Standardized Catch Rate Indices for Gulf of Mexico Gray Triggerfish (*Balistes capriscus*) Landed During 1986-2013 by the Headboat Fishery

Matthew W. Smith, Daniel Goethel, Adyan Rios, and Jeff Isley

SEDAR43-WP-06

20 March 2015

Please cite this document as:

Smith, M.W., D. Goethel, A. Rios, and J. Isley. 2015. Standardized Catch Rate Indices for Gulf of Mexico Gray Triggerfish (*Balistes capriscus*) landed during 1986-2013 by the Headboat Fishery. SEDAR43-WP-06. SEDAR, North Charleston, SC. 18 pp.

Standardized Catch Rate Indices for Gulf of Mexico Gray Triggerfish (*Balistes capriscus*) Landed During 1986-2013 by the Headboat Fishery

by

Mathew W. Smith
Daniel Goethel
Adyan Rios
Jeff Isley

20 March 2015

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

INTRODUCTION

The National Marine Fisheries Service (NMFS) Southeast Zone Headboat Survey data set was updated through 2013 and the procedures outlined in SEDAR 9 were implemented to provide standardized abundance indices. Catch-per-unit effort (CPUE) is derived from the headboat data using total fish caught on a given trip divided by the amount of angler-hours spent fishing, where:

$$\ln(CPUE) = \ln\left[\frac{Number\ of\ Fish}{Anglers*Hours\ Fished}\right]. \quad \textbf{(1)}$$

For the headboat data set, effort was estimated in angler-hours where the number of hours spent fishing (i.e., 5, 7, 10 or >10 days) coincided with the type of trip (i.e., half, three-quarters, full or multi-day, respectively). Two indices (east and west Gulf of Mexico) were calculated based on geographic area (east or west of the Mississippi delta) to better represent the variance and abundance trends in each zone, because effort can vary significantly from year-to-year between the two areas. Trips were eliminated if they had missing values for any of the key factors, were in anyway incomplete, appeared to be misreported (e.g., reported zero anglers) or represented multiple entries for a single trip.

METHODS
Species Associations

An indirect method was necessary to infer targeting behavior of fishermen, because no direct information was available. Previously in SEDAR 9, the species associates subsetting routine proposed by Stephens and MacCall (2004) was implemented to select trips for use in the analysis. During the 2011 update assessment a new approach was adopted, which involved identifying a guild of species that frequently co-occur with gray triggerfish. The guild was defined as all fish in the NOAA reef fish management plan. The guild approach to trip selection was used in the SEDAR 43 assessment.

Index Standardization

A two-step delta-lognormal general linearized model (GLM; Lo et al. 1992) was used to standardize for variability and non-randomness in CPUE data collection methods not caused by the year effect (i.e., to factor out year to year variations in CPUE not due to changes in abundance). The combined approach first modeled the frequency with which trips caught the species of interest (i.e., proportion positive) using a logit regression assuming a binomial distribution of the response variable. In the second step, the logarithm of CPUE on successful trips (those that caught the target species) was used as the response variable assuming a normal distribution and an identity link function. The two models were then combined to provide the final standardized index of abundance.

A forward stepwise regression approach was utilized within the GENMOD procedure of SAS 9.2 (SAS Institute, 2008). In this procedure, potential factors were added to the base model one at a time based on the percent reduction in deviance per degree of freedom. With each run of the model the factor that caused the highest reduction in deviance was added to the base model (assuming the factor was significant based on a Chi-Square test with probability ≤ 0.05) until no factor reduced the percent deviance by the pre-specified level (i.e., 1%). Because the goal of the standardization was to model time trends in abundance, it was necessary to force the year effect as a factor even if it was not deemed significant. Two-way interaction terms were then investigated among each of the significant factors using the same stepwise approach. Higher order interactions were not tested.

The final delta-lognormal model was fit using the factors deemed significant in the GENMOD procedure using the SAS macro GLIMMIX (SAS Institute, 2008). Factors were modeled as fixed effects except for interaction terms involving year, which were modeled as random effects. Results of the binomial (proportion positive) and lognormal (mean CPUE on successful trips) were then multiplied to attain a single index of abundance based on the year effect.

RESULTS

Species Associations

There were 195,831 trip records available in the headboat data base from the Gulf of Mexico with 84,401 encountering gray triggerfish (19,276 from the western Gulf of Mexico and 65,125 from the eastern Gulf of Mexico). The guild approach retained 186,529 trips for use in the index standardization (46,013 from the western Gulf of Mexico and 140,516 from the eastern Gulf of Mexico). The proportion of positive trips before the subsetting routine was applied was 0.431 (0.353 from the western Gulf of Mexico and 0.461 from the eastern Gulf of Mexico), which increased to 0.452 after the subset was taken (0.419 from the western Gulf of Mexico and 0.464 from the eastern Gulf of Mexico).

Abundance Indices

A number of factors were investigated that could potentially influence yearly variations in catch rates including: Year, Season, Red Snapper Season (i.e., open or closed), Day/Night, Trip Duration, and Hours Fished. The Day/Night, Trip Duration and Hours Fished factors were only tested in the binomial model as these factors were confounded with effort for the CPUE response variable in the lognormal model). The levels and potential values for the various factors are provided in Table 1.

