Standardized catch rates of greater amberjack (*Seriola dumerili*) in the southeast U.S. from headboat logbook data

Sustainable Fisheries Branch – NMFS Beaufort Lab
Contact: Eric Fitzpatrick

SEDAR59-WP03

Submitted: 6 December 2018

This information is distributed solely for the purpose of pre-dissemination peer review. It does not represent and should not be construed to represent any agency determination or policy.

Please cite this document as:

SFB-NMFS. 2018. Standardized catch rates of greater amberjack (*Seriola dumerili*) in the southeast U.S. from headboat logbook data. SEDAR59-WP03. SEDAR, North Charleston, SC. 40 pp.

Standardized catch rates of greater amberjack (Seriola dumerili) in the southeast U.S. from headboat logbook data

Sustainable Fisheries Branch[*]

November 2018

This document describes the the development of the SEDAR 59 headboat index for greater amberjack.

General EDA - Headboat effort

The effort patterns for the carolinas (CAR) and Georgia-North Florida (GNF) are fairly constant over time with most trips either being half- or full-day. The majority of trips in South Florida (SF) are half-day trips. The dip in the number of trips from the mid-1990s until the late 2000s in SF is attributed more to reporting than changes in effort (Figure 1). The positive greater amberjack trips are shown in figure 2.

Data Exclusions

1. Outlier removal

Extreme values occur more frequently in self-reported data because there are limited methods for validating data. Recent SEDAR stock assessments have removed values at the extreme upper tail of distribution for cpue and associated fields for self-reported fishery-dependent data. We excluded trips with the largest 0.5% values for catch in number (> 65) and cpue (> 0.333) for trips that caught greater amberjack. The number of anglers on a trip can also influence cpue when calculated as fish/angler-hour. Trips with the largest 0.5% values for reported anglers (> 107) were removed. Figure 3 shows the excluded cpue of excluded trips based on outlier definitions by region. Removing a small percentage of the trips with extreme values for variables used to calculate CPUE is an unbiased method to correct for potential errors in self-reported data.

2. Cutoff for number of trips per vessel and number of anglers

Logbooks submitted by vessels that participated infrequently in the fishery are likely to be less accurate and may add noise to the data. Even if a vessel fished infrequently for one year, the number of trips should be greater than 30. We removed vessels that had fewer than 30 trips in the logbook database. It is rare for a headboat to fish with few anglers. There is anecdotal information that headboats would sometimes fish with just the crew and that logbooks for these trips were submitted. Experienced crew are likely to be more efficient at catching fish than paying customers. Captains may also limit distance to reduce fuel costs for trips with few paying customers. Trips with 6 or fewer anglers were excluded.

3. Starting year

For SEDAR 15 the starting year for the headboat index was 1978. The headboat program increased it's range to south Florida by 1978 but the reporting appears to be in the burn-in phase, especially for full-day trips, until 1980. The number of trips reported may not be important if the subsample is unbiased. However, it takes time for vessel captains and crew to develop consistent and accurate reporting skills. This may be especially true for south Florida due to higher species diversity. The total number of reported trips in SF increases dramatically from 1978 - 1980 (Figure 1). The number of reported positive greater amberjack half-day trips in south Florida increases from 62 in 1978 to 250 in 1979 and then starts to drop in the late 1980s. The number of positive greater amberjack full-day trips were highest in the 1980s (Figure 2). Figure

[*]National Marine Fisheries Service, Southeast Fisheries Science Center, 101 Pivers Island Rd, Beaufort, NC 28516

16 suggests there were issues with reporting in the Georgia-north Flordia region in 1978 abd 1979. Due to these reporting issues 1980 was chosen as the start year.

4. Terminal year - spawning closure exclusion

The seasonal closures occured in 2016 and 2017. Comparisons of the median cpue by region for all months and May-Oct shows little difference in median cpue across regions (Figure 4). Removing trips from these months allows us to extend the headboat logbook index until the terminal year of the assessment (2017). The peaks in the number of positive greater amberjack trips are similar between the seasonal closure and open months prior to the 2010 seasonal closure by region (Figure 5). 2017 was chosen as the terminal year with the removal of all trips from November to April across all years.

5. Trip types

For SEDAR 15, the relatively few multi-day trips were combined with full-day trips and the 3/4-day trips were combined with the half-day trips. Figure 6 shows the noisy positive greater amberjack median cpue associated with these trip types for the Carolinas (CAR) and Georgia-North Florida (GNF) while South Florida(SF) is fairly stable. It is difficult to determine the number of hours spent fishing on multi-day trips which can add noise to the cpue. Combining trip types with either small sample size or high uncertainty with the more reliable half and full-day trips increases the noise in cpue (Figure 7). There are relatively few positive greater amberjack trips across all areas for 3/4-, and multi-day trips. Multi-day and 3/4-day trips were removed from all areas. Trip type was retained to calculate the CPUE unit of fish/angler-hour (where hour is defined by trip type), but was dropped as an explanatory variable. This has the added benefit of not using trip type both as a factor, and to calculate the response variable.

Nominal catch rates

Nominal catch rates of positive greater amberjack trips by year and region from the data as filtered for input into the Stephens and MacCall analysis are shown in figure 8.

Evaluation of explanatory variables

YEAR - Year was necessarily included, as standardized catch rates by year are the desired outcome. Years modeled were initially 1978-2017 as 1978 was the start year in SEDAR 15. Figure 16 suggests there were issues with reporting in the Georgia-north Flordia region in 1978 abd 1979. Due to these reporting issues 1980 was chosen as the start year.

AREA - The two areas modeled for the SEDAR 15 (north and south of Cape Canaveral, FL) headboat logbook index were changed for SEDAR 59. The three areas include the Carolinas (CAR), Georgia and North Florida (to Cape Canveral, FL), South Florida (South of Cape Canaveral, FL). These areas were defined due to shelf characteristics and associated fishing behavior as well as species compositions.

SEASON - Month was used as the within-year factor for SEDAR 15. For SEDAR 59, half of the months were dropped due to the seasonal closures in the last two years (Figure 9). The seasonal pattern in cpue across months seems consistent across areas (Figures 10 and 11). Season was chosen as the explanatory variable.

VESSEL SIZE (vsize) - For SEDAR 15, a factor was developed for the number of anglers using the quartiles of the number of anglers across all trips as breaks for the factors. Given the large range of vessel sizes, a trip with 20 anglers could be either almost full or almost empty. Here we develop a factor for vessel size and crowding separately using the number of anglers. The proxy for vessel size is the maximum anglers reported over all trips for a vessel (Figure 12). This was then divided into four factors based on visual inspection of the density plots into: 1. less than 30 maximum anglers (a.lt30), 2. 30-59 maximum anglers (b.30-59), 3. 60-89 maximum anglers (c.60-89), and 4. 90 or more maximum anglers (d.ge90),(Figure 13).

PERCENT FULL (pctfact) The number of anglers reported for a trip was divided by the maximum number of anglers for a vessel to obtain an estimate of crowding. This was then divided into 4 equally spaced factors;

1. less than 25% full (a.lt25), 25-49% full (b.25-49), 50-74% full (c.50-74), and 75% or more full(d.ge75). The density of percent full by area and the density of cpue associated with each factor are shown in figure 14.

Analytical decisions

1. Subsetting trips - Use Stephens and Maccall (2004) method
2. Species included in Stephens and MacCall approach: limit to snapper-grouper complex and remove species with full-year closures, ID issue, or large shifts in desirability over the index period
3. Apply Stephens and MacCall to Carolinas, Georgia-N.Florida, and S. Florida with Cape Canaveral, FL separating North and South Florida

Subsetting trips

Effective effort was based on those trips from areas where greater amberjack were available to be caught. Without fine-scale geographic information on fishing location, trips to be included in the analysis must be inferred, which was done here using the method of (Stephens and A. 2004). The method uses multiple logistic regression to estimate a probability for each trip that the focal species was caught, given other species caught on that trip. The method was applied separately for the three regions considered due to species composition shifts. A zoogeographic boundary is apparent near Cape Canaveral (Shertzer, Williams, and J. 2009) which is the break between GNF and SF areas. Another break between the CAR and GNF areas was included to limit the influence of species at the edge of their range (e.g. scup in the North or yellowtail snapper to the South). To avoid undue influence of rare species on regression estimates, species included in each analysis were limited to those occurring in 4% or more of trips. A range of 1% to 5% was considered with 1% including too many species and 5% too few, especially in SF. However, the cutoff had little influence on the trips selected because the species with the highest probabilities (positive and negative) were always included. We limited the species to the snapper-grouper species that were on the headboat logbook forms across all years included in the index. The species listed on logbook forms for the entire period differed by region (Table 1). Species with management closures were also omitted because the potential for erroneously removing trips likely to have caught greater amberjack during years of restrictions (Table 2). Gray triggerfish is a species that was less desirable during the early part of the survey and therefore may not be good indicator of greater amberjack habitat. Figure 15, from a headboat data quality study, shows the trend in reporting where trip reports could be matched to port samplers measurements for gray triggerfish. The increase in gray triggerfish in the early 1990s is associated the increase in desirability of gray triggerfish described by headboat Captains. This study was limited mostly to full-day trips and ends in 2013. The same analysis showed consistent reporting for greater amberjack over time with the peaks in fish missed, as determined by fish lengths taken by port samplers, to fish reported (Figure 16).

A backwards stepwise AIC procedure (Venables and Ripley 1997) was then used to perform further selection among possible species as predictor variables, where the most general model included all listed species as main effects. In this procedure, a generalized linear model with Bernoulli response was used to relate presence/absence of greater amberjack in each trip to presence/absence of other species. Regression coefficients of included species for all areas are given in Appendix 1 and shown in figure 17. A trip was then included if its associated probability of catching greater amberjack was higher than a threshold probability (Figure 17). The threshold was designed to be that which resulted in the same number of predicted and observed positive trips, as suggested by Stephens and MacCall (2004). Retention of positive and zero greater amberjack trips across factors are shown in Figures 18 - 20. The nominal catch rate after the subsetting for SEDAR 19 and SEDAR 53 are fairly similar (Figure 21)

Standardization

CPUE was modeled using the delta-GLM approach (Lo, L., and J. 1992; Dick 2004; Maunder and Punt 2004). This approach combines two separate generalized linear models (GLMs), one to describe presence/absence of the focal species, and one to describe catch rates of successful trips (trips that caught the focal species). Estimates of variance were based on 1000 bootstrap runs where trips were chosen randomly with replacement (Efron and Tibshirani 1993). All analyses were programmed in R, with much of the code adapted from Dick (2004).

Bernoulli submodel

The bernoulli component of the delta-GLM is a logistic regression model that attempts to explain the probability of either catching or not catching greater amberjack on any given trip. Initially, all explanatory variables were included in the model as main effects, and then stepwise AIC (Venables and Ripley 1997) with a backwards selection algorithm was used to eliminate those variables that did not improve model fit. In this case, the stepwise AIC procedure did not remove any explanatory variables. Diagnostics, based on Pearson residuals, suggested reasonable fits of the Bernoulli submodel (Figure 22).

Positive CPUE submodel

Two parametric distributions were considered for modeling positive values of CPUE, lognormal and gamma. For both distributions, all explanatory variables were initially included as main effects, and then stepwise AIC (Venables and Ripley 1997) with a backwards selection algorithm was used to eliminate those variables that did not improve model fit. For both lognormal and gamma distributions, the best model fit included all explanatory variables. The two distributions, each with their best set of explanatory variables (all of them), were compared using AIC. Lognormal outperformed gamma, and was therefore applied in the final delta-GLM. Diagnostics suggested reasonable fits of the lognormal submodel (Figures 23 and 24).

Results

The standardized index was similar to the nominal index with the exception of a few years associated with peaks in the catch rate (Figure 25). Assessment methods that account for changes in catchability could be implemented over time periods where effort may have been influenced by management measures.

Table 1: Species listed on headboat logbook forms in 1980 for North and South Carolina (CAR) and Georgia - Florida (GFL) which are in the snapper-grouper complex.

CAR	GFL
Gag	Almaco.jack
Greater.amberjack	Blackfin.snapper
Knobbed.porgy	Blue.runner
Red.Grouper	Blueline.Tilefish
Red.Hind	Bluestriped.grunt
Red.porgy	Cubera.snapper
Red.snapper	Gag
Rock.Hind	Gray.snapper
Scamp	Gray.triggerfish
Scup	Graysby
Silk.snapper	Greater.amberjack
Snowy.Grouper	Jolthead.por gy
Tomtate	Knobbed.porgy
Vermilion.snapper	Lane.snapper
Warsaw.Grouper	Mutton.snapper
White.grunt	Queen.triggerfish
Whitebone.porgy	Red.Grouper
Yellowfin.Grouper	Red.Hind
	Red.porgy
	Red.snapper
	Rock.Hind
	Sand.tilefish
	Scamp
	Silk.snapper
	Tomtate
	Vermilion.snapper
	White.grunt
	Whitebone.porgy
	Yellowfin.Grouper
	Yellowmouth.Grouper

Table 2: Species removed from the Stephens and MacCall method for defining greater amberjack trips due to seasonal or complete closures or ad-hoc evidence of shifts in desireability.

Species.removed
Lesser.amberjack
Almaco.jack
Banded.rudderfish
Red.porgy
Gray.triggerfish
Red.snapper
Vermilion.snapper
Mutton.snapper
Snowy.Grouper
Gag
Black.sea.bass
Blueline.tilefish

Table 3: Nominal and standardized CPUE for greater amberjack 1978-2017 with CVs for stardardized index of abundance.

Year	N	Nominal.CPUE	Relative.nominal	Standardized.CPUE	CV
1978	472	0.01	0.78	0.57	0.10
1979	534	0.00	0.35	0.56	0.11
1980	648	0.01	1.03	0.87	0.10
1981	606	0.00	0.55	0.55	0.12
1982	833	0.01	0.78	0.77	0.07
1983	1094	0.01	0.83	0.82	0.07
1984	862	0.01	0.70	0.64	0.08
1985	877	0.01	0.88	0.91	0.07
1986	1284	0.01	1.13	0.99	0.07
1987	1226	0.01	1.28	1.35	0.06
1988	921	0.01	0.75	0.80	0.08
1989	587	0.00	0.56	0.50	0.11
1990	582	0.01	0.81	0.83	0.09
1991	671	0.01	0.89	0.82	0.09
1992	1073	0.01	0.80	0.78	0.07
1993	1006	0.01	0.94	0.83	0.07
1994	835	0.01	0.85	0.62	0.08
1995	832	0.01	0.74	0.69	0.08
1996	816	0.01	1.14	0.88	0.08
1997	949	0.00	0.57	0.49	0.08
1998	842	0.00	0.63	0.56	0.08
1999	775	0.01	1.29	0.99	0.08
2000	686	0.01	1.41	1.34	0.07
2001	636	0.01	1.20	1.06	0.08
2002	540	0.01	1.48	1.26	0.08
2003	581	0.01	1.66	1.43	0.09
2004	558	0.01	1.06	1.06	0.08
2005	529	0.00	0.63	0.64	0.10
2006	506	0.01	0.73	0.59	0.11
2007	460	0.01	1.65	1.65	0.08
2008	350	0.01	1.22	1.33	0.11
2009	409	0.01	1.54	1.86	0.09
2010	523	0.01	1.11	1.05	0.09
2011	504	0.01	0.73	0.73	0.11
2012	437	0.01	0.70	0.68	0.13
2013	461	0.01	1.40	1.59	0.10
2014	341	0.01	1.13	1.51	0.10
2015	465	0.01	1.51	1.89	0.08
2016	460	0.01	1.61	2.21	0.08
2017	426	0.01	0.91	1.27	0.09

