Standardized Abundance Indices for Gulf of Mexico Gray Triggerfish (*Balistes capriscus*) Based on Catch Rates as Measured by the NMFS Southeast Zone Headboat Survey

by

Joshua Sladek Nowlis

23 June 2005

NOAA Fisheries
Southeast Fisheries Science Center
75 Virginia Beach Drive
Miami, FL 33149

Sustainable Fisheries Division Contribution No. SFD 2005 - ###

INTRODUCTION

Stock assessment models require indices of stock abundance. Ideally we would know exactly how many fish there were over the course of an extended time period but we never have comprehensive stock monitoring for marine fish stocks. Instead, we have to rely on sampling schemes to estimate abundance. Whenever possible, it is preferable to rely on stratified random fishery-independent sampling. However, such indices are limited for triggerfish to specific life stages, short timeframes, or both. In the meantime, we have to rely heavily on indices of abundance produced from catch rates in various fisheries. These indices require that we have measures of catches and of effort.

The NOAA-NMFS Southeast Zone Headboat Survey collects information on the catch and effort of headboats, on which anglers pay individually to join the fishing expedition. Each trip is characterized for its fishing effort and catch by species in numbers of fish.

Trips were excluded if key data were missing or previously flagged as a likely duplicate or misentry. Trips were further restricted to the Gulf of Mexico (headboat areas 18 to 27).

This survey has high sample sizes, characterizes a fishery which catches multiple species per trip, and has apparently strong data going back to its inception in 1986. It may have a gap in geographic coverage (area descriptions would suggest that Mississippi is not included, and sample sizes for Louisiana are very small) and does not include estimates of discarded fish. As such, this index will underestimate the actual catch per unit effort. This limitation would only present a problem if the discarding rates varied across years or other factors.

METHODS

Species Associations

Using fishery-dependent data to develop abundance indices presents problems. Unlike scientific sampling, fishing trips will vary in their likelihood of catching the species of interest. As a result, the catch rates from an active fishery may be less indicative of stock abundance than scientific sampling. Nonetheless, one can potentially infer abundance if care is taken to classify fishing trips and focus on a set of them that provides some consistency through time and across different locations. Care should be taken to include trips that are likely to catch the species of interest in order to provide adequate samples for statistical analyses.

Stephens and MacCall (2004) developed a statistical approach for identifying a subset of all trips of this sort. Their approach uses logistical regression to categorize trips. It develops correlation coefficients between the presence or absence of the species of interest and the presence or absence of every other species. In our case, we limited our consideration to those species that occurred in at least 1 percent of the recorded trips. These coefficients are then used to assign to each trip a probability that it would catch the species of interest based on the presence or absence of other species. Finally, it uses a minimization procedure to select a cutoff probability for which trips to include or exclude. Their paper provides greater technical detail.

Conceptually, this approach is designed to identify fishing trips that were likely to catch the species of interest using the other caught species as an indicator of habitat, gear, and fishing behavior (e.g., time of day, bait use, etc.). As such, it identifies a subset of all trips that were generally likely to catch the species of interest, whether or not that species was caught. One possible limitation of this technique is its reliance on the occurrence of other species. As a result, trips cannot be incorporated into this technique if they do not catch other species.

Some might criticize a catch rate method that ignores trips that caught the focal species. Yet this concern misconstrues the goal, which is to identify trips based on their consistency and then determine the catch rates of the focal species, not vice versa.

Standardization Procedure

In addition to the challenge of inconsistent fishing behavior, fishery-dependent catch-rate based abundance indices are also likely to suffer from a lack of random sampling. The non-randomness comes partly in the form of fishing behavior, which may not correspond to abundance. This challenge cannot be addressed without some fishery-independent measure of abundance. The non-randomness also comes in the form of the waxing and waning of fishing and sampling across time, space, and other factors (e.g., gear). Several statistical techniques exist to standardize catch rates to account for this latter challenge.

We used a method developed by Lo and colleagues (1992). This delta-lognormal technique uses standard generalized linear models (GENMOD; Version 8.02 of the SAS System for Windows © 2000, SAS Institute Inc., Cary, NC, USA) to identify time, space, and other factors that are likely to influence catch rates. It also combines two forms of information: the frequency with which trips catch the species of interest and, on those trips that were successful, the catch per unit time. It assumes a binomial distribution for logit-transformed success data and a normal distribution for ln-transformed catch per unit effort (CPUE) data. The end result is a standardized catch rate per year with an associated standard error.

