Preliminary standardized index of Southeast US Atlantic cobia (*Rachycentron canadum*) from headboat data. (Revised 4/5/19)

Sustainable Fisheries Branch, National Marine Fisheries Service,
Southeast Fisheries Science Center

SEDAR58-DW09

April 10 2019

Please cite this document as:

Sustainable Fisheries Branch, National Marine Fisheries Service,
Southeast Fisheries Science Center. Revised April 5 2019. Standardized index of Southeast US Atlantic cobia (Rachycentron canadum) from headboat data.. SEDAR58-DW09. SEDAR, North Charleston, SC. 29 pp.

Standardized index of Southeast US Atlantic cobia (*Rachycentron canadum*) from headboat data.

Sustainable Fisheries Branch, National Marine Fisheries Service,
Southeast Fisheries Science Center,
101 Pivers Island Rd, Beaufort, NC 28516

Revised April 5 2019

Abstract
A standardized index of abundance was developed using cobia count data from the Southeast headboat survey trip records (logbooks) for 1991-2015. The analysis included areas from North Carolina through Georgia. The index is meant to describe population trends of fish in the size/age range of fish landed by headboat vessels. Data filtering and subsetting steps were applied to the data to model trips that were likely to have directed cobia effort.

Background

The headboat fishery in the south Atlantic includes for-hire vessels. The fishery uses hook and line gear, generally targets hard bottom reefs as the fishing grounds, and generally targets multiple species in the snapper-grouper complex. One of the key characteristics defining a headboat from other recreational fishing such as charter boats is the number of anglers. Prior to 2000 headboats were defined as vessels carrying 15 or more recreational anglers. This criteria changed to 7 or more passengers in 2000 in the Atlantic (Ken Brennan, pers. comm. Dec. 2011).

Headboats in the south Atlantic are sampled from North Carolina to the Florida Keys.
Data have been collected since 1972, but logbook reporting did not start until 1973. In addition, only North Carolina and South Carolina were included in the earlier years of the data set. In 1976, data were collected from North Carolina, South Carolina and northern Florida. Full coverage began in 1978 with data collected from southern Florida while Georgia didn't have active vessels in the fishery until 1993 (Areas 1-17, Figure 1). Variables reported in the data set include year, month, day, area, location, trip type, number of anglers, species, catch, and vessel id. Biological data and discard data were recorded for some trips in some years.

A bag limit of 2 cobia/person/day has been in place since August 20 1990 in federal waters.

Issues

Cobia are a semi-large coastal migratory pelagic species and thus not ideally targeted by snapper-grouper bottom fishing gear.

However, not all anglers on a headboat use typical bottom fishing gear (e.g. the surface line fishers off the stern). So, one unknown is what proportion of cobia caught on headboats are caught on bottom rigs versus surface line rigs? If the predominant cobia interaction is on surface line rigs, then how often are surface lines deployed on headboat trips? We have no data to examine this gear issue.

A necessary assumption is that the probability of encountering a cobia on a headboat is fairly constant with respect to surface line use. Let's ignore that issue and assume that the proportion of surface lines deployed is constant on all headboat trips. We know that cobia are a seasonal migrant and thus we expect the encounter rate with headboats to change with seasons. We also know that cobia migrate spatially, so that the area is also a contributing factor to determining the probability of catching a cobia on a headboat. Ideally, we can use the history of cobia catches on all headboats to determine the time and space strata where cobia have a reasonable chance of being caught. However, there is an underlying assumption that you have a period of time in the data where cobia were abundant enough to fully fill out the time and space strata (i.e. identify the time/space strata). With low catch rates of cobia and the history of fishing for the species, this assumption may be violated.

On any given headboat trip the captain makes a choice once they leave the dock as to where the boat is going to fish. Catching cobia is unlikely to factor into that decision, rather the choice is going to be more likely based on maximizing the catch of bottom fishes. So, another assumption is that within the space/time strata where cobia are caught the captain's location selection has the same probability of capturing cobia. This may depend on the site variability within a captains fishing area.