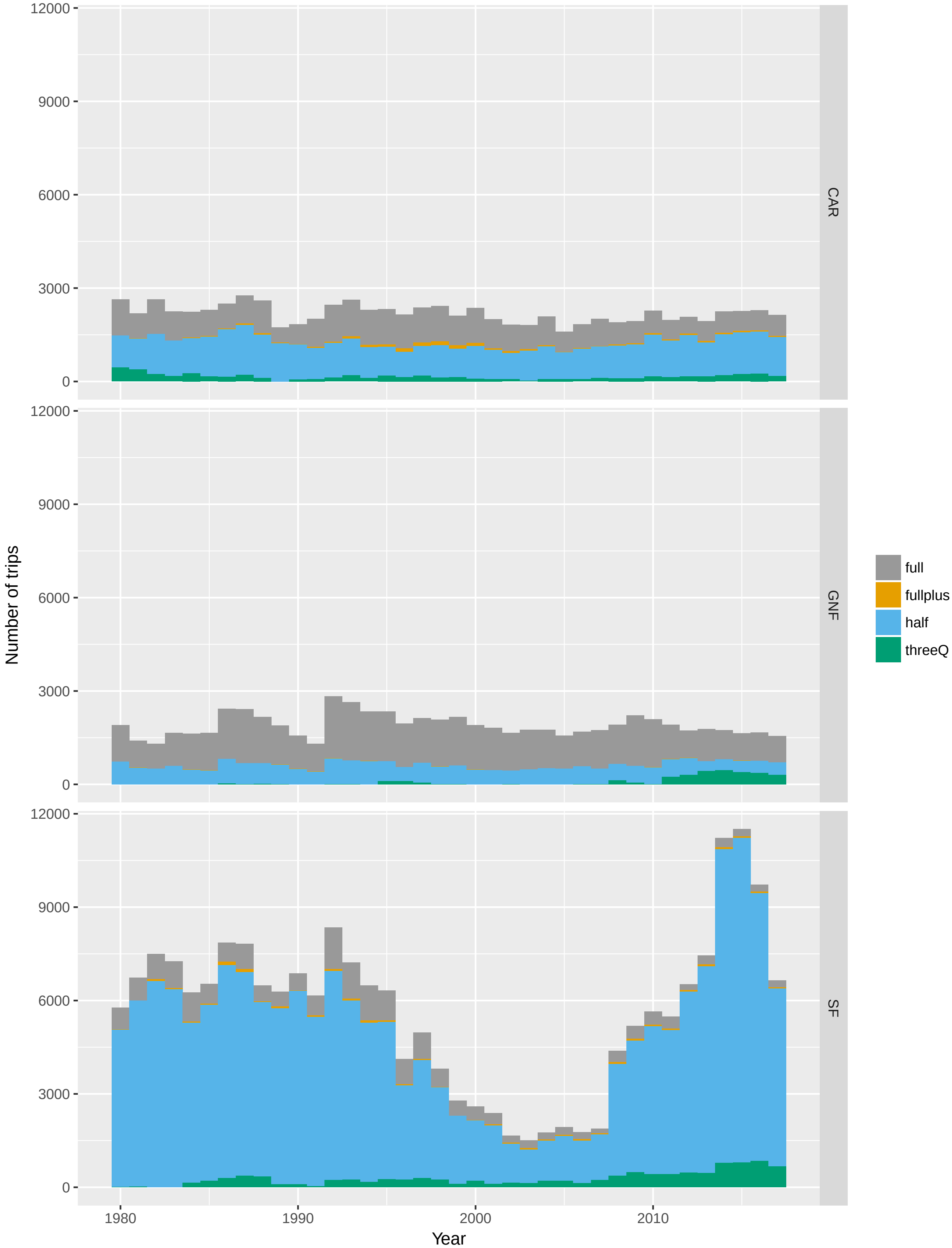


Figure 1: Number of headboat trips that submitted logbooks 1978-2017 for half-day trips (half), full-day trips (full), three-quarter day trips (threeQ), and multiple-day trips ranging from 1.5 to 7 days (fullplus) by region (CAR-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida).

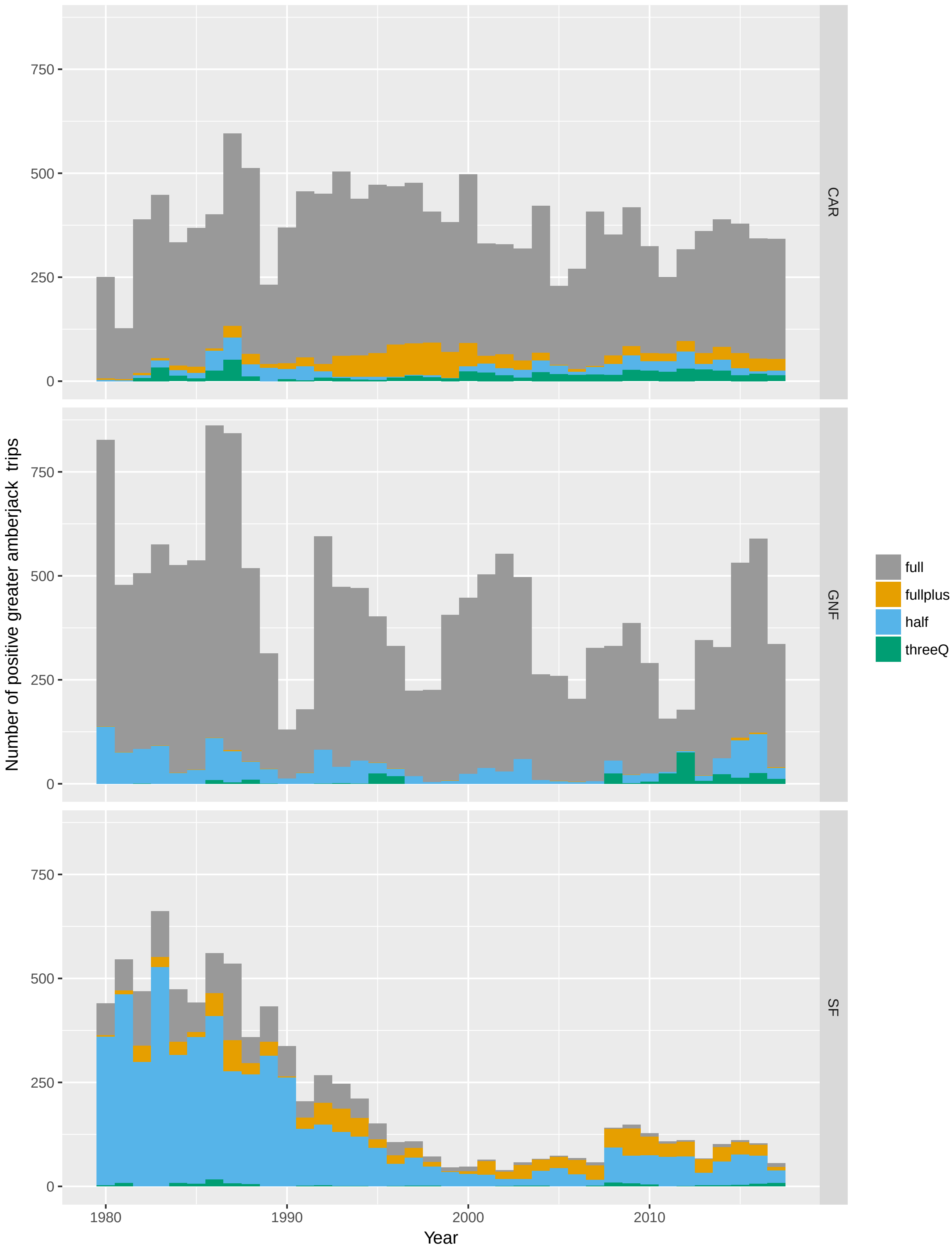


Figure 2: Number of positive greater amberjack headboat trips that submitted logbooks 1978-2017 for half-day trips (half), full-day trips (full), three-quarter day trips (threeQ), and multiple-day trips ranging from 1.5 to 7 days by region (CARr-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida).

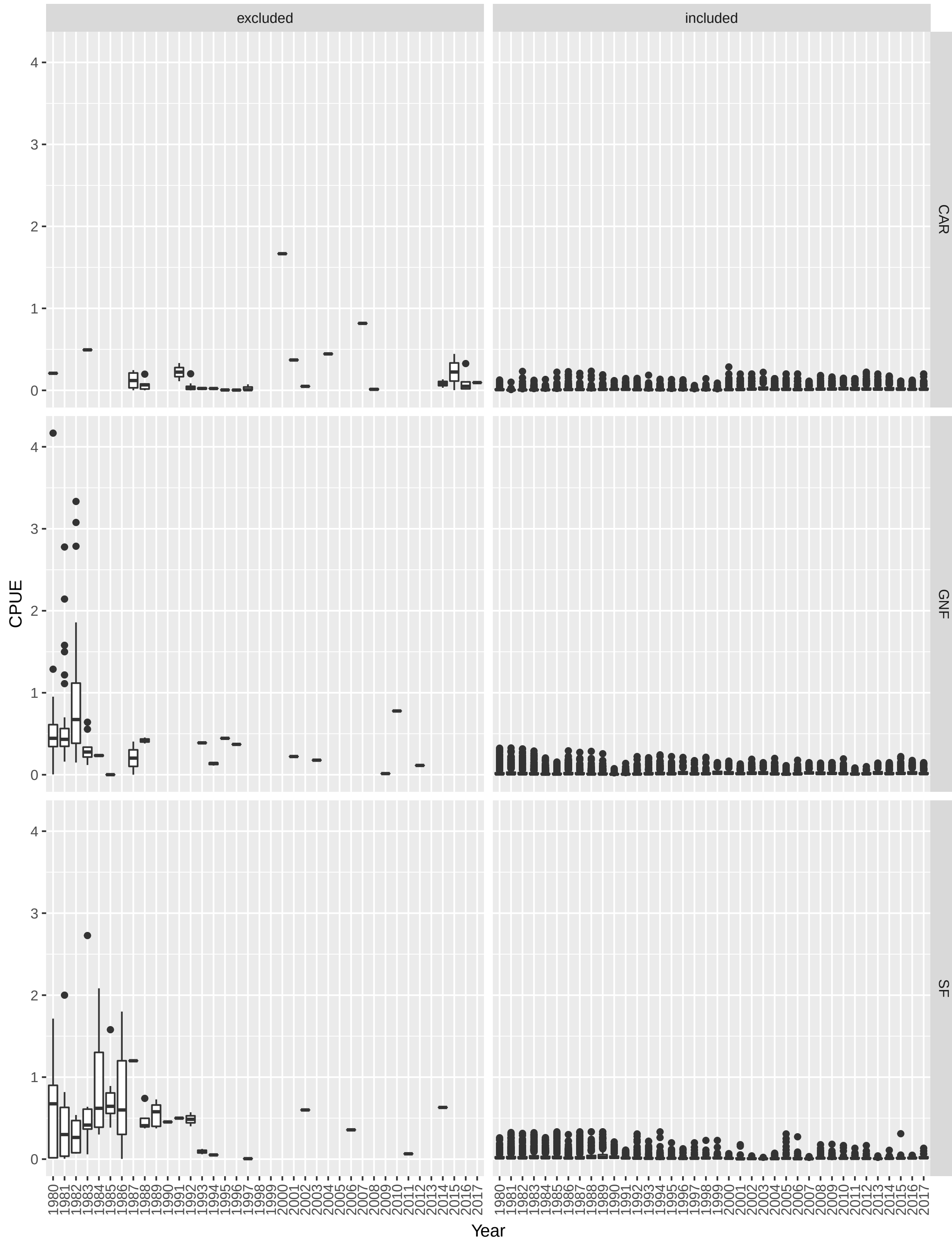


Figure 3: Records determined as outliers (excluded) based on removal of values above the 99.5th percentile for anglers, number of fish caught, and cpue.

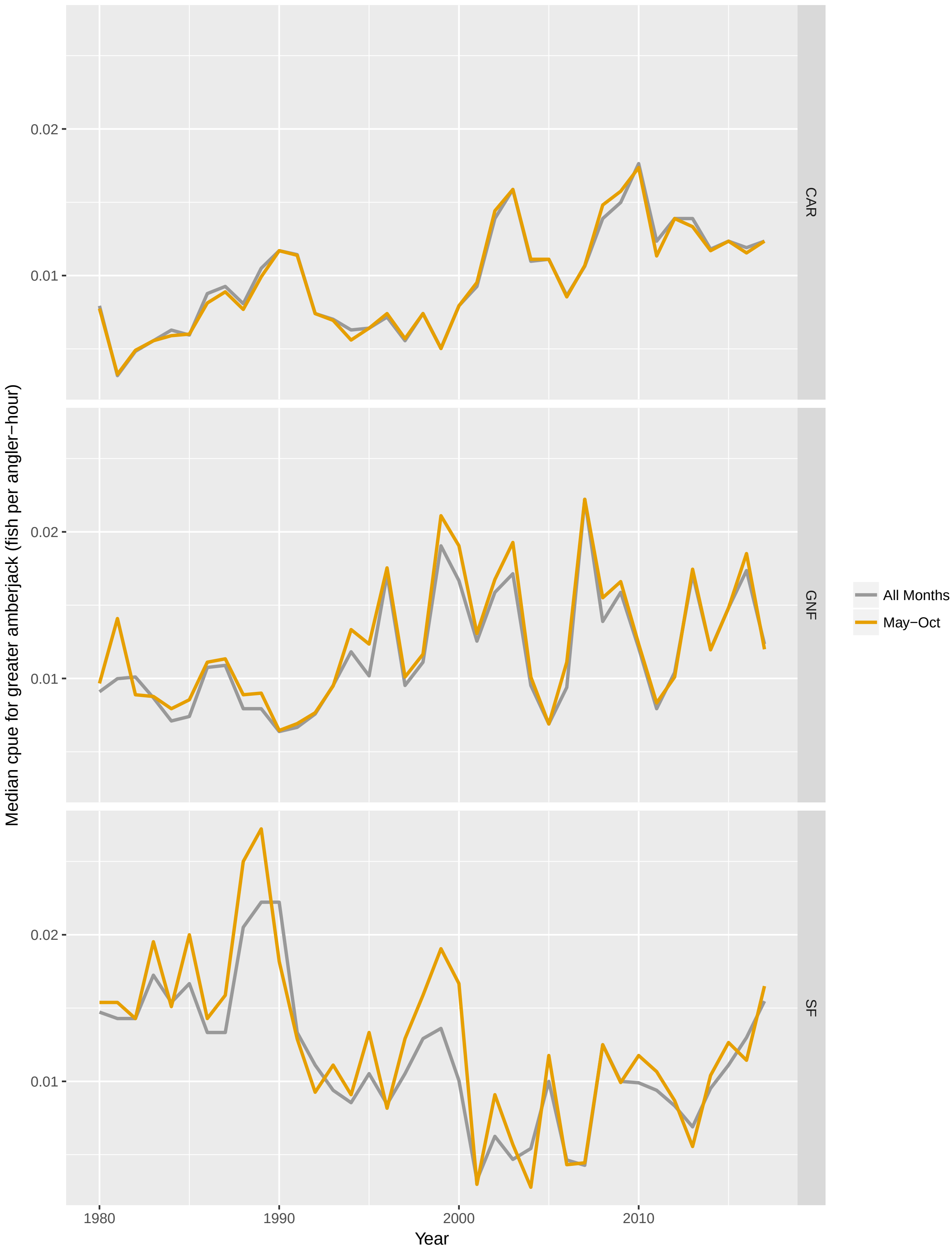


Figure 4: Median nominal greater amberjack catch rates by region (CAR-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida) for all months and just May-Oct.