Six factors were considered for inclusion in the delta-lognormal model. These included year, which was forced in the model due to our interest in generating patterns through time; season, which was defined using Jan-Mar as winter, Apr-Jun as spring, Jul-Sep as summer, and Oct-Dec as autumn because this appeared to fit the data as well as any pattern (Fig. 1); state; vessel, limited to those that had at least 30 trips identified in the species association technique; time of day, defined as daytime, nighttime, or unknown/both; and trip duration, defined as a half-day, full day, or multi-day.

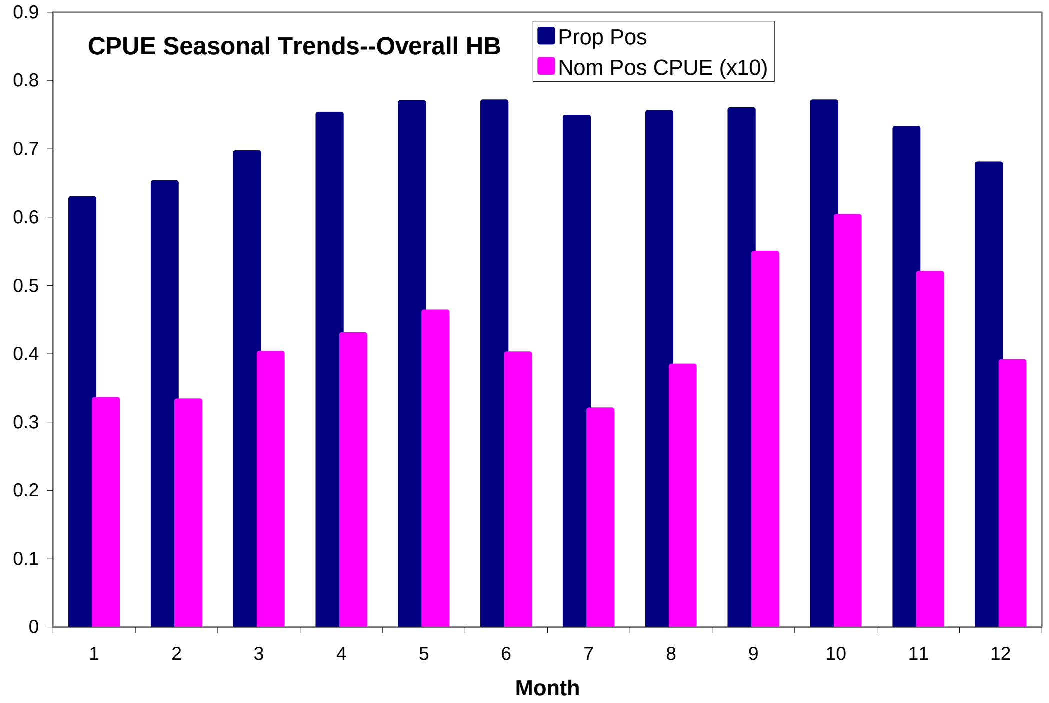

Fig. 1—Seasonal nominal catch rates.

These factors, and their two-way interactions, were tested in the standardization procedure, but only included if they provided a significant improvement in fit to the model. A significant improvement was defined as a significant Chi-square statistic (at the $\alpha = 0.05$ level) and an overall improvement in fit to the model of at least 1 percent reduction in deviance per degree of freedom.

The headboat survey provided 137,768 trips of potential interest from the Gulf of Mexico, characterized by 689,178 records (species by trip). Of all species, gray trigger was encountered on 63,788 of these trips—46% of the time, 1st among all species. When records of rare species (landed in < 1% of trips) were eliminated, there were 137,403 trips consisting of 659,937 records for 47 species. When the species association procedure was run, it identified 64,006 trips as likely to have caught gray trigger, 47,338 (74%) of which actually caught gray trigger. Finally, vessels were eliminated from consideration if they made fewer than 30 trips from within this data set. This procedure reduced the number of vessels under consideration from 161 to 103, while lowering the number of trips from 64,006 to 63,391, 47,072 (74%) of which landed gray trigger.

RESULTS

A number of species were likely to co-occur with gray triggerfish, while others were likely to indicate that gray trigger were not present (Table 1). Vermilion snapper was most strongly associated with gray triggerfish, followed by red snapper. Red porgy and bank seabass were also fairly closely associated. Blackfin tuna was least likely to co-occur with gray trigger, followed by crevalle jack, Atlantic sharpnose shark, pinfish, king mackerel, great barracuda, and greater amberjack. Other species correlations are listed in Table 1.