Western Gulf of Mexico

For the binomial component of the western Gulf of Mexico index the significant factors were Year and Trip Duration along with the Year*Trip Duration interaction term. A Vessel factor had been included in SEDAR 9, but complications resulted from using the Vessel factor within the binomial model and it was dropped as an explanatory variable to ensure model convergence. In the lognormal model Year, Red Snapper Season, and Season were found to be significant along with the Year*Season interaction terms. The final models were:

$$\begin{aligned} Proportion\ Positive &= Year + Trip\ Duration + Year * Trip\ Duration \\ \ln(CPUE) &= Year + Season + Red\ Snapper\ Season + \\ &\quad Year * Season. \end{aligned} \tag{2}$$

The final nominal and standardized abundance index for the western Gulf of Mexico (both provided as relative indices where each value is divided by the timeseries mean) along with confidence intervals and coefficients of variation (*CV*s) for the standardized index are given in Table 2.

The binomial model of proportion positive fit the data fairly well. In general the model tended to marginally underestimate the proportion positive early in the time series (prior to 2000) and marginally overestimate the proportion positive thereafter (Figure 1A). The frequency with which gray triggerfish are caught on a trip is relatively high (around 30-50%) during much of the timeseries, but tails off dramatically to around 15% over the last five years (Figure 1A). This may be associated with the forced conversion to circle hooks and implementation of the gray triggerfish rebuilding plan during this period. Residual analysis of the binomial model indicated no obvious patterns in the residuals by year (Figure 2A). Results from the lognormal model suggest a relatively strong fit and indicated that the assumption of a log normal distribution for positive catch was appropriate for the headboat west data (Figure 3A and 3B).

Nominal and standardized CPUE (scaled to the timeseries means) show similar trends for most of the timeseries with the nominal values within the 95% confidence intervals of the estimated values (Figure 6A). The first six years are characterized by a general increase in both nominal and standardized CPUE followed by sharp declines until around 2000. While the index did increase during the early to mid-2000's, it is currently near time series lows. Standardized CPUE has been below the mean level since 2009 and is currently at around one third of the timeseries mean.

Continuity

Trends were similar among the SEDAR 9, update, and SEDAR 43 indices (Figure 7A). Although attempts have been made to maintain modeling approaches, some alterations have been required since the SEDAR 9 benchmark due primarily to data updates causing changes in the significance of GLM factors. Vessel was included in the SEDAR 9 model, but due to lack of convergence, was dropped from the SEDAR 43 model and was assumed to have been dropped from the update assessment although no record of this could be found. Likewise, SEDAR 9 did not list Red Snapper Season as a significant factor for the lognormal model, but it was included in the current analysis. Another discrepancy among models was that during SEDAR 9, Hours Fished was included in the lognormal model. The variable Hours Fished is confounded with the response variable (mainly the effort term of angler-hours) and is inappropriate to use as an explanatory variable for the lognormal model. It was not included as a factor during the SEDAR 43 analysis.

SEDAR 9 Model (Nowlis, 2005):

$$Proportion\ Positive = Year + Vessel$$

$$\mathbf{ln}(\boldsymbol{CPUE}) = \boldsymbol{Year} + \boldsymbol{Vessel} + \boldsymbol{Season} + \boldsymbol{Hours\ Fished} + \boldsymbol{Year} * \boldsymbol{Season} + \boldsymbol{Year} * \boldsymbol{Vessel}\ . \qquad \textbf{(3)}$$

Results from the 2011 update assessment are poorly documented in regards to the standardization of CPUE indices and the factors used for the standardization could not be found. However, because of the slight discrepancies between the SEDAR 9 and 2011 update indices, it is unlikely that the same model was used. As stated above it is likely that Vessel was excluded and exploratory analyses indicated that Area was likely also excluded from the update analysis.

2011 Update Model

$$Proportion\ Positive = Unknown$$

$$\mathbf{ln}(\boldsymbol{CPUE}) = \boldsymbol{Unknown}\,. \qquad \textbf{(4)}$$

Eastern Gulf of Mexico

For the binomial component of the eastern Gulf of Mexico index the significant factors were Year and Area along with the Year*Area interaction term. Complications resulted from using the Vessel factor within the binomial model and it was dropped as an explanatory variable to ensure model convergence (see the Continuity Section below for a more detailed discussion). In the lognormal model Year, Vessel, and Season were found to be significant along with the Year*Season, Year*Vessel, and Season*Vessel interaction terms. Interactions involving vessel were excluded from the analysis due to model convergence issues. The final models were:

$$Proportion\ Positive = Year + Area + Year * Area$$

$$\ln(CPUE) = Year + Vessel + Season + Year * Season. \qquad \textbf{(5)}$$

The final nominal and standardized CPUE (both provided as relative indices where each value is divided by the timeseries mean) along with confidence intervals and coefficients of variation (*CV*s) are given in Table 3.

Observed trends in the proportion of positive trips demonstrates a fairly consistent variation between 40-60% with a sharp decline to around 25% after 2008. The model consistently overestimated proportion positive, in some cases substantially so, for all but the last few years of the timeseries (Figure 1B). Results from the lognormal model suggest a decent fit to the data and indicated that the assumption of a log normal distribution for positive catch was appropriate for the headboat east data (Figures 4A and 4B).