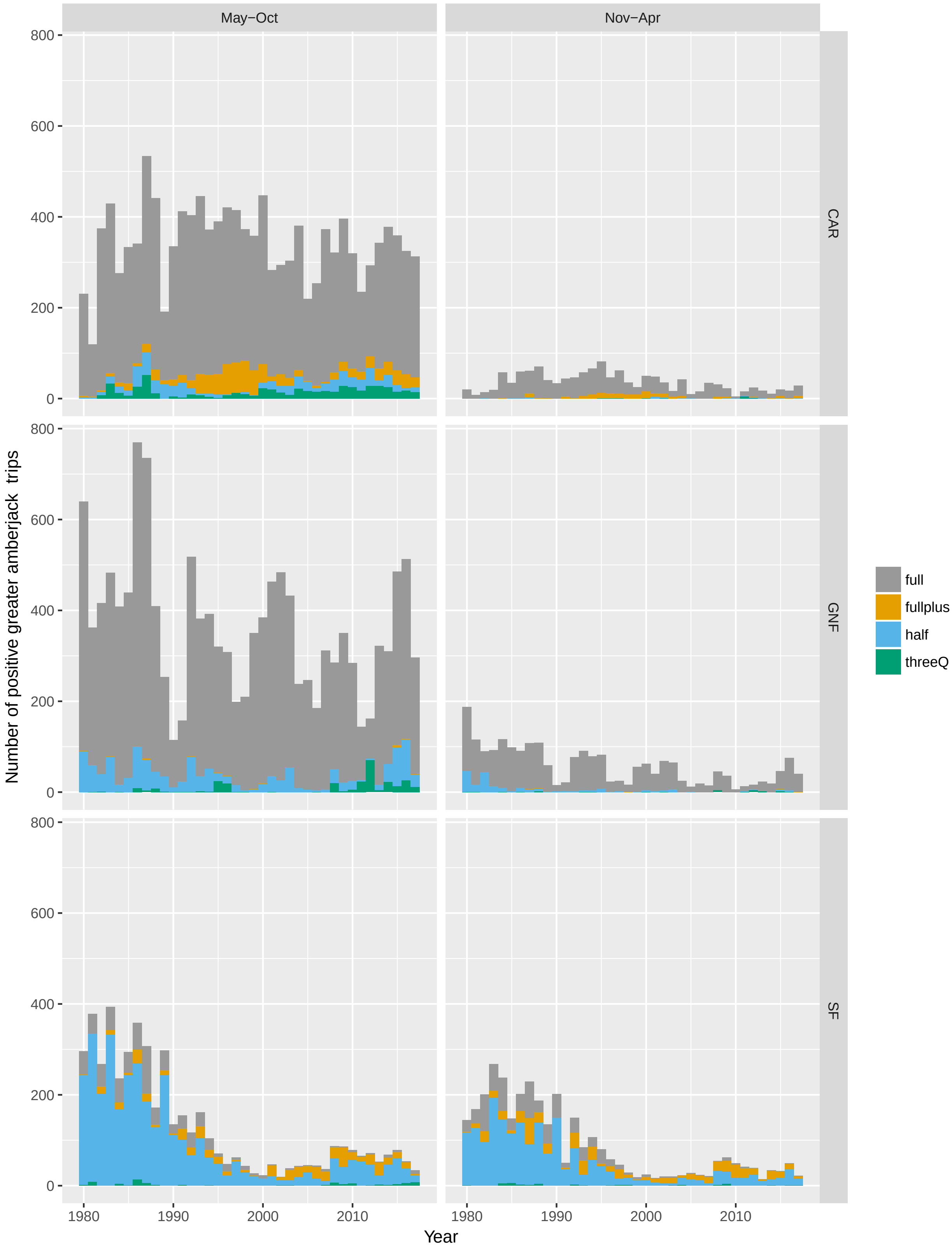


Figure 5: Positive greater amberjack trips by region (CAR-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida) and season. November to April is the closure season that began in 2016.

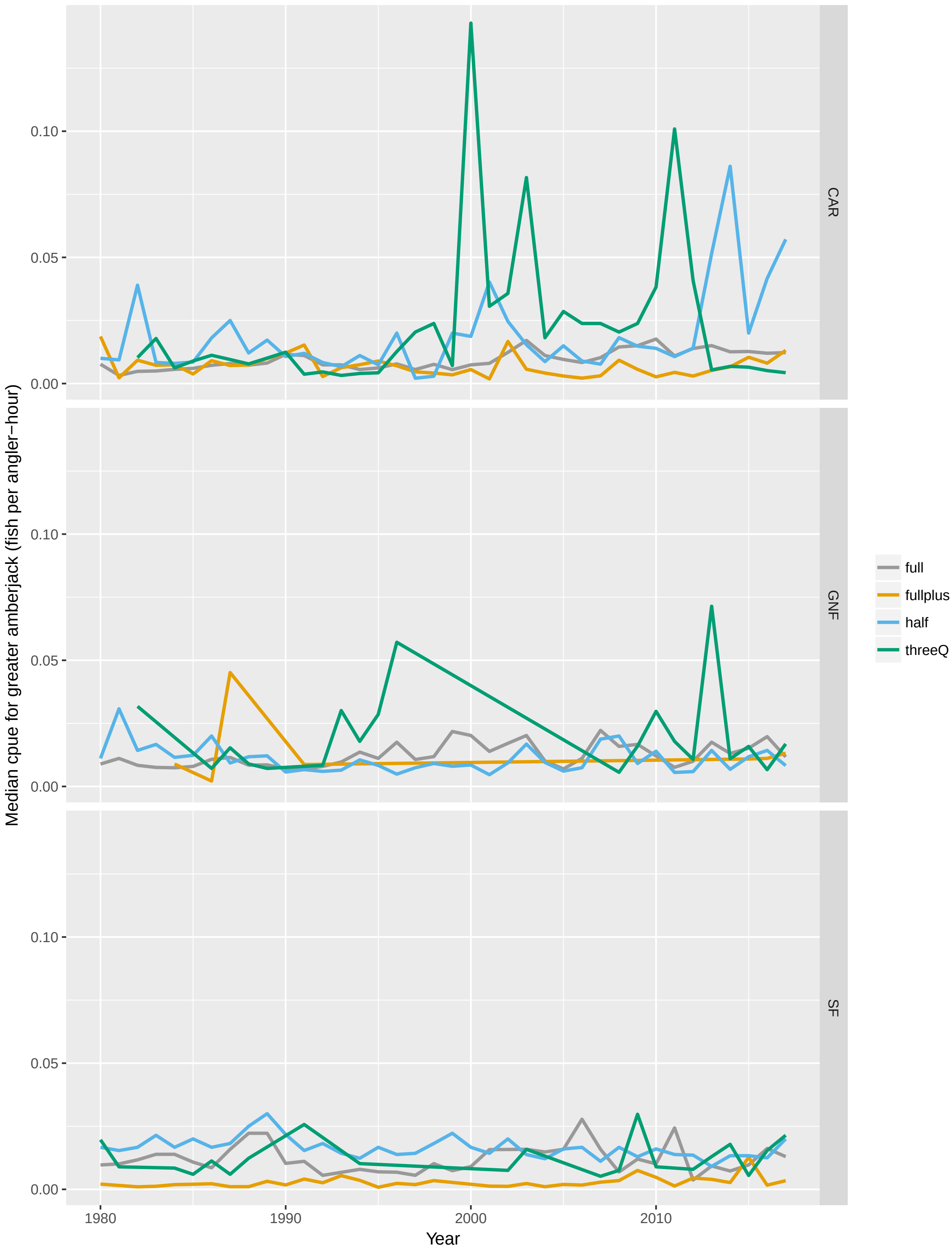


Figure 6: Greater amberjack cpue by region (CAR-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida) for May - Oct.

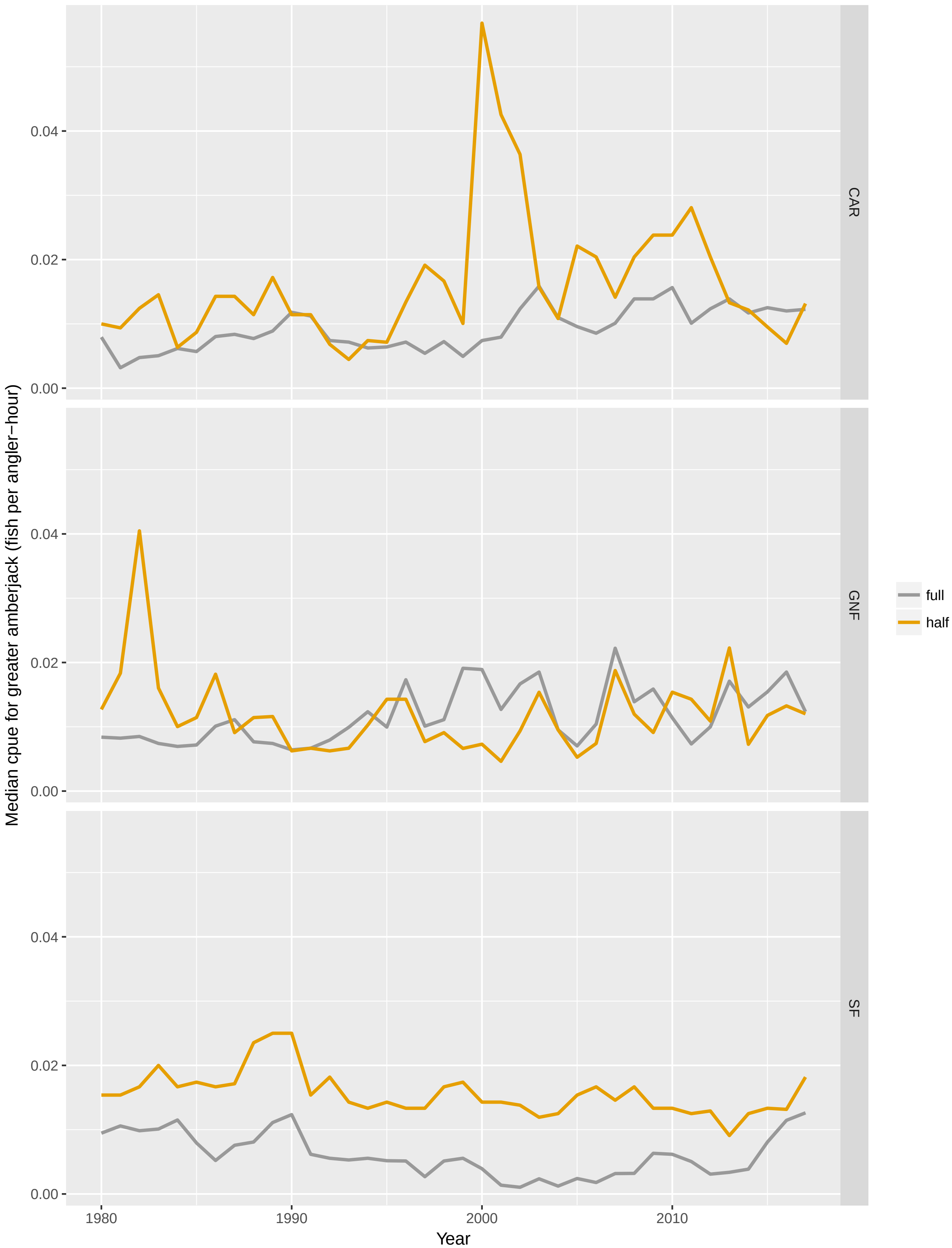


Figure 7: Greater amberjack cpue by region (CAR-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida) for May - Oct. Trips are aggregated into full- (includes multi-day trips) and half- (includes 3/4-day) trips. November to April is the closure season that began in 2016.

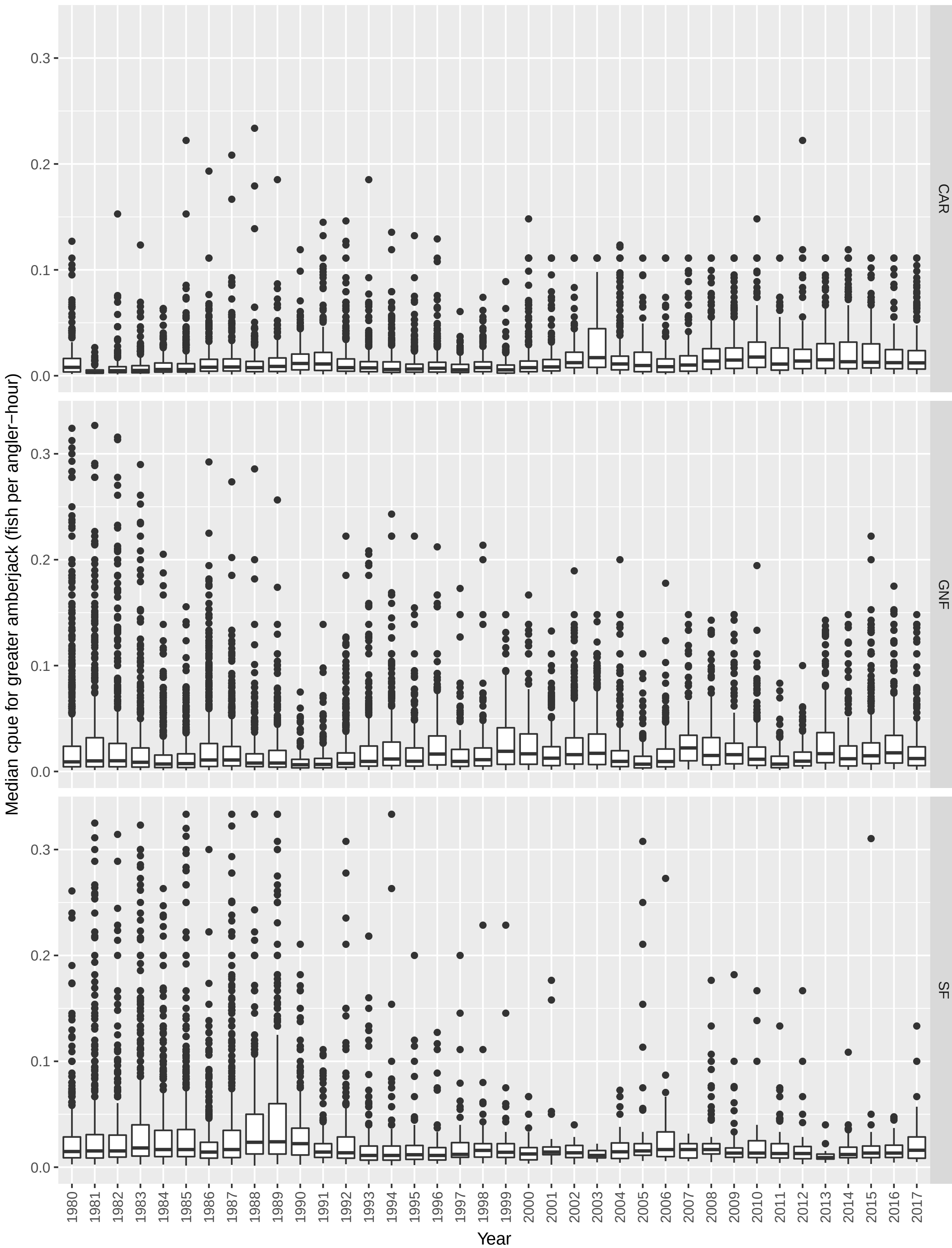


Figure 8: Greater amberjack cpue by region (CAR-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida) and season. Multi-day and 3/4-day trips are removed for all regions. Half-day trips are removed for CAR and GNF. Half-day trips and full-day trips are aggregated for SF. Years are limited to 1978-2017.

