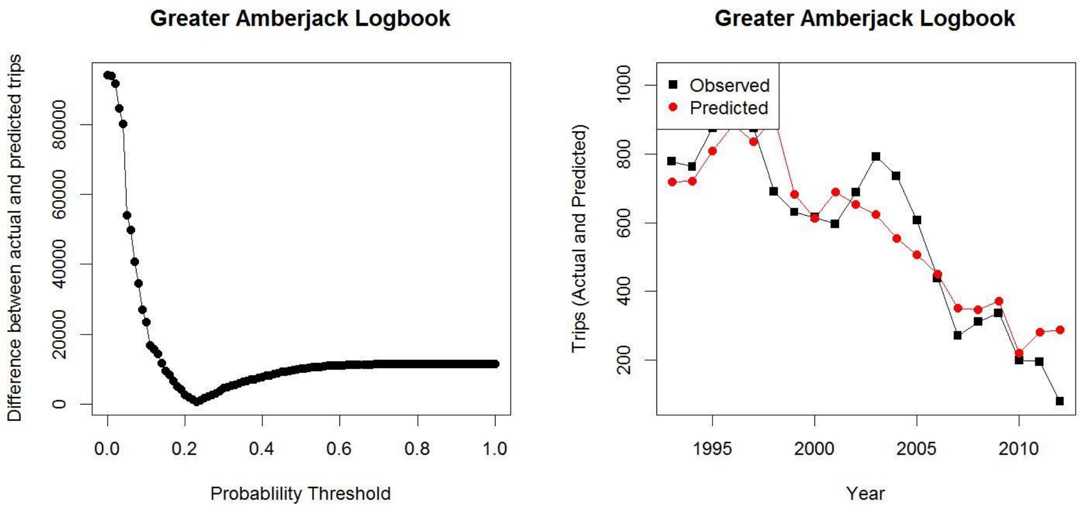

Figure 3: Stephens and MacCall (2004) model diagnostics from the binomial trip selection procedure for the handline fleet.

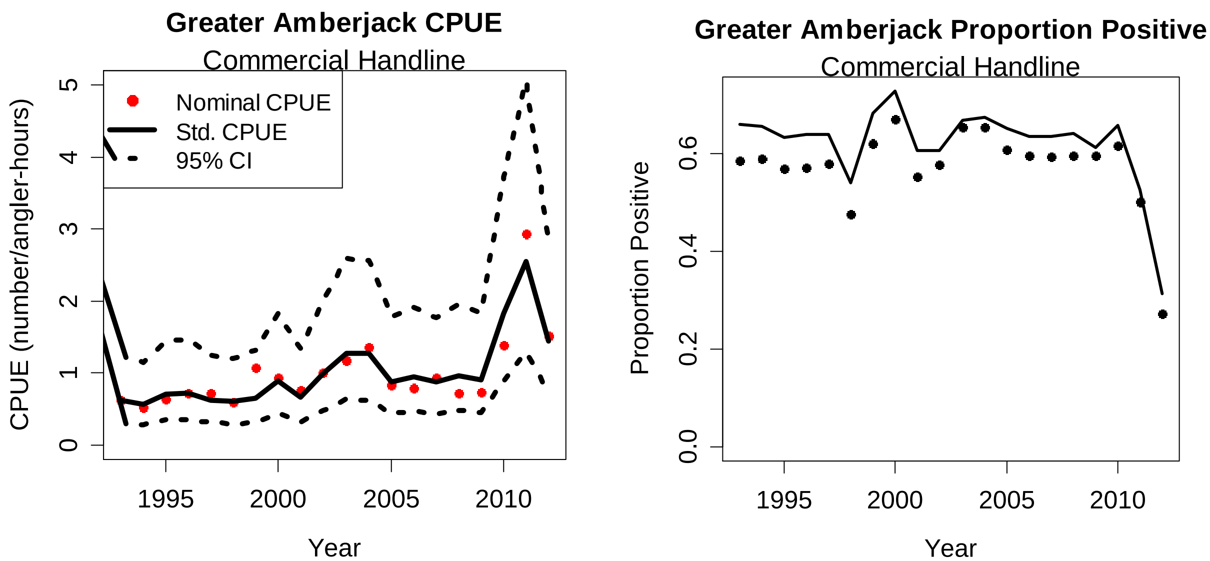

Figure 4: Standardized index with observed nominal CPUE values, and upper and lower confidence intervals for the recommended handline index on the left. Fit of the binomial proportion positive model to the observed proportion positive values for the recommended handline index on the right.