Another issue with cobia is they are a thigmotaxic species that are drawn to structure. To what degree is the catch rate of cobia on a headboat affected by this behavior?

Cobia has not been a popular game fish throughout the whole time period of headboat data. Popularity of this fish rose sometime in the 1990s and 2000s so that there is likely an overall desirability that may have affected both retention rates (keeping or discarding caught cobia), reporting rates of cobia by captains and gear usage (percentage of anglers deploying surface lines).

Identifying a period of time (seasonally and temporally) where these issues are relatively stable would be necessary in the development of an index of abundance for cobia.

Exploratory Data Analysis

Headboat records were examined to determine if sufficient data exists to develop a standardized index of abundance for south Atlantic cobia.

Positive cobia trips represent a small fraction of the overall composition in the south Atlantic headboat fishery (~1-7% annually) (Table 1). Since 1981, an average 111 cobia were captured per year in the south Atlantic headboat fishery (NC-GA). Data filtering steps were applied to identify trips that likely had directed cobia effort.

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 for trips that caught cobia. 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 were removed. Figure 2 shows the excluded trips based on outlier definitions by region. Removing a small percentage of the trips with extreme values 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. Core Vessels

To identify headboat trips that best characterize the cobia fishery, vessels that consistently caught cobia were selected (26 headboats representing 90% (prior to any filtering) of cobia effort and landings). Cobia trips from these 'core' vessels increased from 4% (all data) to 6% (full and half day trips, model input) (Table 1). Prior to the DW, proportion positive averaged near 10% from 1991-2015. Additional subsetting methods were explored (identify vessels using a proportion positive cutoff (Figure 3) but led to a reduction in positive trips and convergence issues since this approach identified vessels that had fewer overall trips but more cobia interactions. Selecting data using a core group of vessels while removing vessels that inconsistently or never reported cobia more appropriately reflects directed cobia effort in the headboat fishery.

4. Starting year

During SEDAR 28 the starting year for the headboat index was 1981. Upon further examination, a different start year should be considered.

Figure 4 illustrates an increase of the 'core' headboats entering the fleet from the 1980s to 1991 when the percentage of 'core' vessels active in the fishery tends to stabilize near 60% in 1991 (Figure 5). Figure 6 shows an increase in cobia cpue over time from the 1980s to the more recent years, which may indicate low abundance in these early years, or a shift in desirability that may have influenced reporting of cobia by captains. Also, cobia was not listed on the logbook form until 1984.

Due to the following concerns, 1991 was chosen as the start year for the cobia headboat index:

- Inconsistent reporting in the early years,

- the addition of new ‘core’ vessels in the middle to late 1980s,
- concerns about write-ins prior to cobia being listed on the catch record form in 1984

5. Terminal year - spawning closure exclusion

Seasonal closures occurred in 2016 (closed June 19) and 2017 (closed January 23). 2015 was chosen as the terminal year due to these regulations.

6. Trip types

For SEDAR 58, the relatively few multi-day trips and the 3/4-day trips were filtered. Figure 7 shows the variability associated with these trip types by region. Trips by region for full and half day trips are presented in Figure 8.

Stephens & McCall
Applying methods described by Stephens & McCall (2004) for cobia resulted in an approximate 67% reduction in positive cobia trips. A large reduction in positive cobia trips and an inflation of zero cobia trips was anticipated due to the infrequency of cobia in the headboat fishery, therefore a more appropriate method was pursued.

Evaluation of explanatory variables

YEAR - Year was necessarily included, as standardized catch rates by year are the desired outcome. Years modeled were 1991-2015.

TRIP TYPE – Full and half day trips were included in the standardization.

SEASON - For SEDAR 58, seven of the months (September-March) were dropped due to inconsistent cobia trips leaving 2 levels for season in the model. The seasonal pattern in cpue across months seems consistent across regions (Figure 9).