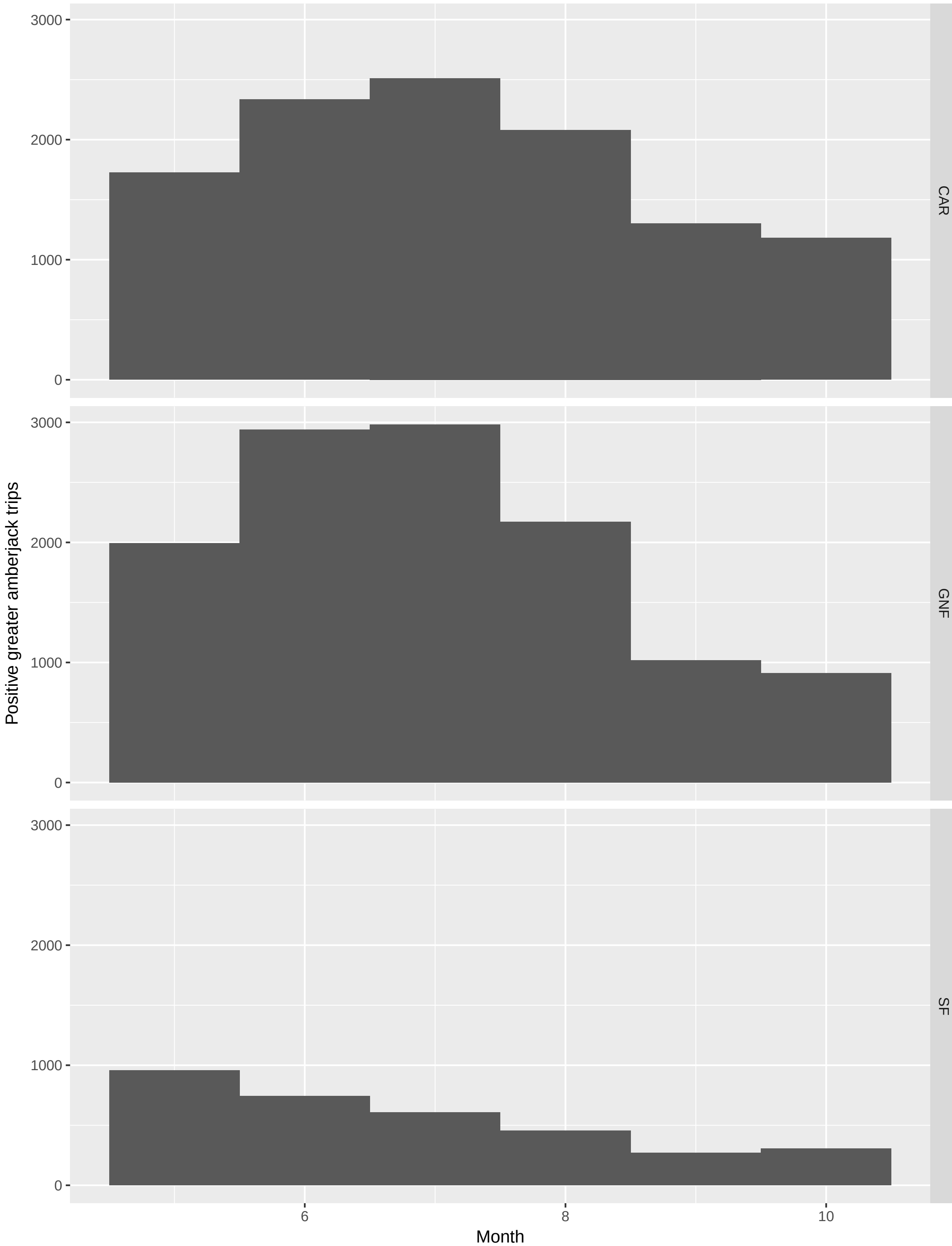


Figure 9: Positive greater amberjack trips by month and region (CAR-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida).

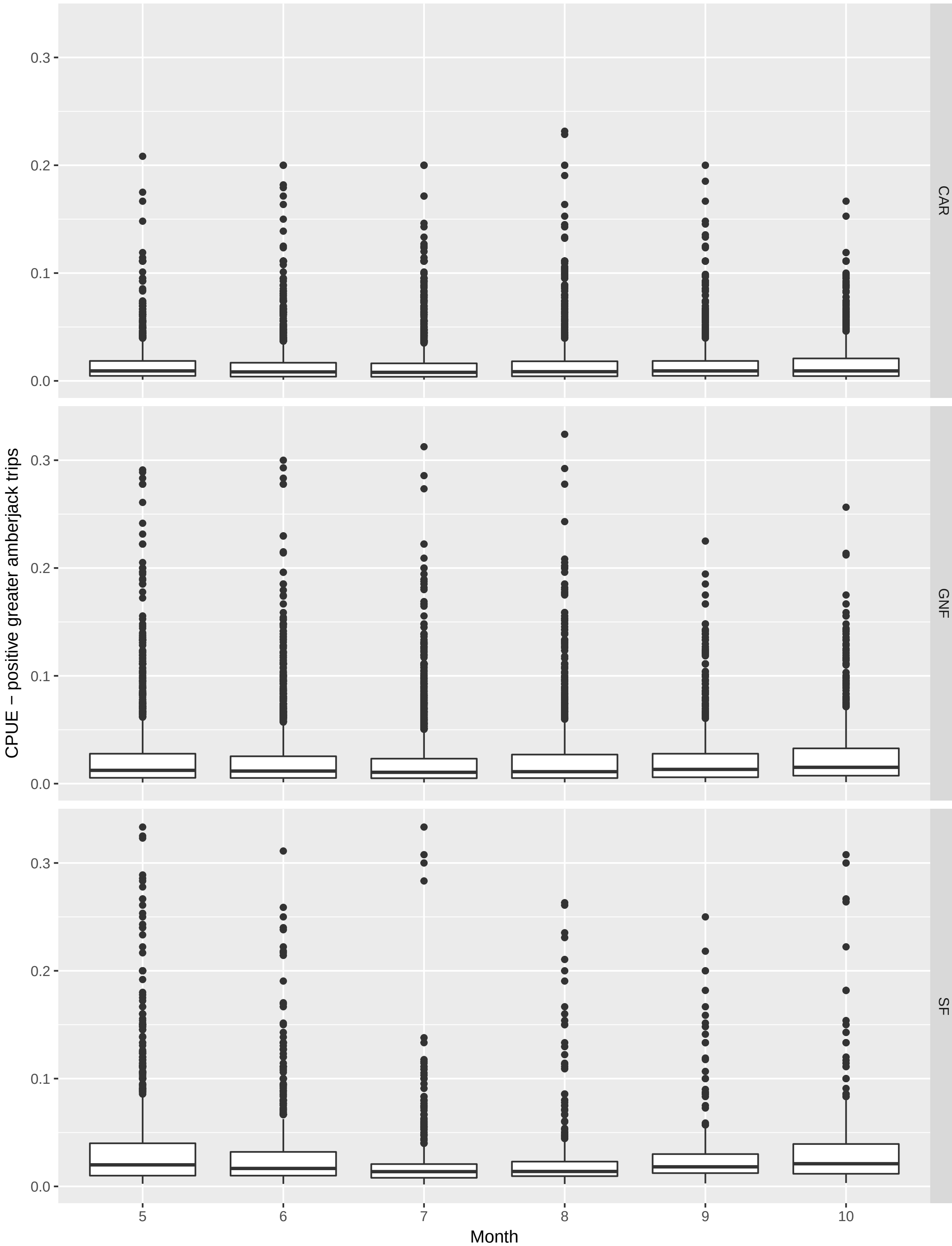


Figure 10: Greater amberjack cpue for positive trips by month and region (CAR-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida).

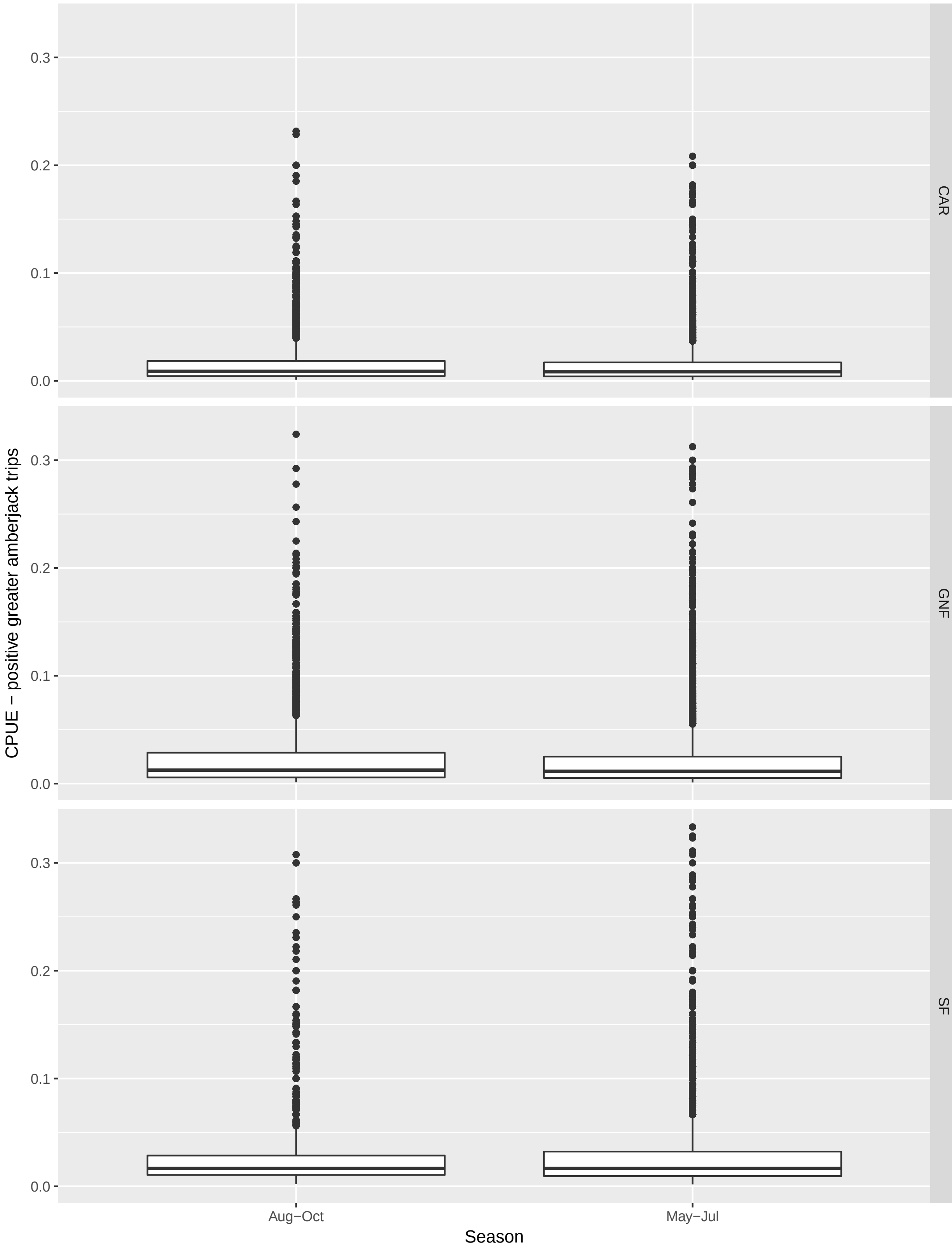


Figure 11: Greater amberjack cpue for positive trips by season and region (CAR-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida).

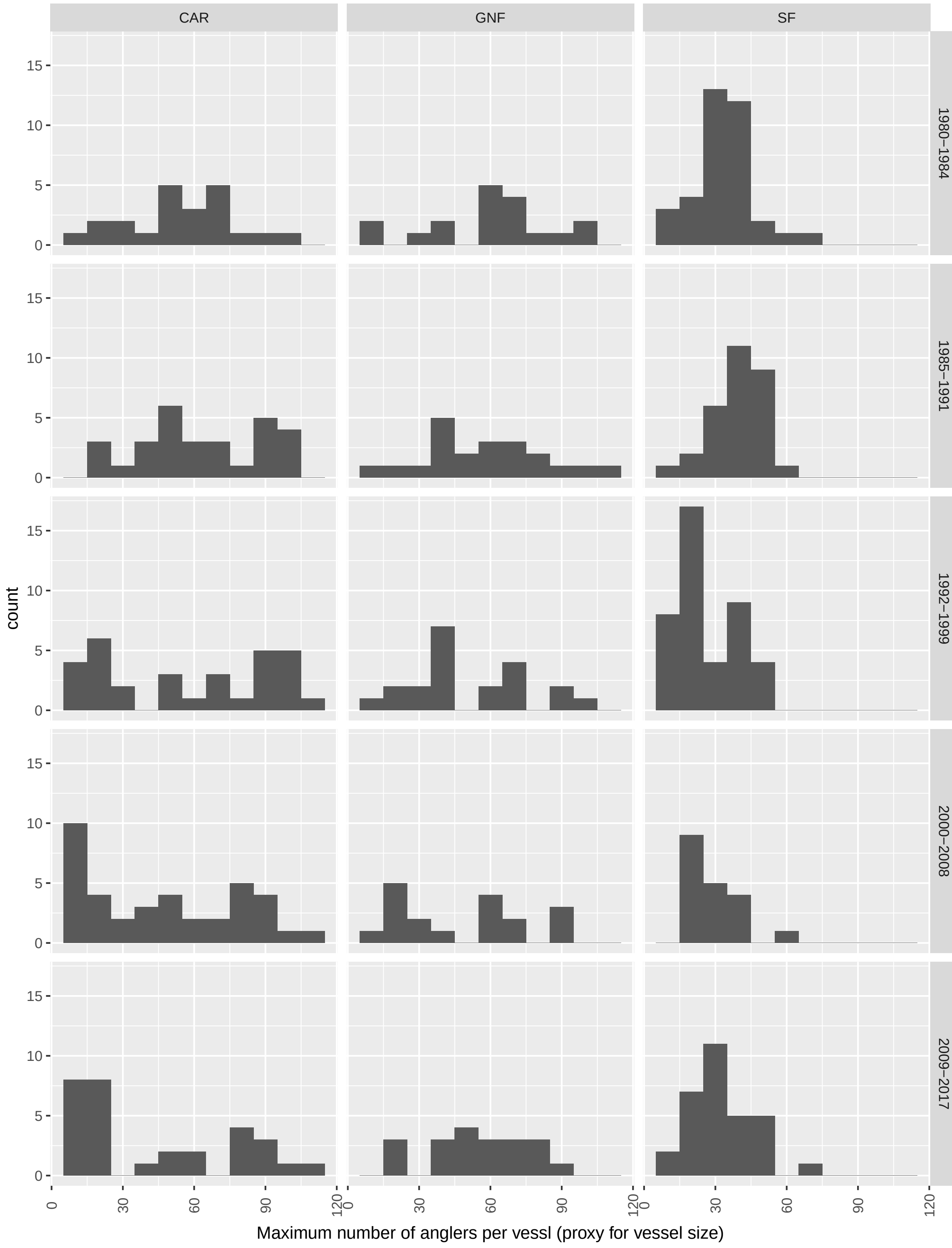
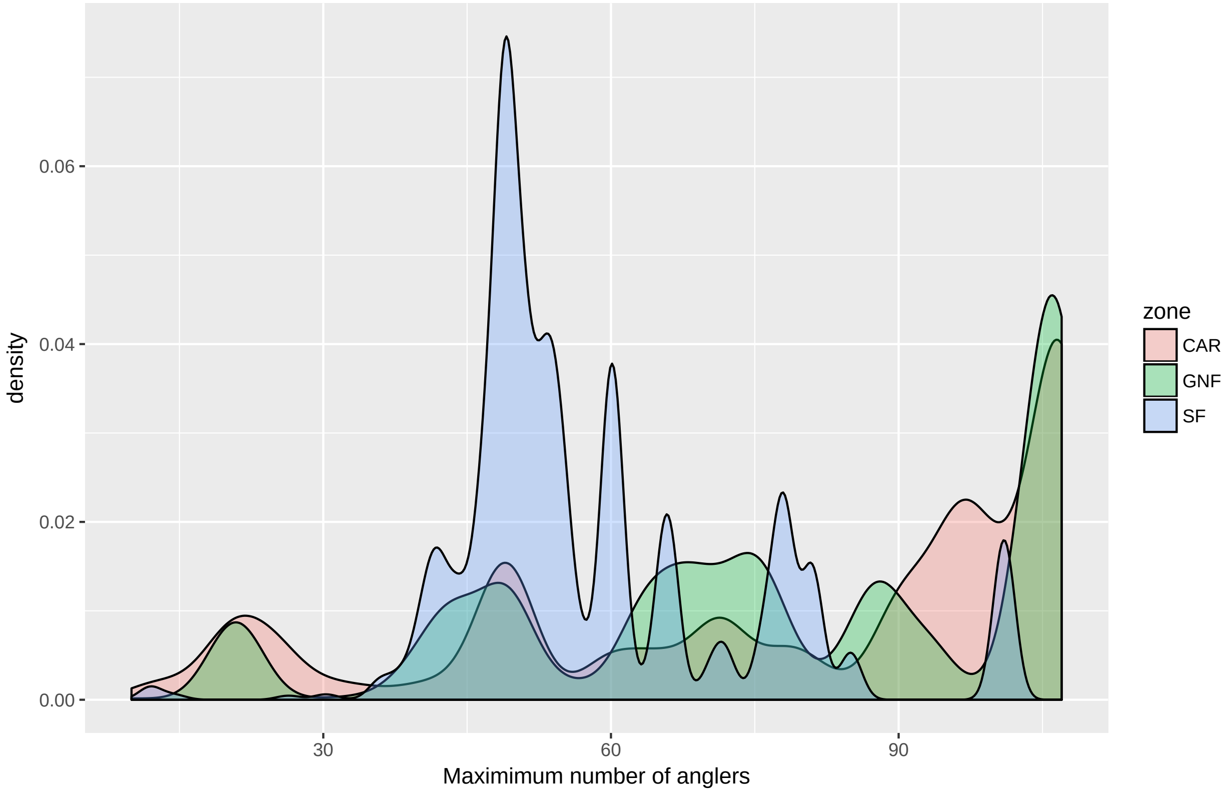


Figure 12: Maximum number of anglers as a proxy for vessel size (single value for each vessel) by region (CAR-carolinas, GNF - Georgia to Cape Canaveral, FL, SF - South florida).