Once appropriate trips were identified, sample sizes were examined (Table 2). These were adequate for most strata across all factors, with a few notable exceptions. There were fewer than 50 trips per year prior to 1986 and far fewer samples from the West, both problems that were raised earlier. Additional concerns include the relative lack of samples in the winter, which might be addressed through the change in seasons suggested above. Private boat samples were also relatively rare, as were samples from 1989 and 19990, although they are probably sufficient for analysis.

The proportion of trips that caught gray triggerfish (ProPos) and the catch per unit effort on positive trips (CPUE) are shown in Fig. 2. These data suggest the stock may have increased early in the time series and decreased more recently.

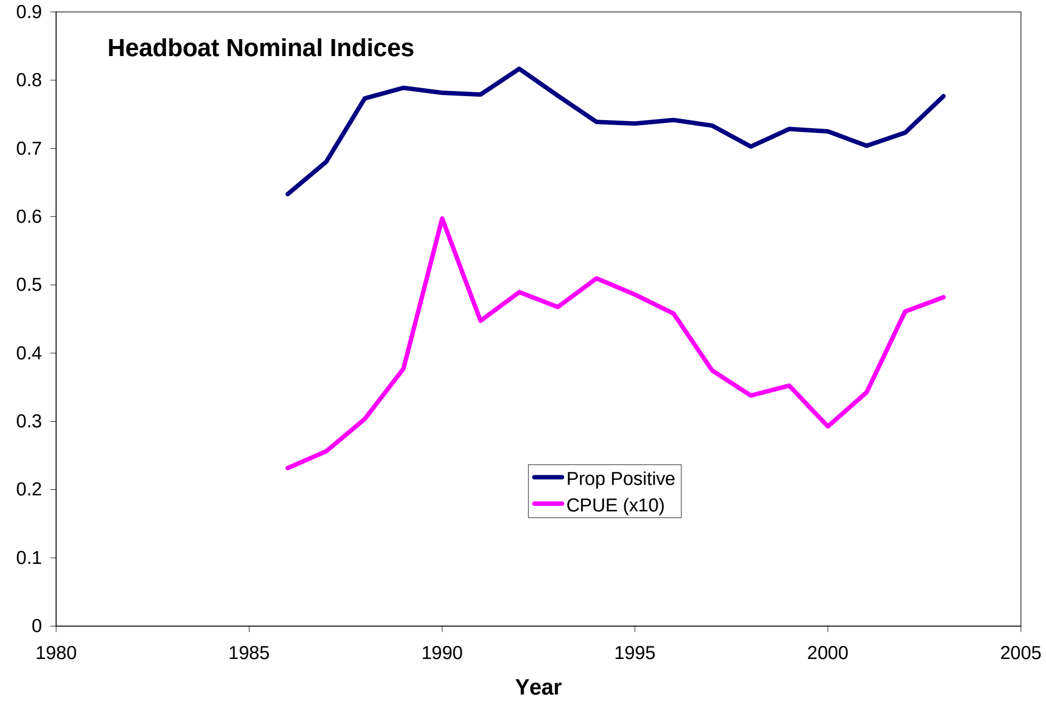

Fig. 2— Nominal CPUE for Gulf of Mexico gray triggerfish. Dark line shows proportion of trips that caught gray trigger (Prop Pos) while light line shows the catch per unit effort for those positive trips.

Table 1—Species associations. Correlations in occurrence between gray trigger and other species.

Species	Correlation Coefficient
vermilion snapper	2.71009856
red snapper	1.736526524
red porgy	1.013877186
bank sea bass	0.914282234
jolthead porgy	0.682701009
littlehead porgy	0.564090907
atlantic spadefish	0.506504284
black sea bass	0.469062215
whitebone porgy	0.40295644
bluefish	0.343316721
gag	0.329953963
sandbar shark	0.294650191
knobbed porgy	0.249420357
grass porgy	0.249229598
warsaw grouper	0.214694573
summer flounder	0.212405108
spottail pinfish	0.203990874
rock hind	0.187333398
blue runner	0.114727484
lane snapper	0.094931465
pigfish	0.064537441
almaco jack	0.057268321
spanish mackerel	0.025397751
black grouper	-0.00968372
scamp	-0.049221459
sand perch	-0.093124719
tomtate	-0.099299618
gray snapper	-0.127402275
saucereye porgy	-0.139881406
white grunt	-0.143799987
sand seatrout	-0.152783352
dolphin	-0.15344849
little tunny	-0.235152534
blacktip shark	-0.265482832
yellowtail snapper	-0.291654234
cobia	-0.384340233
gulf flounder	-0.555454051
red grouper	-0.63409283
banded rudderfish	-0.685458697
greater amberjack	-0.75035656
great barracuda	-0.763480472
king mackerel	-0.833688501
pinfish	-0.882145796
atlantic sharpnose shark	-1.28327846
crevalle jack	-1.677967534
blackfin tuna	-1.70116068

Table 2—Sample sizes. Number of trips examined by various factors.