REGION - The regions modeled for the SEDAR 28 headboat logbook index were modified for SEDAR 58 and were more evenly split based on annual trips by each inlet instead of by state alone due to the southern boundary designation and limited samples from Georgia. The three regions include inlets from NC to GA (St. Mary’s GA.- Murrell’s Inlet SC (1), Little River, SC – Carolina Beach NC(2), Masonboro Inlet NC – Oregon Inlet NC (3)) (Figure 10).

VESSEL SIZE (vsize) - A factor was explored for the vessel size using the quartiles of the maximum number of anglers across all trips as breaks for the factors. The proxy for vessel size is the maximum anglers reported over all trips for a vessel (Figure 11). Due to limited data and convergence issues, vessel size was modified to two levels: ‘small’ or ‘large’.

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 but

subsequently led to convergence issues due to low sample sizes and therefore was modified to two levels: ‘partial’ or ‘full’. The density of percent full by region and the density of cpue associated with each factor are shown in figure 12.

PARTY SIZE (party size)
The number of anglers reported for a trip was divided into 4 equally spaced factors but led to convergence issues due to low sample sizes and therefore was modified to two levels: ‘small’ or ‘large’.

Following filtering and subsetting, trips retained for model input are presented in Figures 13 and 14.

Standardization

Zero-inflated models are valuable tools for modeling distributions that do not fit standard error distributions due to excessive number of zeroes. These data distributions are often referred to as “zero-inflated” and are a common condition of count based ecological data. Zero inflation is considered a special case of over-dispersion that is not readily addressed using traditional transformation procedures (Hall 2000). Due to the high proportion of zero counts found in our data set (Figure 15), we used a zero-inflated mixed model approach that accounts for the high occurrence of zero values, as well as the positive counts. The model does so by combining binomial and count processes (Zuur et al. 2009).

The modeling approached used here was similar to that used in SEDAR41 for gray triggerfish and red snapper for the video index. We initially considered a full null model (1) using both a zero-inflated Poisson (ZIP) and a zero-inflated negative binomial (ZINB) formulation,

$$Count = y + i + t + s + p + v + a \mid y + i + t + s + p + v + a \quad (1)$$

In this formulation, variables to the left of the “|” apply to the count sub-model, and variables to the right apply to the binomial sub-model. In this analysis, we favored a simpler null model because of the relatively small proportion of positive counts for cobia,

$$Count = y \mid y \quad (2)$$

which allowed us to add covariates using a step-wise forward selection process (rather than the backward selection). However, prior to adding covariates we compared ZIP and ZINB formulations. We compared the variance structure of each model formulation using AIC and likelihood ratio tests (Zuur et al 2009) to determine the most appropriate model error structure for the development of a cobia headboat index. The results of these tests (Table 2) support the ZINB formulation (similar results were obtained when using the full null model). These results concur with our expectations based on the over dispersion within the headboat data. A comparison between the fitted and original data for the ZIP and ZINB model formulations is shown in Figure 16. The rootogram (Kleiber and Zeileis 2017) in the lower panels of Figure 16 extends the Tukey (1977) rootogram to regression models. These plots are useful as diagnostics specific to overdispersion and/or excess zeros in count data models.

We used a step-wise forward model selection procedure to systematically include important covariates in our model formulation. In this procedure, we added each explanatory variable one at a time, alternating between the count (negative binomial) and binomial components. The variable with the largest ΔAIC was added, and the process repeated until no variable resulted in ΔAIC>2. The final cobia ZINB model formulation included year, party size, season and region in the negative binomial component, and year, region, percent full and season in the binomial component,

$$Count = y + t + s + v + a \mid y + s + v + t + i \qquad (3)$$

Diagnostics of the final model showed no clear patterns of association between Pearson's residuals and fitted values, or between the fitted values and original data (Figure 17) indicating acceptable model choice (Zuur et al 2009). Finally, a comparison of predicted values against the original data distribution (Figure 18) demonstrates how the model fits the original data.