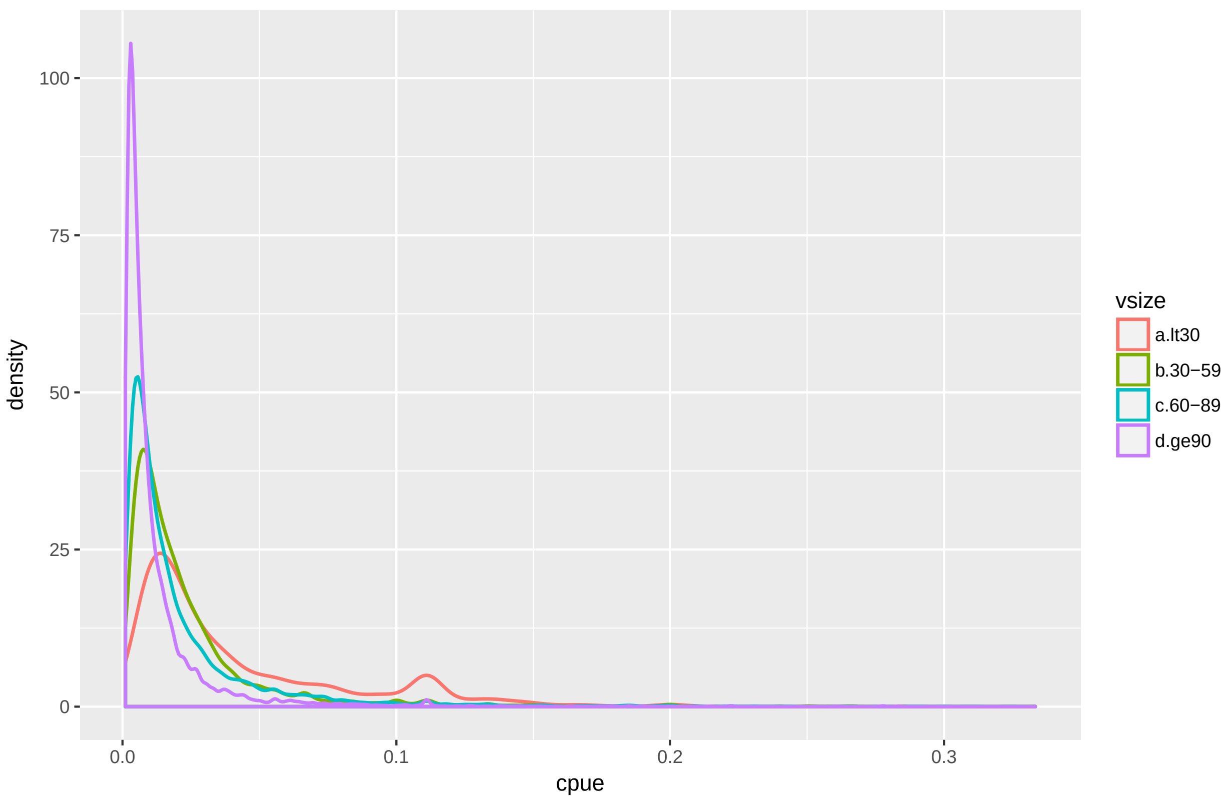


Figure 13: Density of maximum number of anglers across areas and cpue associated the factors for maximum anglers as a proxy for vessel size.

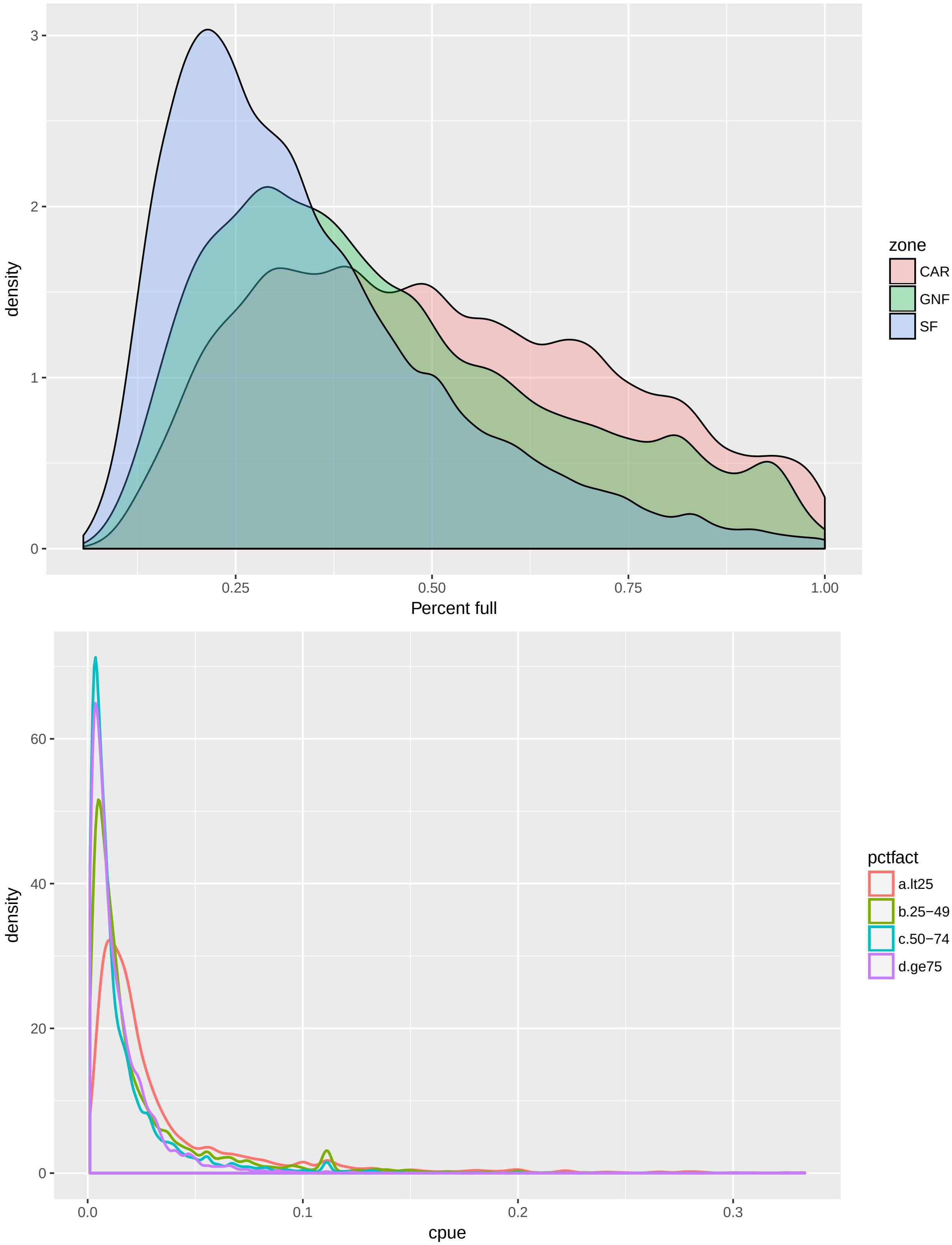


Figure 14: Density of percent full across areas and cpue associated the factors for percent full.

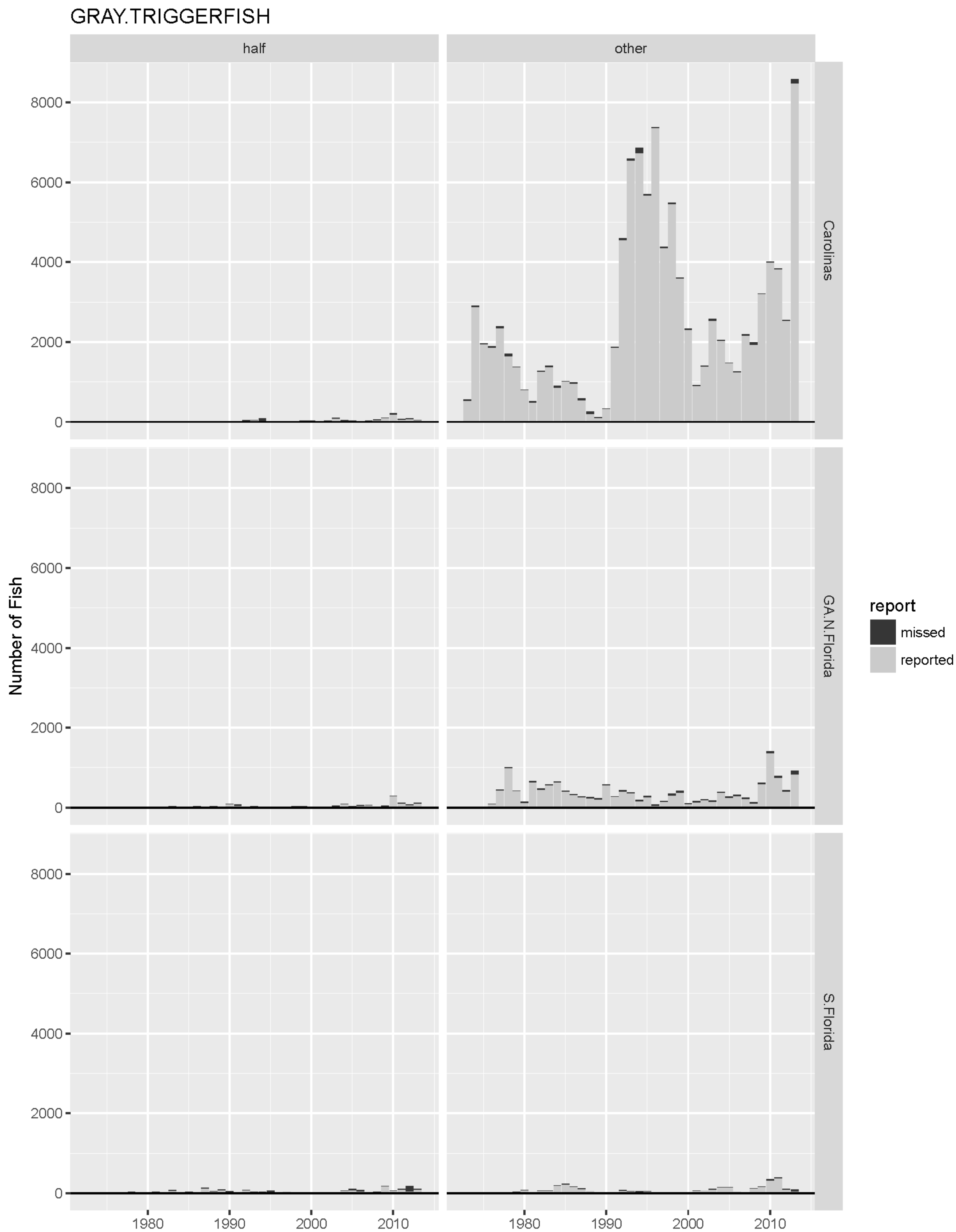


Figure 15: Comparison of the number of fish measured by a port sampler for a trips that could be matched to the number of fish reported for gray triggerfish. A general field for triggerfish was on the logbook form for in all areas from the beginning of the survey however headboat captains reported that they were mostly discarded during the early part of the survey because of angler preference. Gray Triggerfish was added to the for for the Carolinas in 1984.

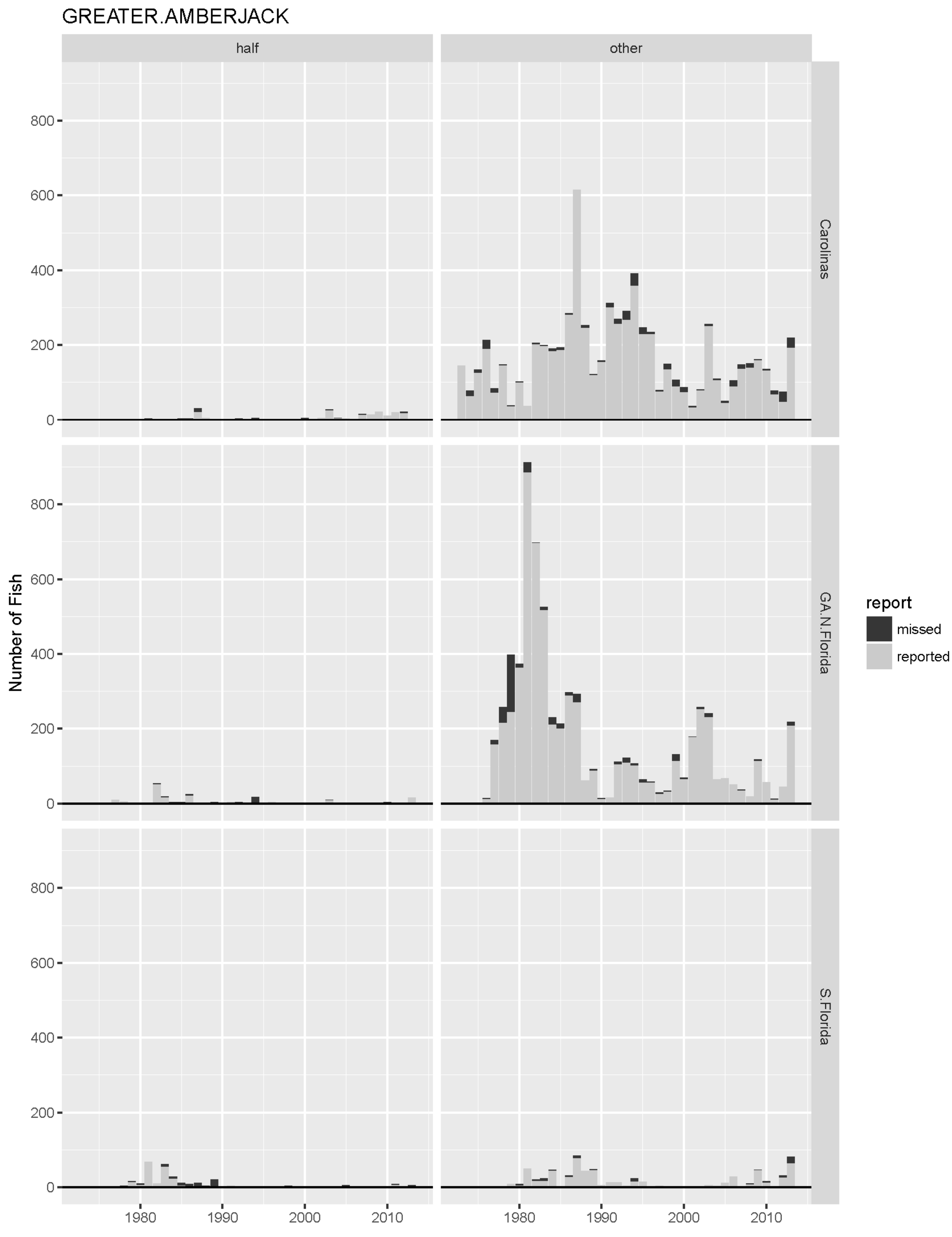


Figure 16: Comparison of the number of fish measured by a port sampler for a trips that could be matched to the number of fish reported for greater amberjack. Greater amberjack was on the logbook form for in all areas from the beginning of the survey and reporting is consistent over time.