Nominal and standardized CPUE (scaled to the timeseries means) show similar trends, but rates of increase or decrease can vary substantially. However, the nominal values still fall within the 95% confidence intervals for most of the timeseries (Figure 6B). The first five years are characterized by a general increase in both nominal and standardized CPUE followed by a decline until around 2000. This was followed by a period of moderate increases until the mid-2000s, but CPUE has fallen sharply to timeseries lows since 2006. Standardized CPUE has been below the mean level since 2006 and is currently at around one half of the timeseries mean.

Continuity

The SEDAR 9 index generally declined throughout the timeseries and the SEDAR 43 index closely matched these results (Figure 7B). On the other hand, the update index tracked the nominal index closely and did not demonstrate any of the strong peaks in the early 1990s unlike the SEDAR 9 and SEDAR 43 indices (Figure 7B). Results from the 2011 update assessment are poorly documented in regards to the standardization of CPUE indices and the factors used for the standardization could not be found.

SEDAR 9 Model (Nowlis, 2005):

$$\begin{gathered} Proportion\ Positive = Year + Vessel \\ \ln(CPUE) = Year + Vessel + Season + Year * Season + Year * Vessel\ . \end{gathered} \quad \textbf{(6)}$$

For the purposes of this assessment, a continuity index that followed the methods of the update index was required. Details on the model used during the update assessment were not available for this analysis, requiring the use of an exploratory approach in order to create the continuity index. The Vessel and Area factors were included in SEDAR 9 and SEDAR 43, respectively, and were the factors that explained the largest amount of variation in their respective models. Removal of these two variables from the list of candidate variables resulted in the selection of Year, Hours Fished, and Season for the binomial model and only Year for the lognormal model. The selection of only Year in the lognormal model forced the resulting index to closely match the nominal index (Figures 6C and 7B).

SEDAR 43 Continuity Model:

$$\begin{gathered} Proportion\ Positive = Year + Hours\ Fished + Season \\ \ln(CPUE) = Year\ . \end{gathered} \quad \textbf{(7)}$$

The binomial model of proportion positive fit the data very well. In general the model tended to marginally underestimate the proportion positive throughout the time series (Figure 1C). Residual analysis of the binomial model indicated no obvious patterns in the residuals by year (Figure 2C). Results from the lognormal model suggest a decent fit to the data and indicated that the assumption of a log normal distribution for positive catch was appropriate for the headboat east data (Figures 5A and 5B).

Nominal and standardized CPUE for the SEDAR 43 continuity index along with confidence intervals and coefficients of variation (*CVs*) are given in Table 4. Nominal and standardized CPUE are nearly identical in the continuity run with the standardized values showing very tight confidence limits (Figure 6C).

Given that the update model closely resembled the nominal CPUE, it is unlikely that the update index followed the same protocol as SEDAR 9. It is probably that the update followed similar procedures as the SEDAR 43 continuity model with only Year included as a factor in the lognormal portion of the model.

DISCUSSION

The results and model diagnostics indicate that the headboat data can be used to develop reliable indices of CPUE that can be used in the SEDAR 43 assessment of gray triggerfish. There are a few areas of concern that warrant future research, the most important of which is the impact of the 2008 regulatory changes. Because management period would likely be confounded with the Year effect, it would be difficult to use a regulatory variable as an explanatory factor in the standardization. However, future investigations should consider splitting the series in 2008 as more years of data become available or dealing with these changes directly in the stock assessment model by allowing time-varying catchability or fishery selectivity. The convergence issues caused by the Vessel terms also needs to be further investigated. Finally, the gray triggerfish bag limit enacted in 2013 is likely to impact future catch rates and should be more carefully considered within the GLM in future updates. Due to the discrepancies between the 2011 update model and the SEDAR 9 and SEDAR 43 models along with the lack of documentation of how the update model results were obtained, it is suggested that care be taken when choosing whether to use the SEDAR 43 or SEDAR 43 continuity model for stock assessment purposes.

LITERATURE CITED

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

Nowlis, J.S. 2005. Updated fishery-dependent indices of abundance for the Gulf of Mexico gray triggerfish (*Balistes capriscus*). SEDAR 9-AW07. SFD-2005-031.

SAS Institute Inc. 2008. *SAS/STAT 9.2 User's Guide*. Cary, NC: SAS Institute Inc.

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

TABLES

Table 1: Levels and values of the factors investigated in the GLM

Factor	Levels	Values
Year	28	1986 - 2013
Season	4	Jan-Mar, Apr-Jun, Jul-Sep, Oct-Dec
Red Snapper Season	2	Closed, Open
Day/Night	2	Day, Mixed
Trip Duration	3	Half Day, Full Day, Multi Day
Hours Fished	9	5, 7, 10, 18, 24, 36, 48 ,60, 72

Table 2: SEDAR 43 index for the headboat dataset in the western Gulf of Mexico along with confidence intervals and CVs for the model estimates.