All data manipulation and analysis was conducted using R version 3.5.0 (R Core Team 2018). Modeling was executed using the **zeroinfl** function in the **pscl** package (Kleiber and Zeileis 2017), available from the Comprehensive R Archive Network (CRAN).

Uncertainty

Uncertainty in the index was computed using a bootstrap procedure with *n*=1000 replicates. In each replicate, a data set of the original size was created by drawing observations (rows) at random with replacement. This was done by year, to maintain the same annual sample size as in the original data. The model (Equation 3) was fitted to each data set, and uncertainty (CVs) was computed from those fits that converged.

Changes and justification made at DW

During the DW, trip type (full day and half day trip) was included as a covariate in the final model run and was very similar to the initial index that only included full day trips. By including half day trips, the bootstrap convergence rate increased from 74% to 98% and appears to reduce the possibly unrealistic changes in population size in a few years while the average proportion positive decreased from 10% to 6%.

Results and discussion

Annual standardized index values for cobia including CVs are presented in Table 3. The relative nominal index fell within the 2.5% and 97.5% confidence intervals of the standardized index and tracked closely with the standardized index (Figure 19).

Literature cited

Hall, D. B. 2000. Zero-Inflated Poisson binomial regression with random effects: a case study. Biometrics, 56: 1030-1039.

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

Tukey J.W. 1977. Exploratory Data Analysis. Addison-Wesley Publishing Company, Phillipines.

Zeileis A. and C. Kleiber. 2017. countreg: Count Data Regression. R package version 0.2-0/r34, URL http://R-Forge.R-project.org/projects/countreg/.

Zeileis A., C. Kleiber, and S. Jackman. 2008. "Regression Models for Count Data in R." Journal of Statistical Software, 27(8), 1–25. doi: 10.18637/jss.v027.i08.

Zuur, A.F., E.N. Ieno, N.J. Walkder, A.A. Saveliev, and G.M. Smith. 2009. Mixed Effects Models and Extensions in Ecology with R. Spring Science and Business Media, LLC, New York, NY.

Table 1. Total number of headboat trips and positive cobia trips in the south Atlantic by year and region for the raw data compared to the model input (St. Mary's GA.- Murrell's Inlet SC (1), Little River, SC – Carolina Beach NC(2), Masonboro Inlet NC – Oregon Inlet NC (3)) Confidential data.