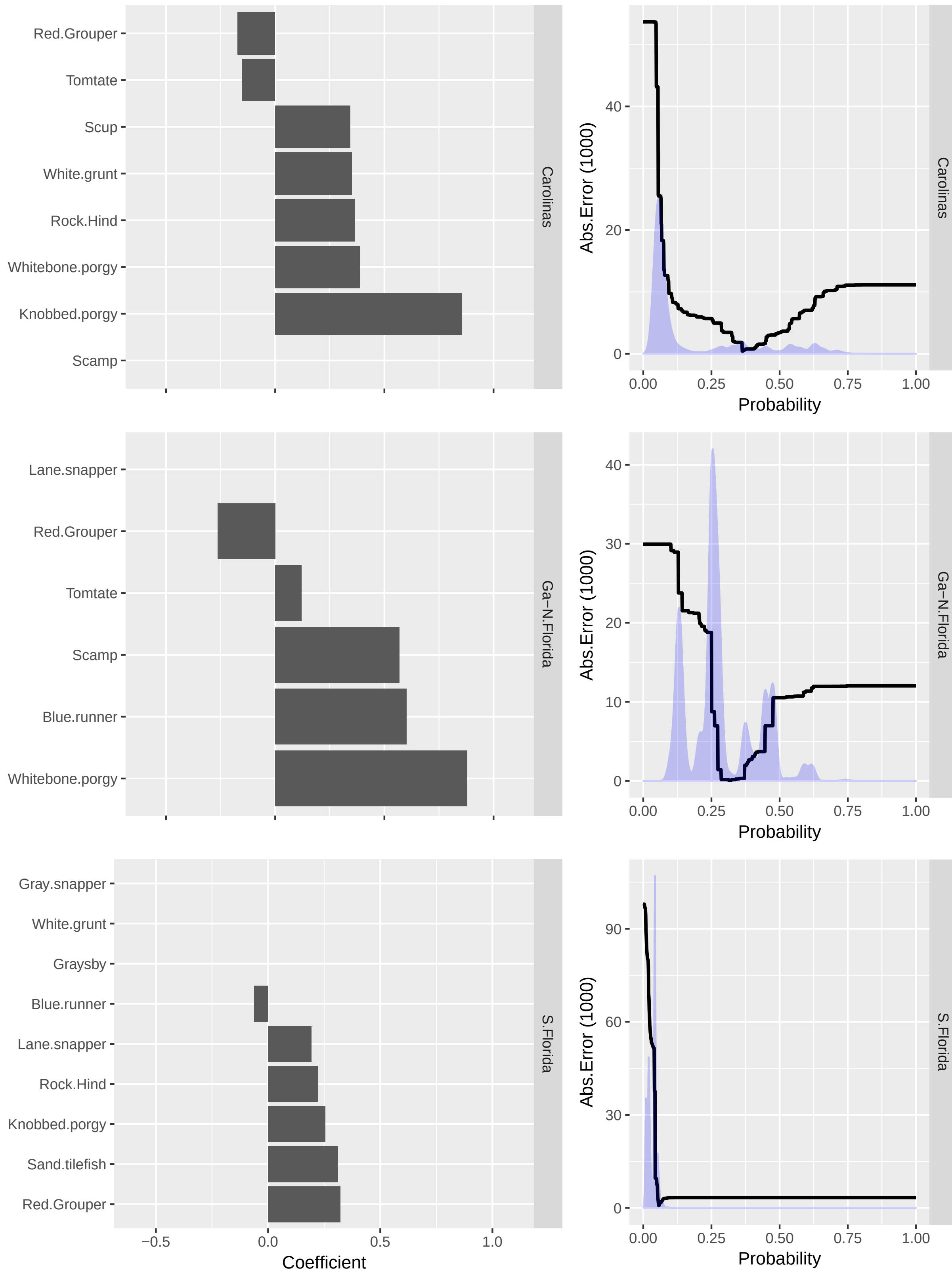


Figure 17: Estimates of species-specific regression coefficients used to predict each trip's probability of catching the focal species on the left panel. The right panel shows the absolute difference between observed and predicted number of positive trips across a range of probability cutoff values. .

Figure 18: Positive and zero greater amberjack trips retained after subsetting using Stephens and MacCall approach by year.

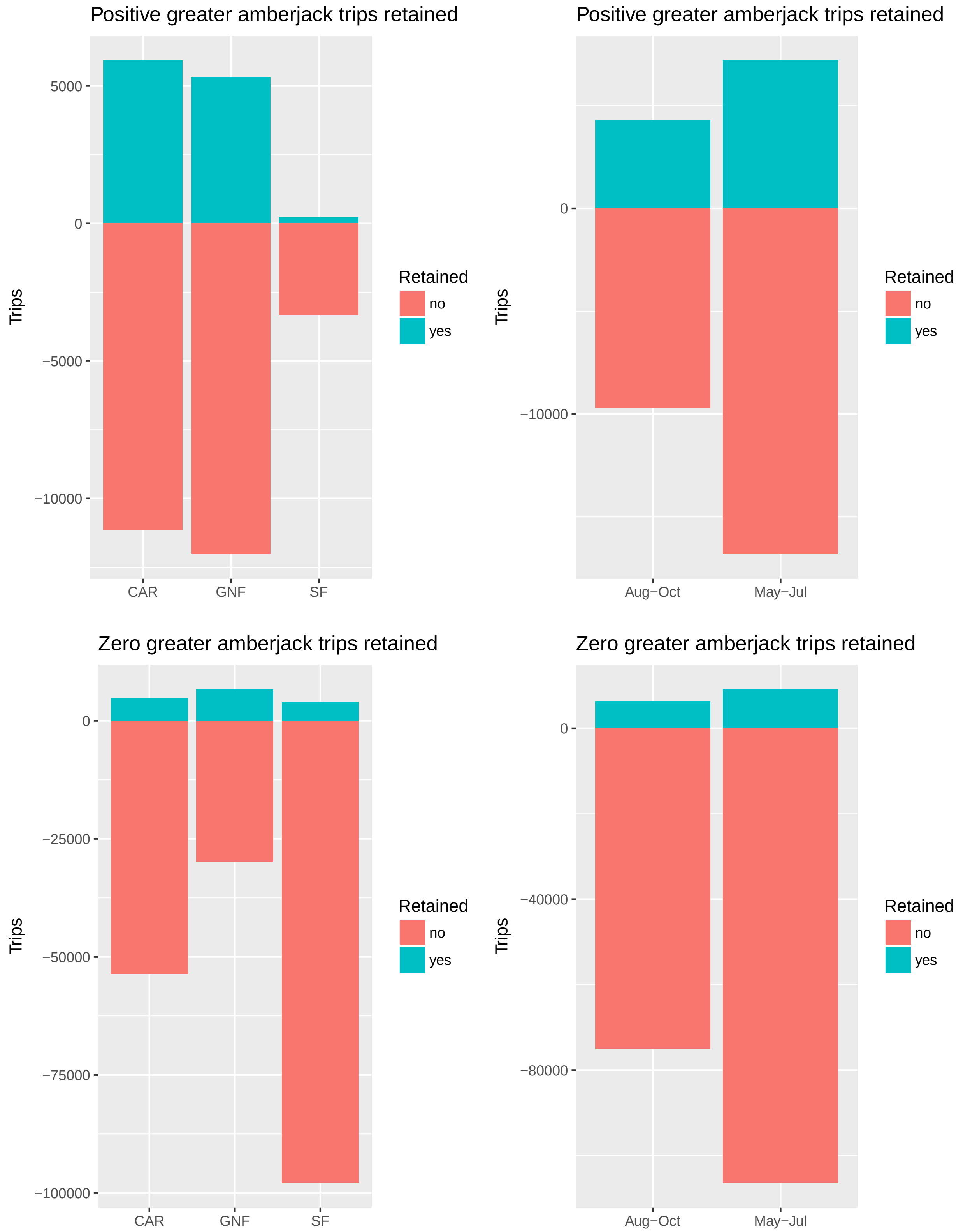


Figure 19: Positive and zero greater amberjack trips retained after subsetting using Stephens and MacCall approach by area and season.

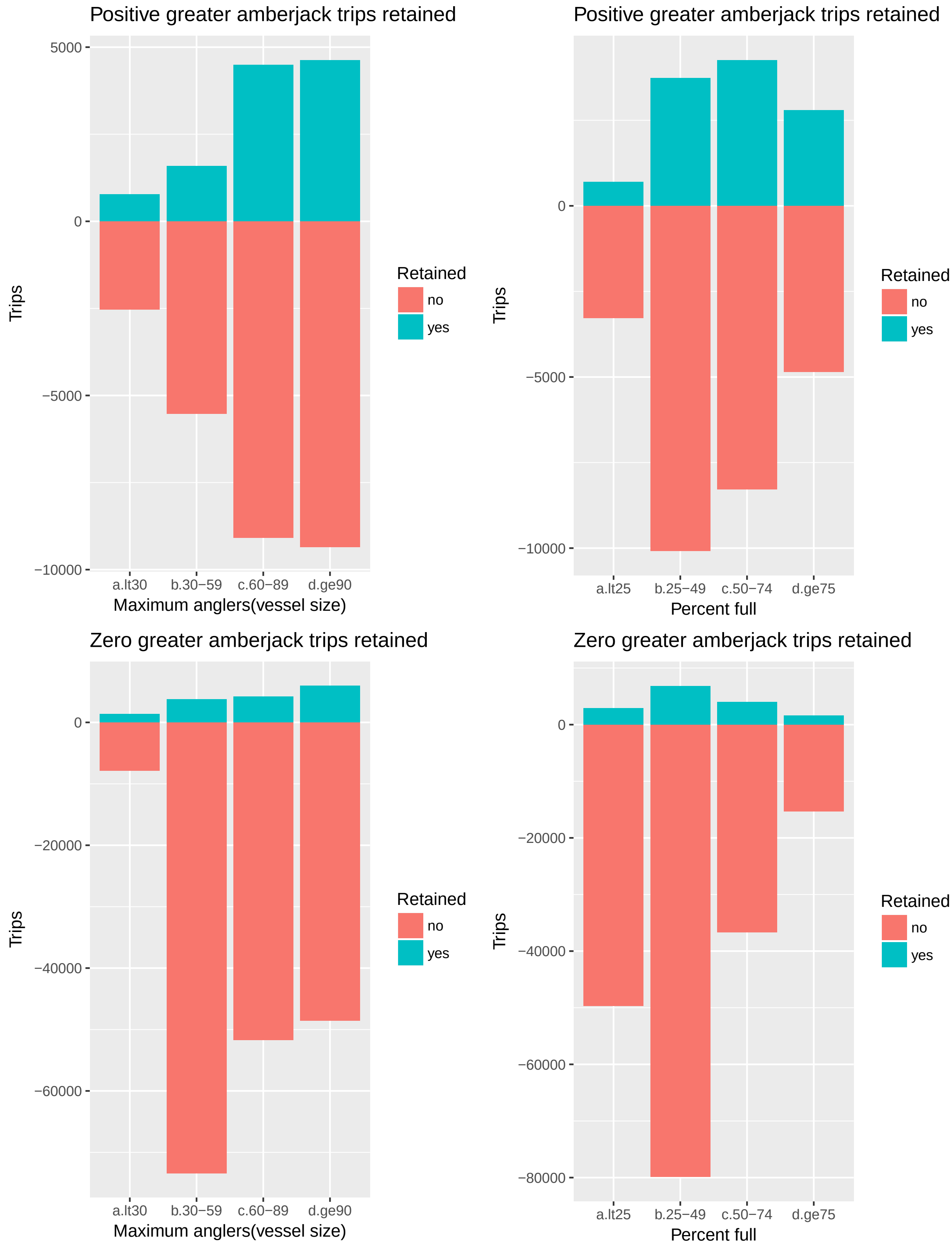


Figure 20: Positive and zero greater amberjack trips retained after subsetting using Stephens and MacCall approach by factors for maximum anglers and percent full.

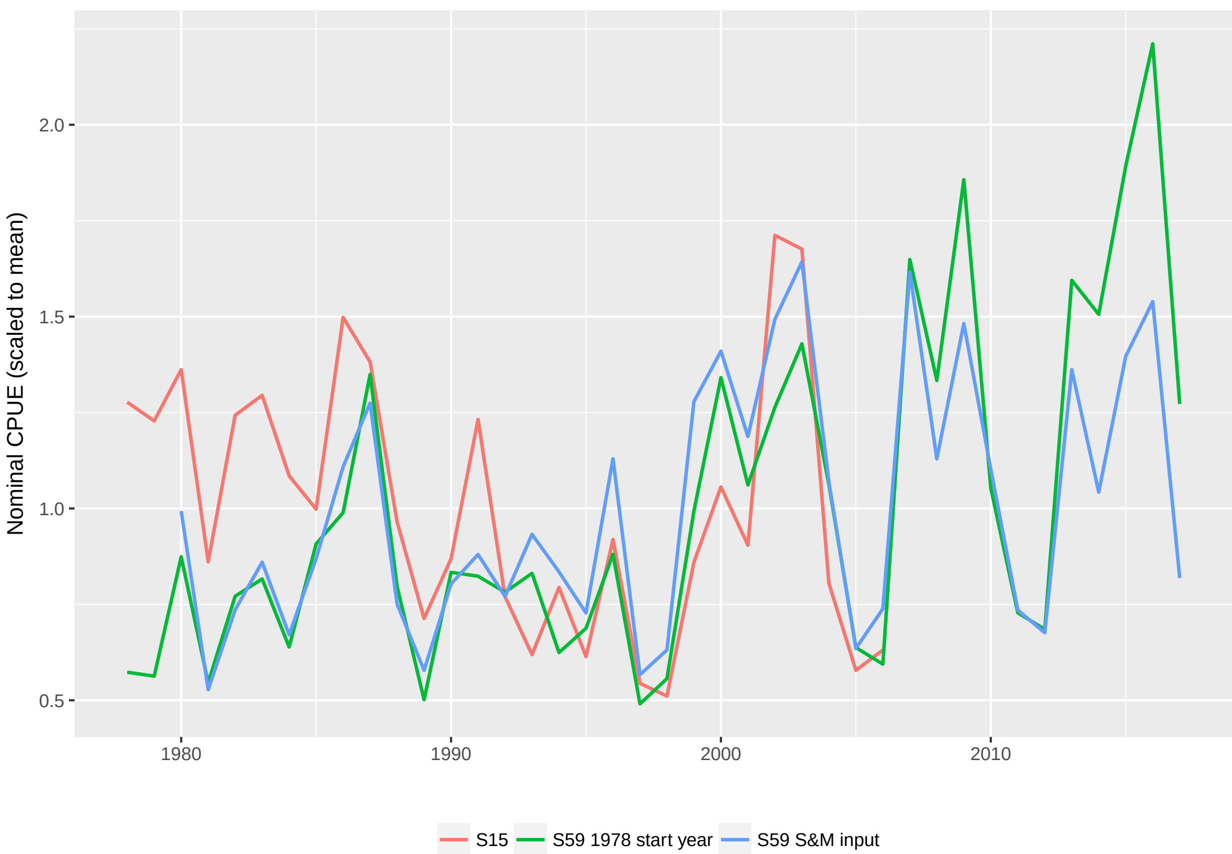


Figure 21: Nominal greater amberjack cpue for subsetted trips for SEDAR 15, SEDAR 59, and all positve trips.