	Index		Confidence Limits (95%)		
Year	Standardized	Nominal	Lower	Upper	CV
1986	0.85	0.97	0.44	1.61	0.33
1987	0.82	0.88	0.44	1.54	0.32
1988	1.21	1.50	0.67	2.19	0.30
1989	1.47	1.44	0.81	2.65	0.30
1990	1.75	1.61	1.00	3.07	0.29
1991	2.96	2.35	1.84	4.78	0.24
1992	2.22	2.04	1.30	3.80	0.27
1993	2.05	1.82	1.19	3.54	0.28
1994	2.04	1.88	1.18	3.52	0.28
1995	1.61	1.53	0.93	2.80	0.28
1996	1.84	1.70	1.05	3.24	0.29
1997	1.24	1.11	0.65	2.35	0.33
1998	0.83	0.82	0.45	1.56	0.32
1999	0.57	0.54	0.29	1.12	0.35
2000	0.32	0.44	0.16	0.65	0.36
2001	0.45	0.58	0.23	0.86	0.34
2002	0.59	0.68	0.31	1.13	0.34
2003	0.75	0.80	0.41	1.38	0.31
2004	0.96	0.96	0.54	1.73	0.30
2005	0.93	0.87	0.54	1.61	0.28
2006	0.75	0.90	0.43	1.32	0.29
2007	1.00	1.07	0.56	1.77	0.29
2008	0.51	1.10	0.27	0.98	0.33
2009	0.08	0.11	0.04	0.16	0.38
2010	0.03	0.04	0.02	0.07	0.40
2011	0.06	0.10	0.03	0.12	0.38
2012	0.07	0.11	0.03	0.17	0.45
2013	0.02	0.04	0.01	0.05	0.47

Table 3: SEDAR 43 index for the headboat dataset in the eastern Gulf of Mexico along with confidence intervals and CVs for the model estimates.

	Index		Confidence Limits (95%)		
Year	Standardized	Nominal	Lower	Upper	CV
1986	0.73	0.31	1.21	0.01	0.26
1987	0.66	0.38	1.15	0.01	0.28
1988	0.73	0.78	1.27	0.01	0.29
1989	1.73	1.21	2.57	0.03	0.20
1990	2.31	2.09	3.39	0.04	0.19
1991	1.97	1.20	2.83	0.03	0.18
1992	2.31	1.59	3.29	0.04	0.18
1993	1.67	1.20	2.58	0.03	0.22
1994	1.25	1.26	2.02	0.02	0.24
1995	1.21	1.36	2.06	0.02	0.27
1996	1.04	1.28	1.77	0.02	0.27
1997	1.13	1.21	1.83	0.02	0.25
1998	1.09	1.21	1.71	0.02	0.23
1999	1.12	1.24	1.66	0.02	0.20
2000	0.71	1.03	1.16	0.01	0.25
2001	0.71	1.41	1.24	0.01	0.29
2002	1.17	1.77	1.93	0.02	0.26
2003	1.10	1.71	1.84	0.02	0.26
2004	1.08	1.27	1.74	0.02	0.24
2005	1.20	1.15	1.81	0.02	0.21
2006	0.68	0.67	1.12	0.01	0.26
2007	0.75	0.53	1.24	0.01	0.25
2008	0.51	0.50	0.87	0.01	0.28
2009	0.24	0.29	0.45	0.00	0.32
2010	0.22	0.32	0.45	0.00	0.37
2011	0.27	0.38	0.51	0.00	0.33
2012	0.23	0.31	0.46	0.00	0.36
2013	0.19	0.35	0.44	0.00	0.42

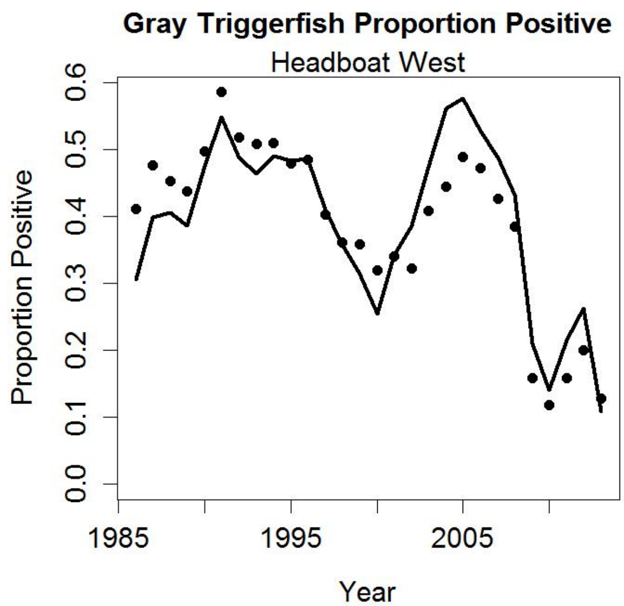
Table 4: SEDAR 43 continuity index for the headboat dataset in the eastern Gulf of Mexico along with confidence intervals and CVs for the model estimates.