	Inlet Region						Inlet Region					
Year	1	2	3	raw data. N.trips	Raw data N.Fish	Raw Data Proportion Positive	1	2	3	Core Vessel N.trips	Core Vessel Proportion Positive	Core Vessel N.fish
1978	[redacted]			1	20	0.0%	[redacted]			0	0.0%	0
1979	[redacted]			8	11	0.3%	[redacted]			2	0.4%	5
1980	[redacted]			9	12	0.3%	[redacted]			4	0.9%	4
1981	[redacted]			26	30	1.2%	[redacted]			23	5.5%	25
1982	[redacted]			33	46	1.2%	[redacted]			27	4.5%	40
1983	[redacted]			34	52	1.5%	[redacted]			22	4.1%	40
1984	[redacted]			29	36	1.3%	[redacted]			13	2.4%	18
1985	[redacted]			35	46	1.5%	[redacted]			17	2.7%	26
1986	[redacted]			62	74	2.4%	[redacted]			31	4.9%	38
1987	[redacted]			78	114	2.7%	[redacted]			52	5.3%	80
1988	50	4	15	69	110	2.6%	36	4	12	52	4.4%	82
1989	21	7	10	38	53	2.1%	15	5	7	27	3.4%	35
1990	[redacted]				72	2.0%	[redacted]				2.0%	24
1991	55	8	38	101	176	4.9%	26	6	30	62	5.9%	115
1992	52	22	57	131	255	5.2%	28	17	39	84	7.0%	134
1993	66	14	54	134	203	5.0%	35	13	42	90	6.6%	124
1994	55	21	55	131	224	5.4%	20	15	37	72	5.9%	98
1995	79	18	74	171	268	6.4%	39	15	61	115	8.9%	178
1996	53	7	28	88	112	3.6%	22	7	18	47	3.9%	57
1997	63	24	37	124	214	4.7%	27	18	16	61	4.8%	75
1998	68	35	29	132	205	5.0%	27	27	15	69	5.8%	92
1999	59	14	27	100	149	4.3%	39	12	13	64	5.4%	90
2000	62	22	31	115	183	4.4%	30	14	15	59	4.6%	94
2001	67	27	20	114	207	5.0%	34	19	5	58	5.2%	69
2002	82	34	28	144	267	7.0%	45	19	11	75	7.2%	111
2003	45	26	21	92	151	4.6%	32	17	16	65	5.8%	82
2004	53	36	40	129	186	5.6%	23	25	24	72	5.5%	99
2005	41	23	25	89	139	4.9%	27	18	17	62	6.4%	98
2006	35	25	24	84	105	4.1%	18	17	20	55	5.0%	70
2007	78	70	15	163	288	7.5%	45	50	15	110	9.5%	154
2008	41	63	26	130	201	6.3%	26	42	17	85	8.7%	112
2009	34	46	4	84	126	3.9%	25	16	3	44	5.1%	59
2010	38	37	11	86	113	3.5%	28	23	6	57	5.1%	73
2011	16	56	11	83	106	3.8%	11	37	9	57	5.6%	74
2012	16	64	20	100	125	4.6%	11	39	10	60	6.5%	73
2013	35	78	33	146	202	6.8%	27	45	16	88	10.6%	130
2014	39	69	27	135	213	5.5%	26	37	10	73	7.6%	96
2015	24	42	16	82	119	3.3%	19	14	11	44	5.0%	75
2016	10	23	10	43	77	1.7%	6	9	5	20	2.7%	37
2017	[redacted]					0.0%	[redacted]				0.1%	[redacted]

Table 2: Preliminary model error structure comparison

	df	Likelihood	AIC	χ^2	df	*p*-value
ZIP	50	-7947	15995			
ZINB	51	-7820	15743	253	1	<0.001

Table 3: The relative nominal *Count*, number of trips, proportion positive, standardized index, and CV for the SEDAR 58 cobia index.

Year	Relative Nominal (Count)	N	Proportion Positive	Standardized index	CV
1991	1.13	603	0.10	1.03	0.16
1992	1.34	778	0.10	1.09	0.16
1993	1.06	846	0.12	0.97	0.18
1994	1.07	766	0.11	0.96	0.14
1995	1.30	850	0.15	1.16	0.17
1996	0.48	791	0.07	0.46	0.11
1997	0.81	788	0.10	0.77	0.15
1998	0.92	718	0.11	0.84	0.12
1999	0.89	727	0.10	0.93	0.12
2000	0.75	767	0.08	0.74	0.15
2001	0.83	678	0.09	0.79	0.16
2002	1.71	629	0.15	1.74	0.18
2003	0.99	617	0.10	0.95	0.16
2004	0.90	649	0.10	0.80	0.16
2005	1.11	504	0.11	1.15	0.14
2006	0.72	565	0.09	0.80	0.16
2007	1.46	624	0.18	1.62	0.18
2008	1.09	553	0.14	1.20	0.10
2009	0.44	442	0.06	0.54	0.12
2010	0.45	518	0.06	0.45	0.20
2011	0.72	498	0.09	0.69	0.19
2012	0.93	400	0.13	1.05	0.17
2013	1.61	438	0.18	1.86	0.15
2014	1.40	499	0.13	1.49	0.12
2015	0.89	422	0.09	0.88	0.18