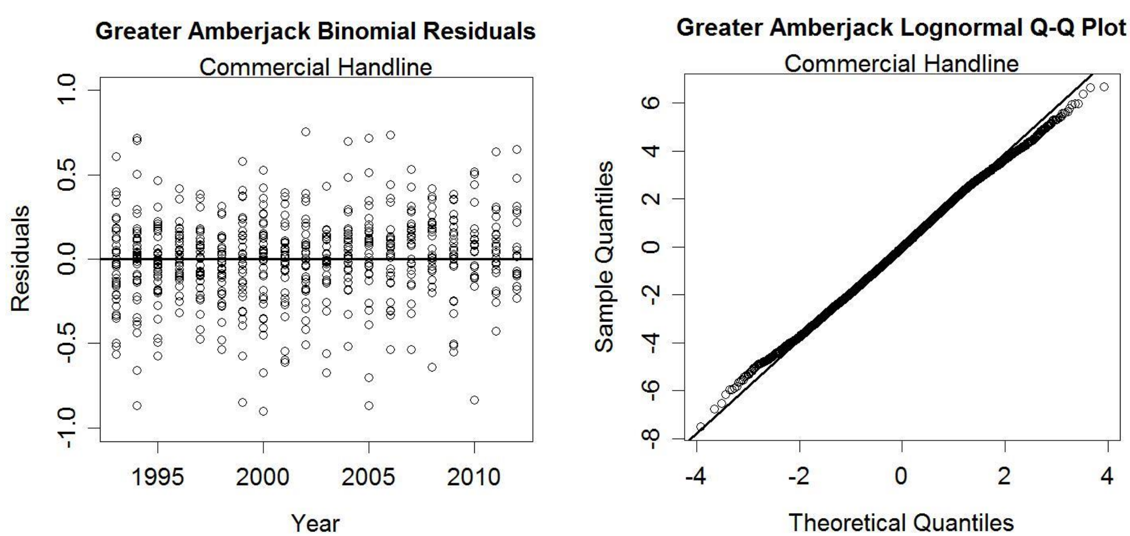

Figure 5: Model diagnostics for the binomial (left) and lognormal (right) fits to the handline index standardization.

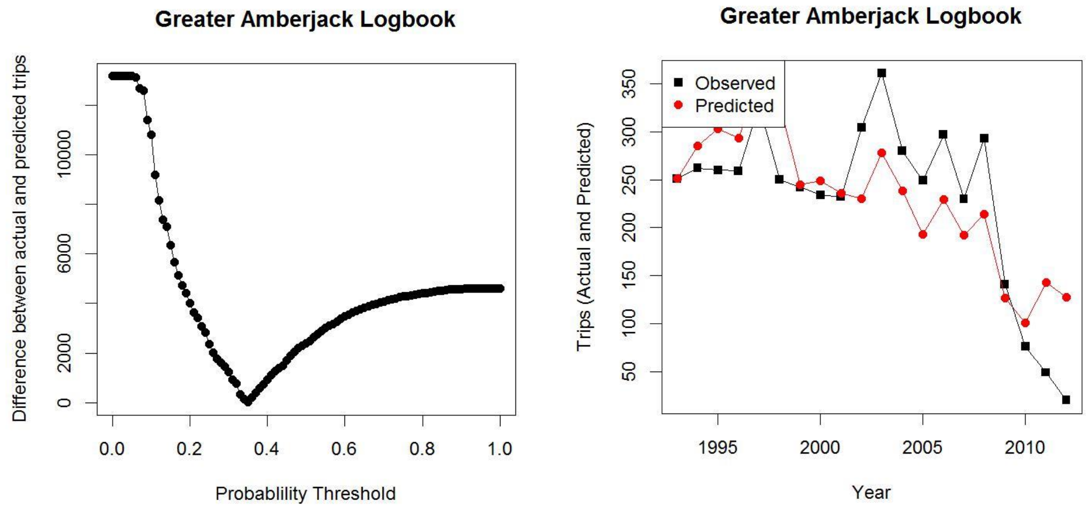

Figure 6: Stephens and MacCall (2004) model diagnostics from the binomial trip selection procedure for the longline fleet.