	Index		Confidence Limits (95%)		
Year	Standardized	Nominal	Lower	Upper	CV
1986	0.33	0.31	0.25	0.43	0.14
1987	0.46	0.38	0.35	0.61	0.14
1988	0.79	0.78	0.67	0.94	0.08
1989	1.19	1.21	1.03	1.37	0.07
1990	1.53	2.09	1.37	1.72	0.06
1991	1.07	1.20	0.94	1.23	0.07
1992	1.36	1.59	1.21	1.54	0.06
1993	1.28	1.20	1.14	1.44	0.06
1994	1.13	1.26	0.98	1.30	0.07
1995	1.33	1.36	1.15	1.53	0.07
1996	1.18	1.28	1.01	1.36	0.07
1997	1.19	1.21	1.03	1.37	0.07
1998	1.17	1.21	1.02	1.34	0.07
1999	1.26	1.24	1.07	1.47	0.08
2000	1.06	1.03	0.90	1.24	0.08
2001	1.30	1.41	1.10	1.55	0.09
2002	1.89	1.77	1.60	2.23	0.08
2003	1.89	1.71	1.60	2.23	0.08
2004	1.32	1.27	1.12	1.56	0.08
2005	1.28	1.15	1.08	1.51	0.09
2006	0.80	0.67	0.66	0.96	0.09
2007	0.61	0.53	0.51	0.73	0.09
2008	0.55	0.50	0.46	0.64	0.08
2009	0.34	0.29	0.28	0.41	0.09
2010	0.38	0.32	0.30	0.48	0.12
2011	0.44	0.38	0.37	0.53	0.09
2012	0.46	0.31	0.35	0.61	0.14
2013	0.43	0.35	0.33	0.56	0.13

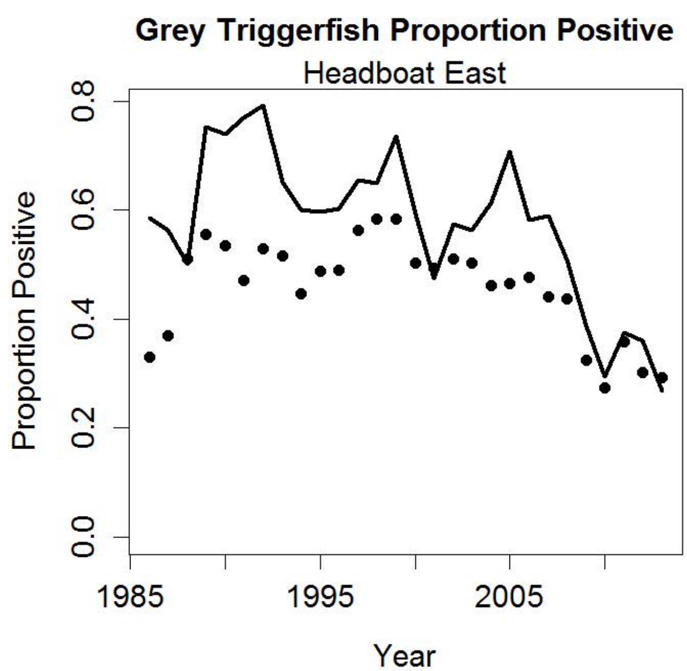
FIGURES

Figure 1: Binomial model results. Observed (black dots) and predicted (black line) proportion positive trips that caught the target species by year for the western Gulf of Mexico (A), the eastern Gulf of Mexico (B), and the continuity run for the eastern Gulf of Mexico (C).

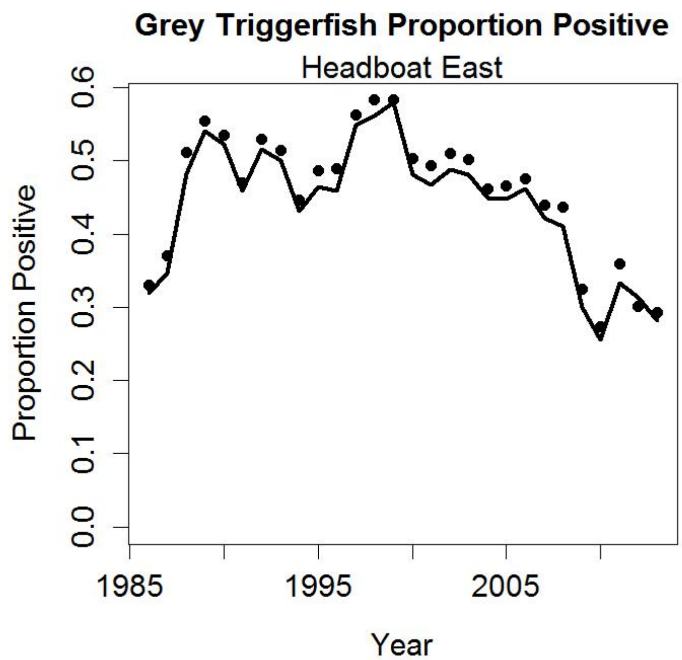
A.



B.



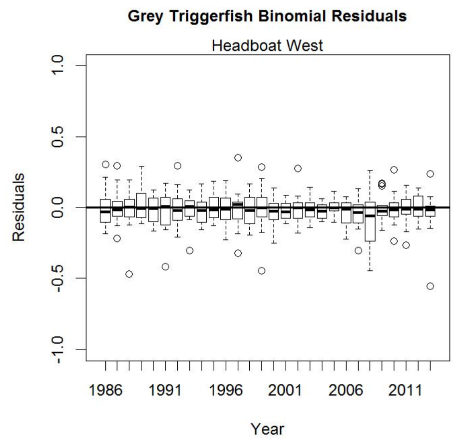
C.