YEAR	Trips
1986	1632
1987	1973
1988	2780
1989	2827
1990	4109
1991	3881
1992	4318
1993	4926
1994	4842
1995	4411
1996	3974
1997	3787
1998	3800
1999	3124
2000	3081
2001	3274
2002	3095
2003	2950

Season	Trips
AUT	8182
SPR	22155
SUM	23516
WIN	8931

Red Snapper Season	Trips
CLSD	3273
OPEN	59511

State	Trips
AL	35952
FL	10097
LA	2088
TX	14647

Diagnostics of the delta-lognormal model indicate the results were robust (Fig. 3). Residuals appear to be evenly distributed and follow a normal distribution.

GLM results are presented in Table 3 and Fig. 4. The delta-lognormal modeling exercise identified the following significant factors and interactions: year, state, and year*state for the proportion positive data; and year, vessel, season, year*vessel, and year*season for the CPUE data. Vessel also significantly improved the fit (19.4% improvement) of the model to proportion positive data, as did the season*vessel interaction (2.2% improvement) to the CPUE data, but the inclusion of either of these factors prevented the delta-lognormal model from converging. Therefore, they were excluded. These results should be viewed as fairly fixed, although data from 2004 may be available and, if so, should be included in the model.

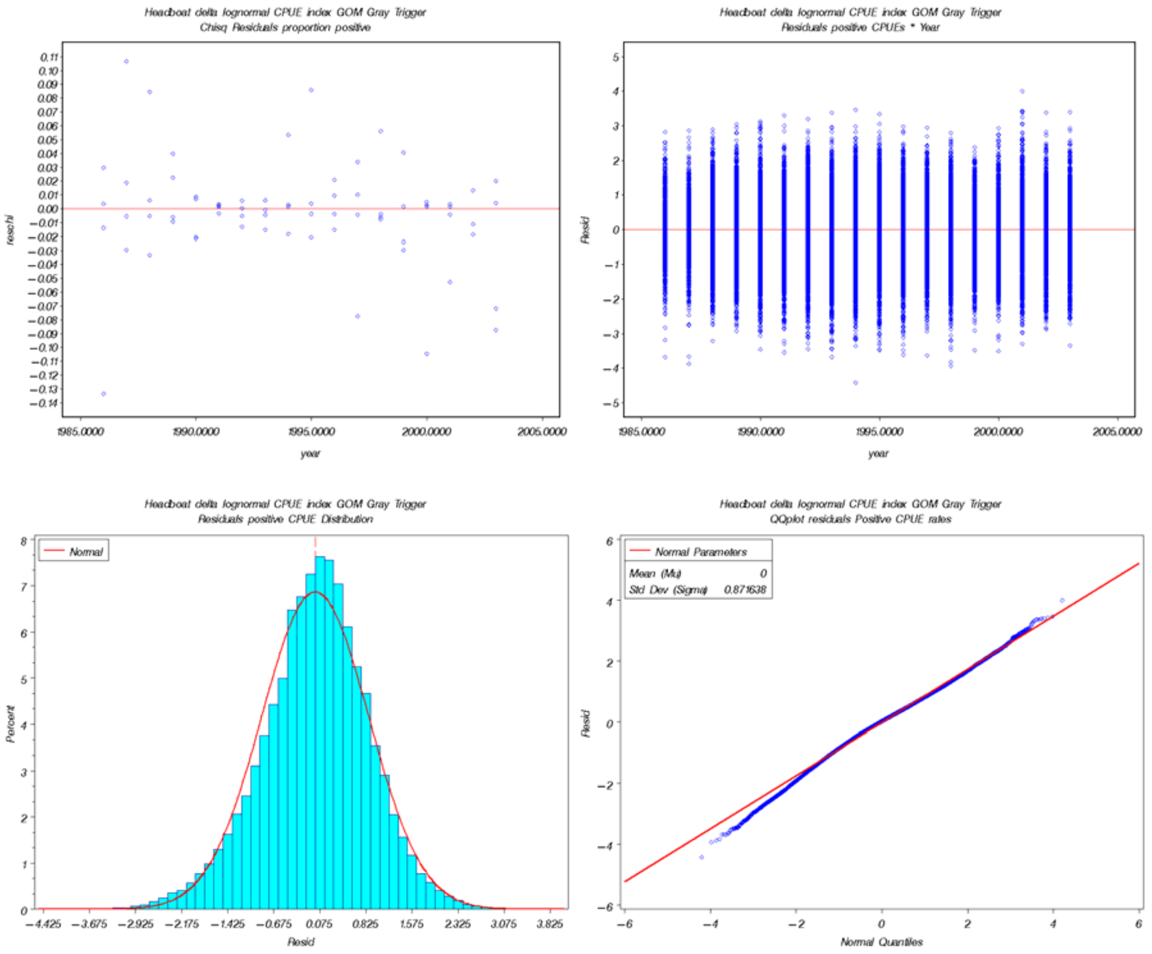