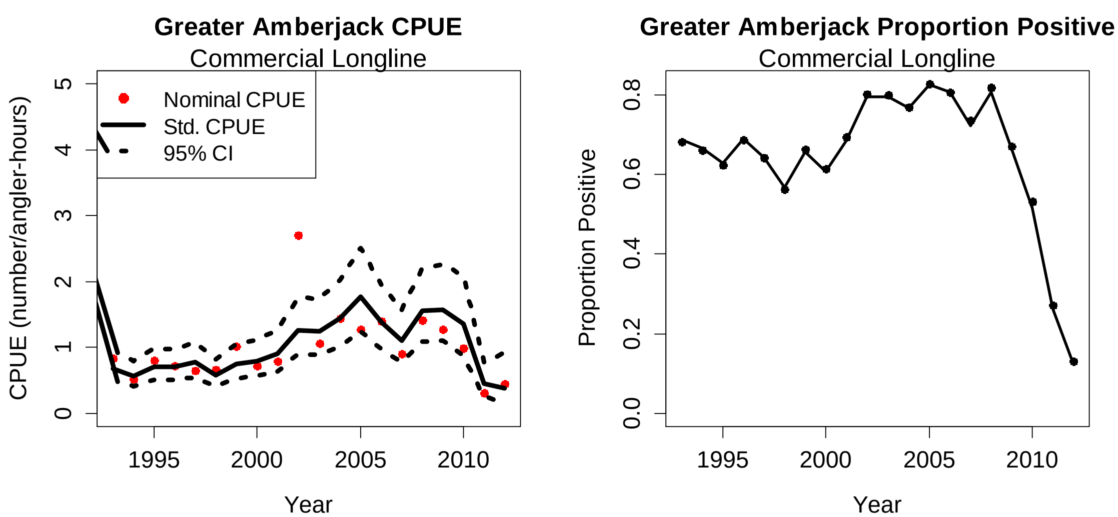

Figure 7: Standardized index with observed nominal CPUE values, and upper and lower confidence intervals for the recommended longline index on the left. Fit of the binomial proportion positive model to the observed proportion positive values for the recommended longline index on the right.

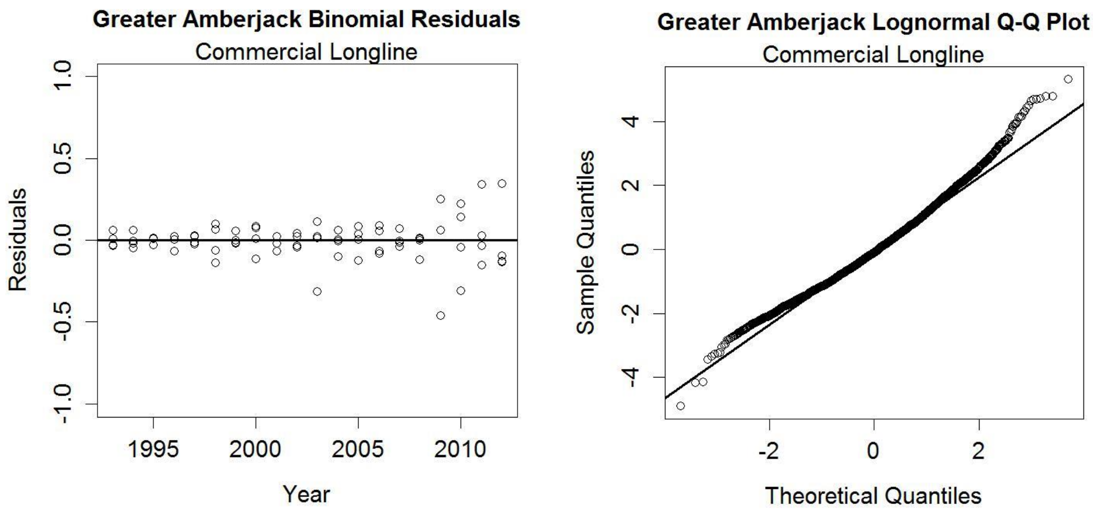

Figure 8: Model diagnostics for the binomial (left) and lognormal (right) fits to the longline index standardization.