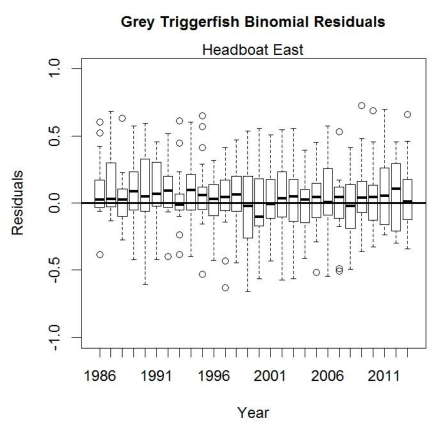
Grey Triggerfish Proportion Positive
Headboat East
Proportion Positive
0.0
0.1
0.2
0.3
0.4
0.5
0.6
1985
1995
2005
Year

Figure 2: Diagnostic plots for the binomial model: A) Box and Whisker plots of residuals by year for the western Gulf of Mexico, B) Box and Whisker plots of residuals by year for the eastern Gulf of Mexico, and C) Box and Whisker plots of residuals by year for the eastern Gulf of Mexico continuity run.

A.



B.



C.

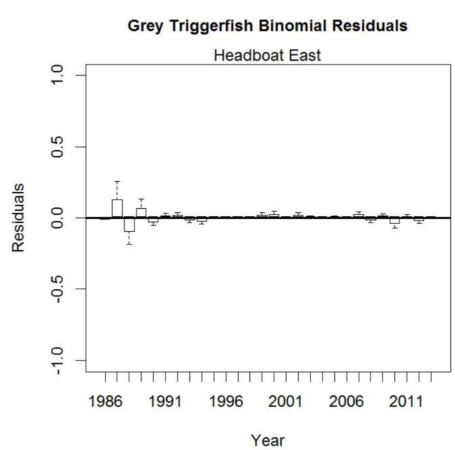
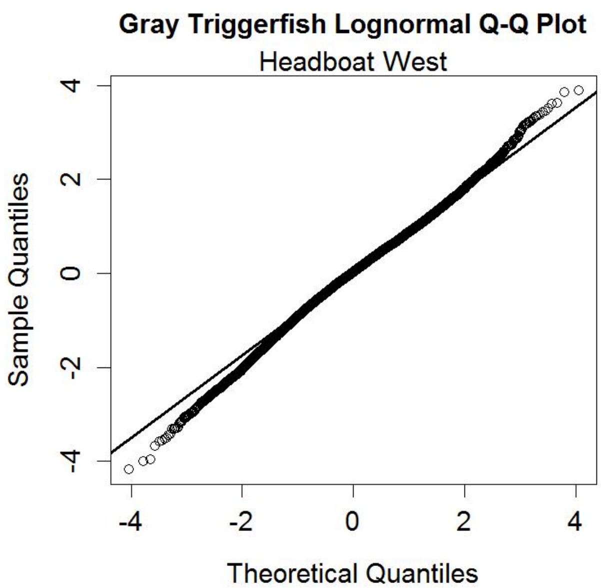


Figure 3: Diagnostic plots for the western Gulf of Mexico lognormal model: A) Q-Q plot, and B) Histogram of lognormal residuals plotted against normal distribution (red line).

A.



B.

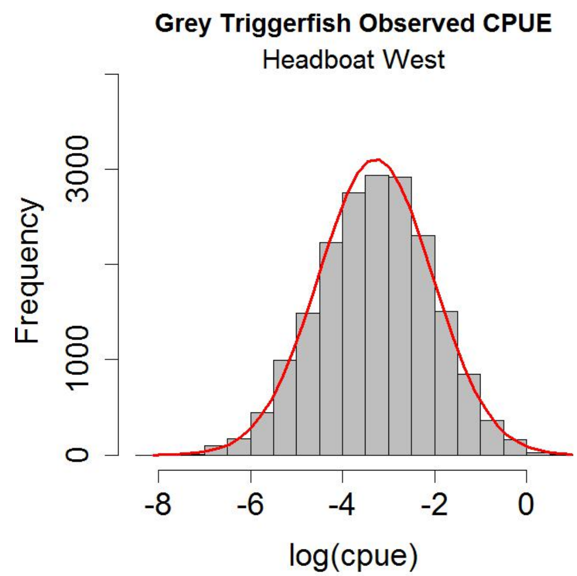
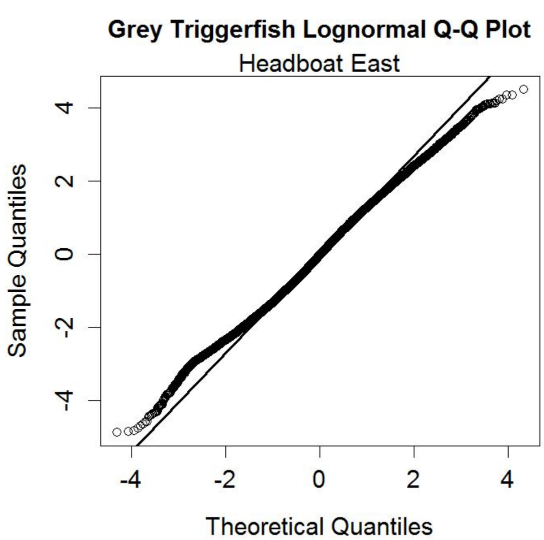


Figure 4: Diagnostic plots for the eastern Gulf of Mexico lognormal model: A) Q-Q plot, and B) Histogram of lognormal residuals plotted against normal distribution (red line).

A.



B.