Fig. 3—Diagnostics of the delta-lognormal model. Residuals by year did not show biases or unidentified problems in either the proportion positive (a) or ln(CPUE) (b) portions of the model. Residuals overall of the ln(CPUE) portion fit well to a normal distribution (c), and a Q-Q plot (d) also validated the assumption of normality.

DISCUSSION

With the exception of the difficulty in fitting the vessel effect into the binomial portion of the model, which fit proportion positive data, these results appear to be robust and provide abundance information with relatively high confidence.

LITERATURE CITED

Lo, NC, LD Jackson, JL 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, A MacCall. 2004. A multispecies approach to subsetting logbook data for purposes of estimating CPUE. Fish. Res. 70: 299-310.

Table 3—Standardized index values per year and confidence intervals.

YEAR	ln(CPUE)	SE
1986	0.358655164	0.237241999
1987	0.435487806	0.23649647
1988	0.634518334	0.234248595
1989	0.892187844	0.230057171
1990	1.001365091	0.21588047
1991	1.072999442	0.216655152
1992	1.217308398	0.214982809
1993	0.92184561	0.210260264
1994	0.758721082	0.208563132
1995	0.557595107	0.213565638
1996	0.458898158	0.217863281
1997	0.370536961	0.221583583
1998	0.349205846	0.218645432
1999	0.346790609	0.226466607
2000	0.225677766	0.231348632
2001	0.125932546	0.229548839
2002	0.192833184	0.245457465
2003	0.348264209	0.245239032

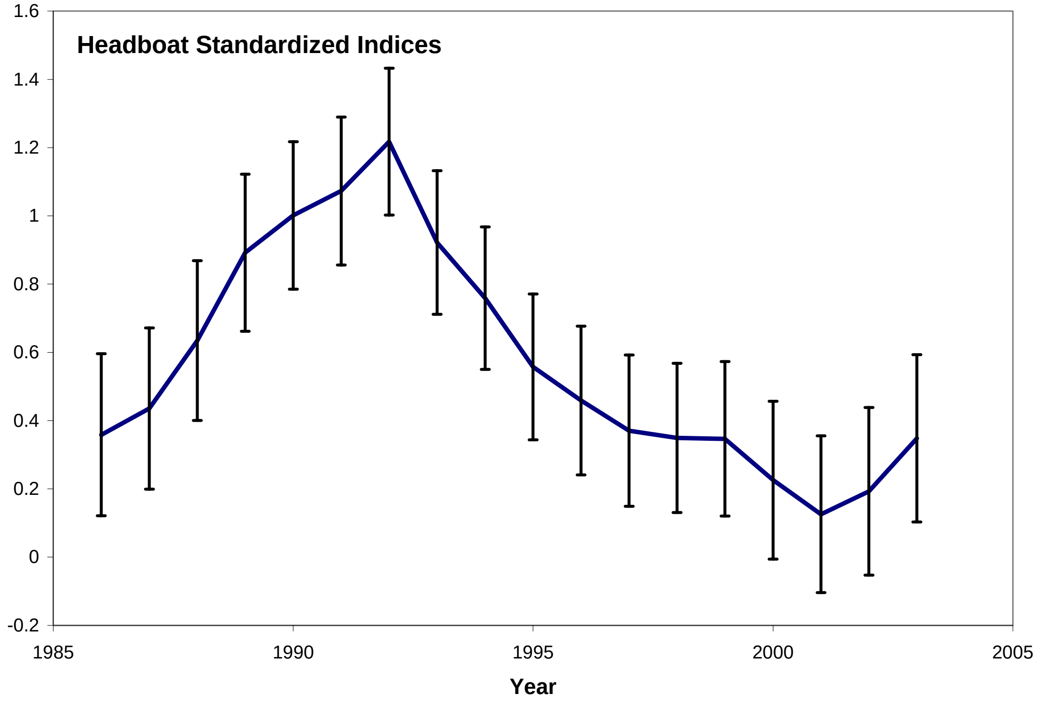

Fig. 4—Standardized index values per year. Ln(CPUE) values shown with error bars representing standard errors.