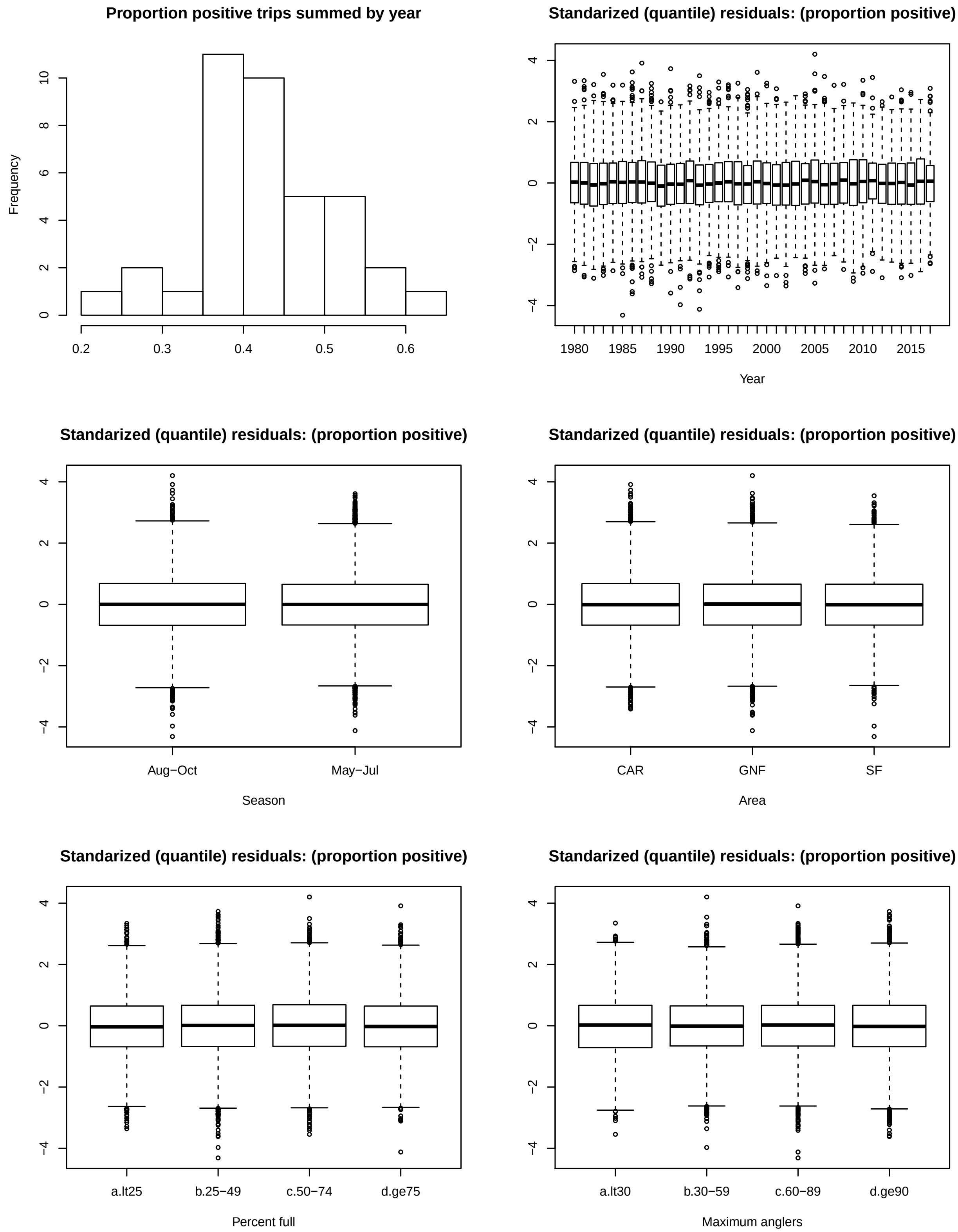


Figure 22: Diagnostics of Bernoulli submodel fits to positive versus zero CPUE data. Box and whisker plots give first, second (median) and third quartiles, as well as limbs that extend to approximately one interquartile range beyond the nearest quartile, and outliers (circles) beyond the limbs. Residuals are standardized (quantile) residuals.

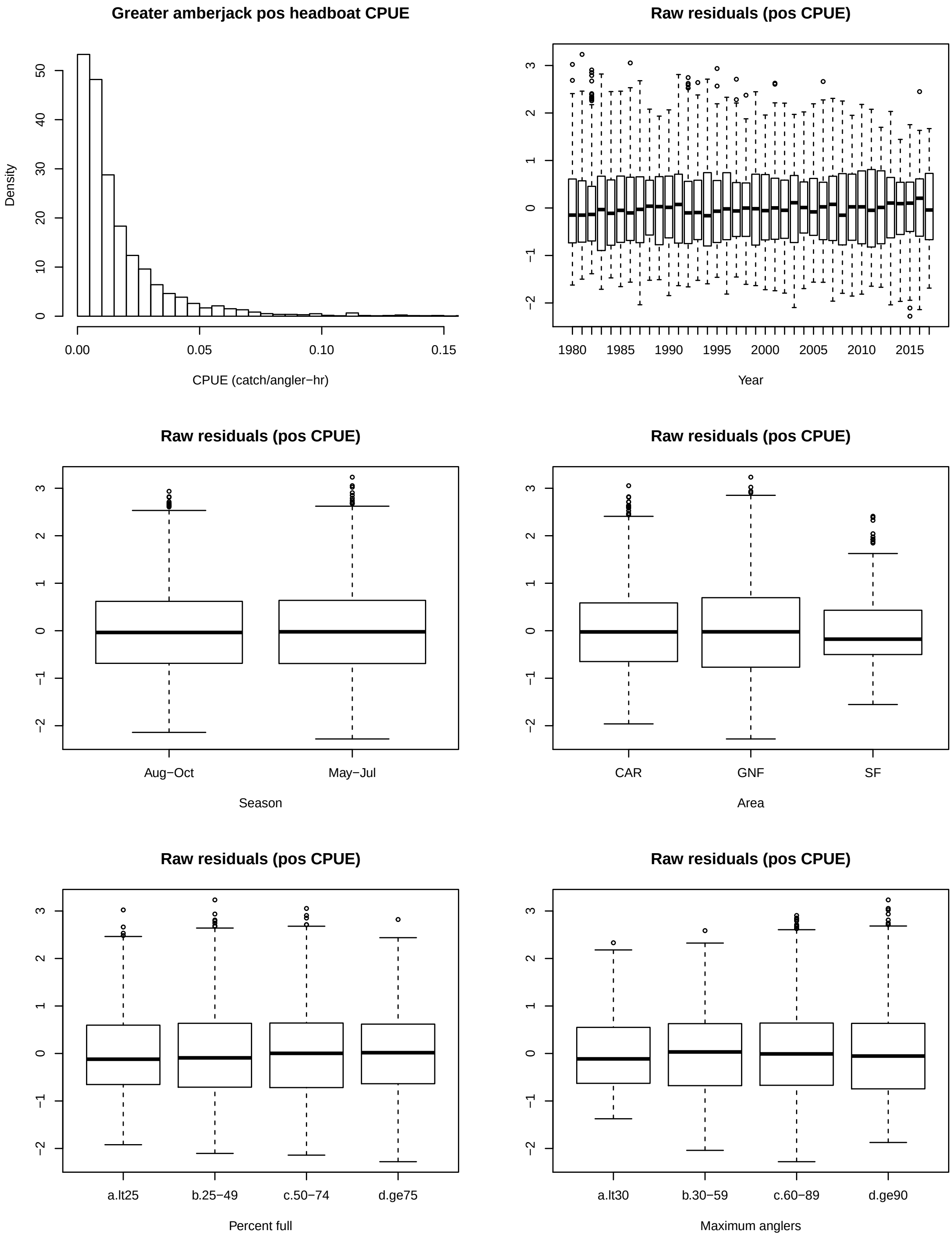
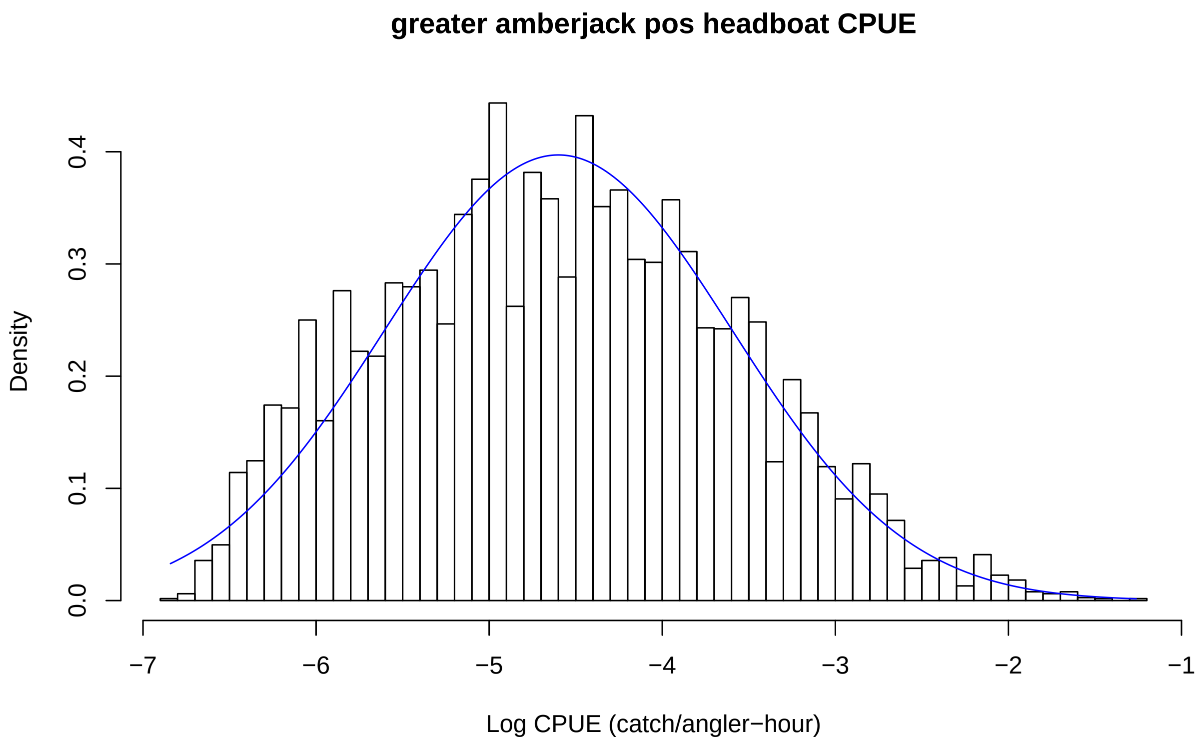


Figure 23: Diagnostics of lognormal submodel fits to positive CPUE data. Top left panel shows the distribution of positive cpue. Box and whisker plots give first, second (median) and third quartiles, as well as limbs that extend to approximately one interquartile range beyond the nearest quartile, and outliers (circles) beyond the limbs. Residuals are raw.



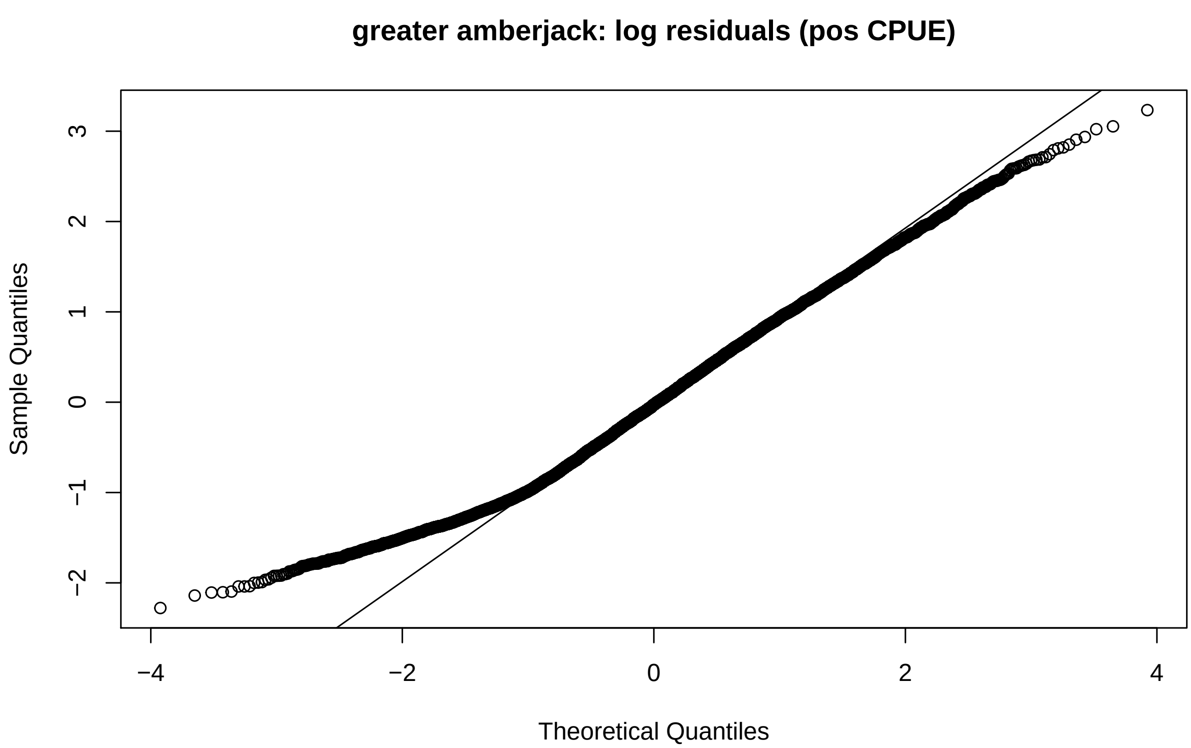


Figure 24: Histogram of empirical log CPUE, with the normal distribution (empirical mean and variance) overlaid. Quantile-quantile plot of residuals from the fitted lognormal submodel to the positive cpue cata.

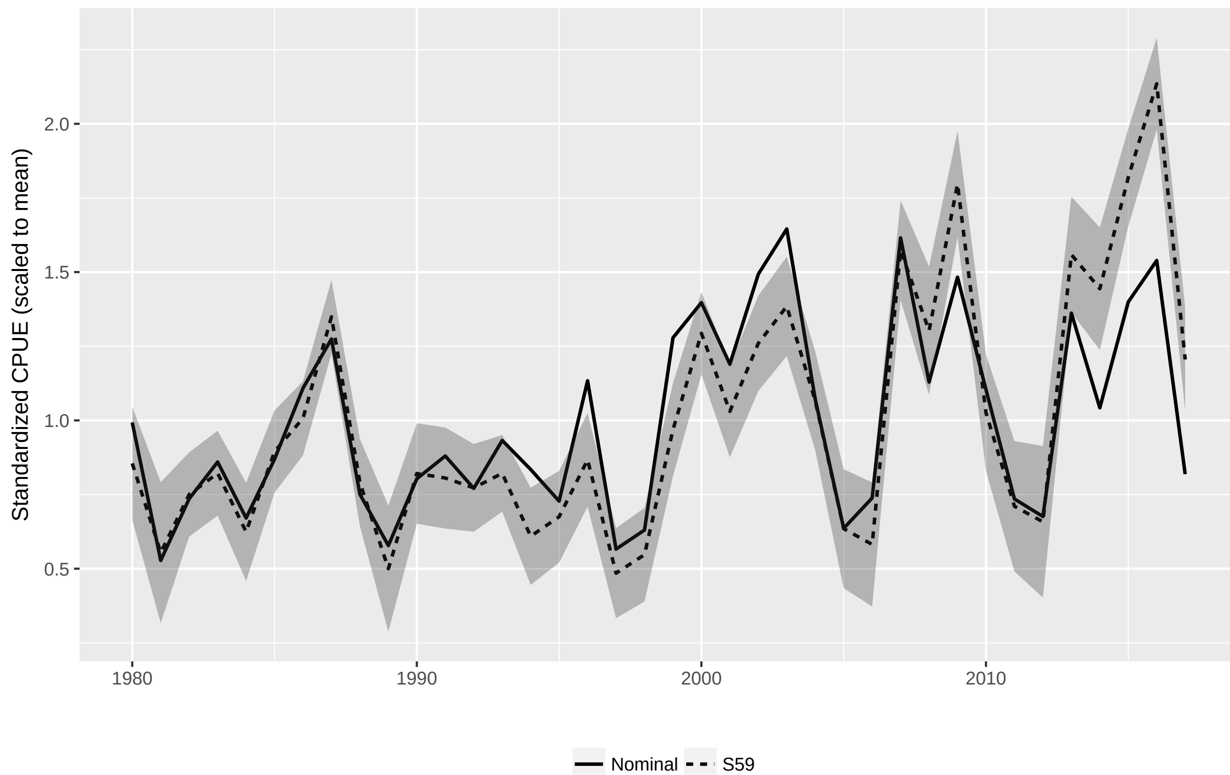


Figure 25: Standardized (dashed) with 95% confidence interval (shaded) and nominal index (solid) greater amberjack catch rate from headboat logbooks.