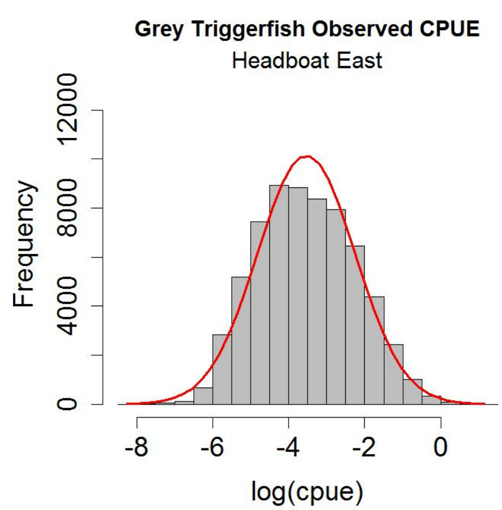


Figure 5: Diagnostic plots for the eastern Gulf of Mexico continuity run lognormal model: A) Q-Q plot, and B) Histogram of lognormal residuals plotted against normal distribution (red line).

A. **B.**

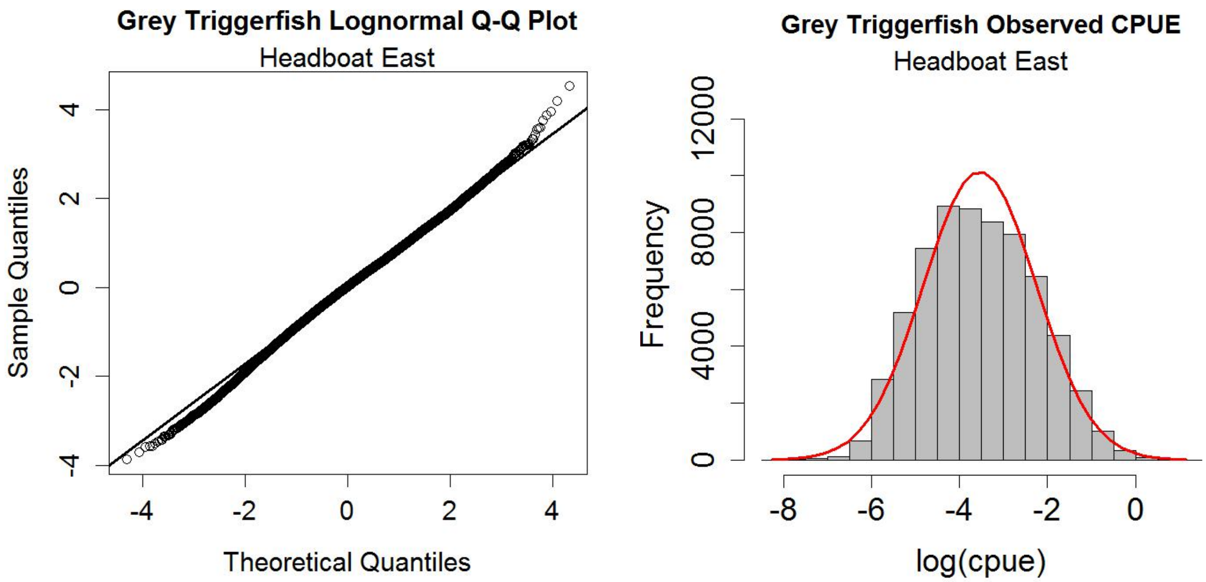
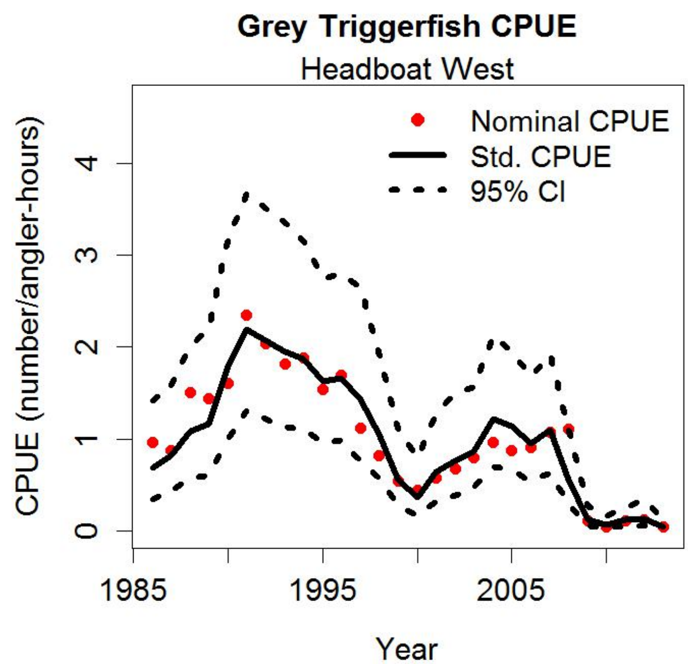
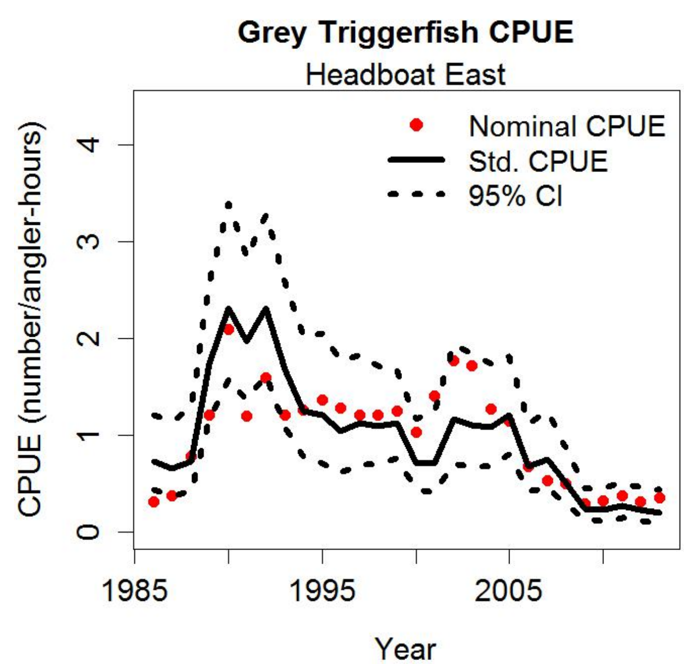


Figure 6: Timeseries plots of nominal (red dots) and standardized (black line) CPUE relative to the mean of the given timeseries for the western Gulf of Mexico (A), eastern Gulf of Mexico (B), and eastern Gulf of Mexico continuity run (C). 95% confidence intervals for the standardized CPUE are given by the dashed lines.

A.



B.



C.

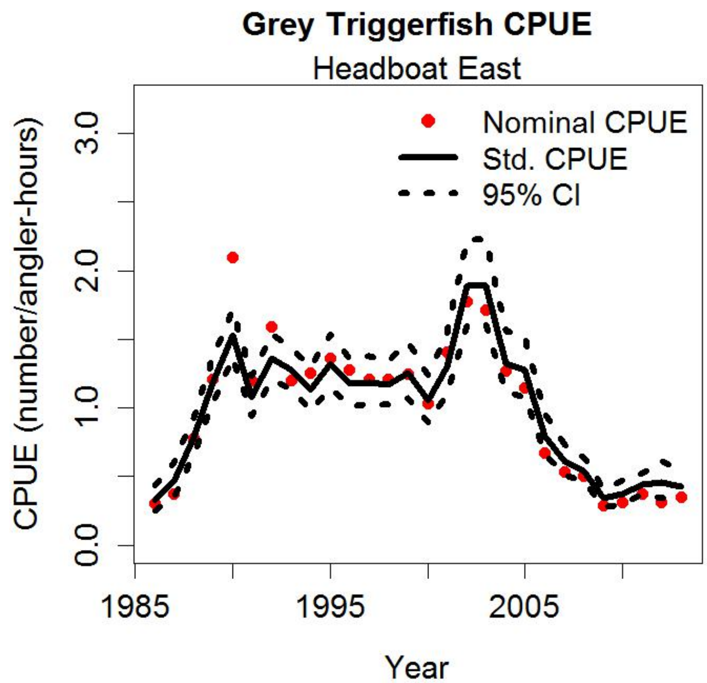
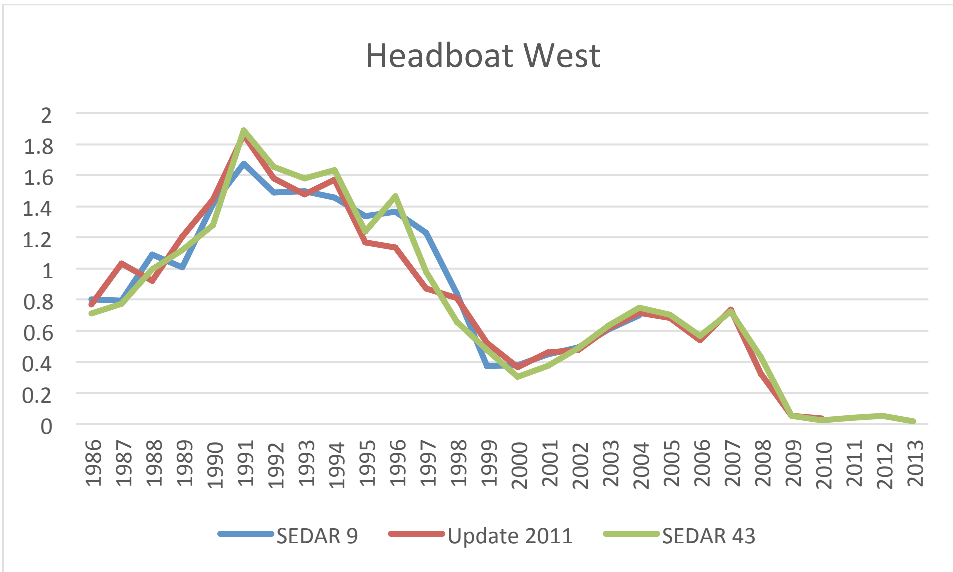

Grey Triggerfish CPUE
Headboat East
Nominal CPUE
Std. CPUE
95% CI
CPUE (number/angler-hours)
0.0
1.0
2.0
3.0
1985
1995
2005
Year

Figure 7: Timeseries plots of the standardized headboat CPUE indices for the western Gulf of Mexico (A) and eastern Gulf of Mexico (B). Shown are the indices for SEDAR 9 (dark blue line), 2011 update (red line), current SEDAR 43 (green line), and SEDAR 43 continuity run (purple line; eastern Gulf of Mexico only).

A.



B.