Appendix

Results of generalized linear model with Bernoulli response to select species associations with greater amberjack for the Carolinas.

```
##
## Call:  glm(formula = Greater.amberjack ~ Tomtate + White.grunt + Whitebone.porgy +
##     Red.Grouper + Scamp + Rock.Hind + Knobbed.porgy + Scup, family = "binomial",
##     data = n.sp.mat.trim)
##
## Coefficients:
##     (Intercept)          Tomtate      White.grunt  Whitebone.porgy
##         -2.8571          -0.1507           0.3521           0.3870
##     Red.Grouper            Scamp        Rock.Hind    Knobbed.porgy
##         -0.1723           1.9424           0.3651           0.8567
##            Scup
##          0.3435
##
## Degrees of Freedom: 64811 Total (i.e. Null);  64803 Residual
## Null Deviance:       59500
## Residual Deviance: 44940     AIC: 44960
```

Results of generalized linear model with Bernoulli response to select species associations with greater amberjack for the Georgia-N.Florida.

```
##
## Call:  glm(formula = Greater.amberjack ~ Tomtate + Whitebone.porgy +
##     Red.Grouper + Scamp + Blue.runner + Lane.snapper, family = "binomial",
##     data = m.sp.mat.trim)
##
## Coefficients:
##     (Intercept)          Tomtate  Whitebone.porgy      Red.Grouper
##         -1.0958           0.1198           0.8794          -0.2637
##           Scamp      Blue.runner     Lane.snapper
##          0.5698           0.6018          -0.8217
##
## Degrees of Freedom: 41975 Total (i.e. Null);  41969 Residual
## Null Deviance:       50280
## Residual Deviance: 47160     AIC: 47170
```

Results of generalized linear model with Bernoulli response to select species associations with greater amberjack for the S. Florida.

```
##
## Call:  glm(formula = Greater.amberjack ~ White.grunt + Red.Grouper +
##     Rock.Hind + Knobbed.porgy + Gray.snapper + Sand.tilefish +
##     Blue.runner + Graysby + Lane.snapper, family = "binomial",
##     data = s.sp.mat.trim)
##
## Coefficients:
##   (Intercept)    White.grunt    Red.Grouper      Rock.Hind  Knobbed.porgy
##      -3.09532       -0.72501        0.32177        0.22113        0.25449
##  Gray.snapper  Sand.tilefish    Blue.runner        Graysby   Lane.snapper
##      -0.82596        0.31157       -0.06201       -0.64940        0.19417
##
## Degrees of Freedom: 101248 Total (i.e. Null);  101239 Residual
```

```
## Null Deviance:       29360
## Residual Deviance: 28540     AIC: 28560
```

Results of lognormal glm to determine factors.

```
##
## Call:  glm(formula = log(cpue) ~ year + zone + season + vsize + pctfact,
##     family = gaussian(link = "identity"), data = dat.pos)
##
## Coefficients:
##    (Intercept)        year1981        year1982        year1983
##      -3.323914       -0.173101       -0.318908       -0.157448
##       year1984        year1985        year1986        year1987
##      -0.261352       -0.146551       -0.056207        0.151768
##       year1988        year1989        year1990        year1991
##      -0.060080       -0.089972        0.191341        0.092080
##       year1992        year1993        year1994        year1995
##      -0.072445       -0.082535       -0.169744       -0.227367
##       year1996        year1997        year1998        year1999
##       0.052492       -0.319297       -0.158363       -0.033939
##       year2000        year2001        year2002        year2003
##       0.093803        0.007696        0.137487        0.473616
##       year2004        year2005        year2006        year2007
##       0.063455       -0.171728       -0.170600        0.188519
##       year2008        year2009        year2010        year2011
##       0.259470        0.396130        0.180277        0.034157
##       year2012        year2013        year2014        year2015
##       0.188754        0.516736        0.444101        0.330816
##       year2016        year2017         zoneGNF          zoneSF
##       0.326994        0.053767        0.150649        0.127613
##  seasonMay-Jul    vsizeb.30-59    vsizec.60-89    vsized.ge90
##      -0.064028       -0.333934       -0.457069       -1.045833
## pctfactb.25-49  pctfactc.50-74  pctfactd.ge75
##      -0.463580       -0.760917       -0.974113
##
## Degrees of Freedom: 11477 Total (i.e. Null);  11431 Residual
## Null Deviance:       11580
## Residual Deviance: 9034  AIC: 29920
```

Results of gamma glm to determine factors.

```
##
## Call:  glm(formula = cpue ~ year + zone + season + vsize + pctfact,
##     family = Gamma(link = "log"), data = dat.pos)
##
## Coefficients:
##    (Intercept)        year1981        year1982        year1983
##       -2.96261        -0.22663        -0.36629        -0.17578
##       year1984        year1985        year1986        year1987
##       -0.39170        -0.27898        -0.14413         0.09475
##       year1988        year1989        year1990        year1991
##       -0.24882        -0.21846         0.06767         0.03351
##       year1992        year1993        year1994        year1995
##       -0.05182        -0.19056        -0.18047        -0.31134
##       year1996        year1997        year1998        year1999
##       -0.06040        -0.46491        -0.32297        -0.04676
```

```
##       year2000       year2001       year2002       year2003
##       -0.02354       -0.07423       -0.01344        0.41988
##       year2004       year2005       year2006       year2007
##       -0.15462       -0.30193       -0.27567        0.06850
##       year2008       year2009       year2010       year2011
##        0.22288        0.32917        0.12654        0.01093
##       year2012       year2013       year2014       year2015
##        0.08950        0.35622        0.20292        0.11569
##       year2016       year2017        zoneGNF         zoneSF
##        0.12768       -0.09593        0.25998        0.12305
##  seasonMay-Jul    vsizeb.30-59   vsizec.60-89    vsized.ge90
##       -0.05890       -0.28953       -0.34156       -0.93010
## pctfactb.25-49  pctfactc.50-74  pctfactd.ge75
##       -0.46172       -0.77056       -1.00694
##
## Degrees of Freedom: 11477 Total (i.e. Null);  11431 Residual
## Null Deviance:       11810
## Residual Deviance: 9305  AIC: -73820
```

Results of binomial glm to determine factors.

```
##
## Call:  glm(formula = cpue ~ year + zone + vsize + pctfact, family = "binomial",
##     data = dat.bin)
##
## Coefficients:
##    (Intercept)       year1981       year1982       year1983
##      -1.157401      -0.335256       0.260937       0.162128
##       year1984       year1985       year1986       year1987
##      -0.070644       0.266568       0.307322       0.432788
##       year1988       year1989       year1990       year1991
##      -0.025380      -0.561778      -0.299518      -0.197751
##       year1992       year1993       year1994       year1995
##      -0.038040       0.057286      -0.220560      -0.012245
##       year1996       year1997       year1998       year1999
##      -0.050838      -0.319454      -0.369190       0.217776
##       year2000       year2001       year2002       year2003
##       0.457283       0.247559       0.351401       0.011247
##       year2004       year2005       year2006       year2007
##       0.215276      -0.165571      -0.278366       0.616965
##       year2008       year2009       year2010       year2011
##       0.221421       0.497632       0.002291      -0.284533
##       year2012       year2013       year2014       year2015
##      -0.566043       0.114658       0.107947       0.621449
##       year2016       year2017        zoneGNF         zoneSF
##       0.904334       0.409975      -0.432898      -2.725776
##   vsizeb.30-59   vsizec.60-89    vsized.ge90  pctfactb.25-49
##       0.482502       0.979899       0.537873       0.396306
## pctfactc.50-74   pctfactd.ge75
##       0.816025       1.295277
##
## Degrees of Freedom: 26758 Total (i.e. Null);  26713 Residual
## Null Deviance:       36550
## Residual Deviance: 31180     AIC: 31270
```

Results of lognormal delta glm to compare models.

```
## $error.distribution
## [1] "Lognormal distribution assumed for positive observations."
##
## $binomial.formula
## [1] "Formula for binomial GLM: cpue ~ year + zone + vsize + pctfact + season"
##
## $positive.formula
## [1] "Formula for gaussian GLM: log(cpue) ~ year + zone + vsize + pctfact + season"
##
## $deltaGLM.index
##            index jackknife
## 1980 0.005240957        NA
## 1981 0.003397819        NA
## 1982 0.004598991        NA
## 1983 0.005038749        NA
## 1984 0.003826514        NA
## 1985 0.005486624        NA
## 1986 0.006174502        NA
## 1987 0.008270299        NA
## 1988 0.004844183        NA
## 1989 0.003065590        NA
## 1990 0.005032611        NA
## 1991 0.004939009        NA
## 1992 0.004736234        NA
## 1993 0.005035627        NA
## 1994 0.003734442        NA
## 1995 0.004138895        NA
## 1996 0.005315461        NA
## 1997 0.002972368        NA
## 1998 0.003354540        NA
## 1999 0.005936373        NA
## 2000 0.007930259        NA
## 2001 0.006316745        NA
## 2002 0.007722470        NA
## 2003 0.008487914        NA
## 2004 0.006532987        NA
## 2005 0.003892055        NA
## 2006 0.003565900        NA
## 2007 0.009645814        NA
## 2008 0.007983526        NA
## 2009 0.011015084        NA
## 2010 0.006286555        NA
## 2011 0.004354922        NA
## 2012 0.004033313        NA
## 2013 0.009557709        NA
## 2014 0.008850826        NA
## 2015 0.011150281        NA
## 2016 0.013087717        NA
## 2017 0.007389206        NA
##
## $pos.effects
## $pos.effects[[1]]
```

```
##        CAR        GNF         SF
## 0.01964763 0.02284211 0.02232194
##
## $pos.effects[[2]]
##     a.lt30    b.30-59    c.60-89     d.ge90
## 0.03412123 0.02443425 0.02160341 0.01199016
##
## $pos.effects[[3]]
##     a.lt25    b.25-49    c.50-74     d.ge75
## 0.03735114 0.02349490 0.01745186 0.01410106
##
## $pos.effects[[4]]
##    Aug-Oct    May-Jul
## 0.02225852 0.02087802
##
##
## $bin.effects
## $bin.effects[[1]]
##        CAR        GNF         SF
## 0.51019946 0.40334645 0.06389044
##
## $bin.effects[[2]]
##    a.lt30   b.30-59   c.60-89    d.ge90
## 0.1806868 0.2631624 0.3700780 0.2741063
##
## $bin.effects[[3]]
##    a.lt25   b.25-49   c.50-74    d.ge75
## 0.1625046 0.2240088 0.3052950 0.4151557
##
## $bin.effects[[4]]
##   Aug-Oct   May-Jul
## 0.2673705 0.2658993
##
##
## $data.filter
## [1] "Data filter threshold set at 2 positive observations."
##
## $levels.deleted.by.filter
## $levels.deleted.by.filter$year
## [1] NA
##
## $levels.deleted.by.filter$zone
## [1] NA
##
## $levels.deleted.by.filter$vsize
## [1] NA
##
## $levels.deleted.by.filter$pctfact
## [1] NA
##
## $levels.deleted.by.filter$season
## [1] NA
##
##
```

```
## $aic
##                        [,1]
## AIC.binomial   3.127108e+04
## AIC.lognormal -7.567882e+04
## sigma.mle      8.871699e-01
```

Results of gamma delta glm to compare models.

```
## $error.distribution
## [1] "Gamma distribution assumed for positive observations."
##
## $binomial.formula
## [1] "Formula for binomial GLM: cpue ~ year + zone + vsize + pctfact + season"
##
## $positive.formula
## [1] "Formula for Gamma GLM: cpue ~ year + zone + vsize + pctfact + season"
##
## $deltaGLM.index
##           index jackknife
## 1980 0.02204905        NA
## 1981 0.01757780        NA
## 1982 0.01528672        NA
## 1983 0.01849473        NA
## 1984 0.01490310        NA
## 1985 0.01668132        NA
## 1986 0.01908946        NA
## 1987 0.02424035        NA
## 1988 0.01719205        NA
## 1989 0.01772204        NA
## 1990 0.02359282        NA
## 1991 0.02280041        NA
## 1992 0.02093561        NA
## 1993 0.01822350        NA
## 1994 0.01840819        NA
## 1995 0.01615021        NA
## 1996 0.02075672        NA
## 1997 0.01385109        NA
## 1998 0.01596343        NA
## 1999 0.02104176        NA
## 2000 0.02153601        NA
## 2001 0.02047172        NA
## 2002 0.02175472        NA
## 2003 0.03355363        NA
## 2004 0.01889037        NA
## 2005 0.01630278        NA
## 2006 0.01673667        NA
## 2007 0.02361231        NA
## 2008 0.02755409        NA
## 2009 0.03064409        NA
## 2010 0.02502336        NA
## 2011 0.02229128        NA
## 2012 0.02411344        NA
## 2013 0.03148431        NA
## 2014 0.02700955        NA
## 2015 0.02475343        NA
```

```
## 2016 0.02505199        NA
## 2017 0.02003217        NA
##
## $pos.effects
## $pos.effects[[1]]
##        CAR        GNF         SF
## 0.01824075 0.02365656 0.02062915
##
## $pos.effects[[2]]
##     a.lt30    b.30-59    c.60-89     d.ge90
## 0.03061931 0.02292202 0.02176001 0.01207981
##
## $pos.effects[[3]]
##     a.lt25    b.25-49    c.50-74     d.ge75
## 0.03627535 0.02286079 0.01678656 0.01325265
##
## $pos.effects[[4]]
##    Aug-Oct    May-Jul
## 0.02134429 0.02012343
##
##
## $bin.effects
## $bin.effects[[1]]
## CAR GNF  SF
##   1   1   1
##
## $bin.effects[[2]]
##  a.lt30 b.30-59 c.60-89  d.ge90
##       1       1       1       1
##
## $bin.effects[[3]]
##  a.lt25 b.25-49 c.50-74  d.ge75
##       1       1       1       1
##
## $bin.effects[[4]]
## Aug-Oct May-Jul
##       1       1
##
##
## $data.filter
## [1] "Data filter threshold set at 2 positive observations."
##
## $levels.deleted.by.filter
## $levels.deleted.by.filter$year
## [1] NA
##
## $levels.deleted.by.filter$zone
## [1] NA
##
## $levels.deleted.by.filter$vsize
## [1] NA
##
## $levels.deleted.by.filter$pctfact
## [1] NA
```

```
##
## $levels.deleted.by.filter$season
## [1] NA
##
##
## $aic
##                          [,1]
## AIC.binomial      94.000000
## AIC.gamma     -73899.384913
## shape.mle          1.376332
```

References

Dick, E.J. 2004. "Beyond Lognormal Versus Gamma: Discrimination Among Error Distributions for Generalized Linear Models." *Fisheries Research* 70: 351–66.

Efron, B., and R. Tibshirani. 1993. *An Introduction to the Bootstrap*. London: Chapman; Hall.

Lo, N., Jacobson L., and Squire J. 1992. "Indices of Relative Abundance from Fish Spotter Data Based on Delta-Lognormal Models." *Canadian Journal of Fisheries and Aquatic Sciences* 49: 2515–26.

Maunder, M., and A. Punt. 2004. "Standardizing Catch and Effort Data: A Review of Recent Approaches." *Fisheries Research* 70: 141–59.

Shertzer, E., K., E. Williams, and Taylor J. 2009. "Spatial Structure and Temporal Patterns in a Large Marine Ecosystem: Exploited Reef Fishes of the Southeast United States." *Fisheries Research* 100: 126–33.

Stephens, A., and MacCall A. 2004. "A Multispecies Approach to Subsetting Logbook Data for Purposes of Estimating Cpue." *Fisheries Research* 70: 299–310.

Venables, W. N., and B. D. Ripley. 1997. *Modern Applied Statistics with S-Plus, 2nd Edition*. New York, New York: Springer-Verlag.