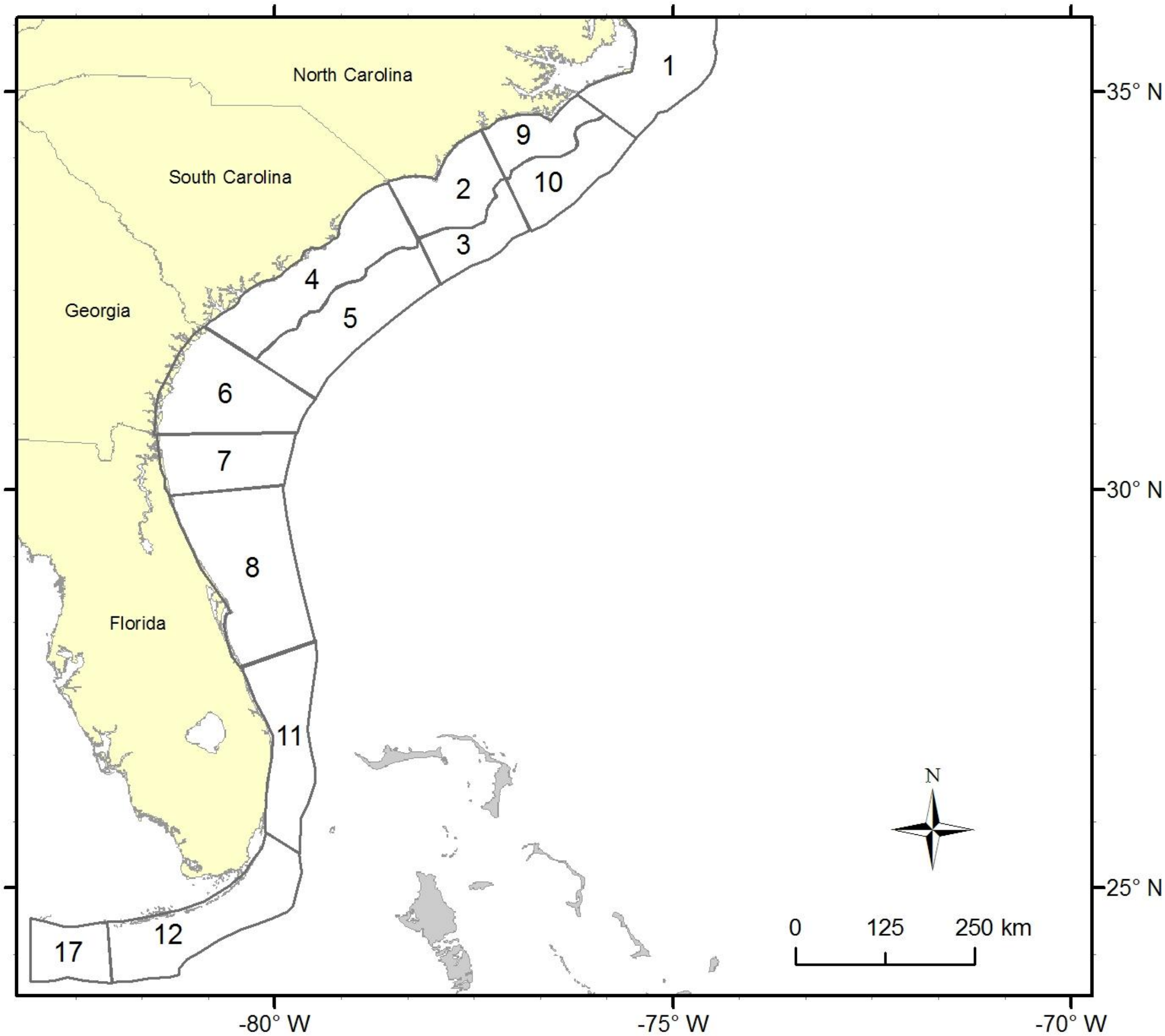

Figure 1. Map of headboat sampling area definition.

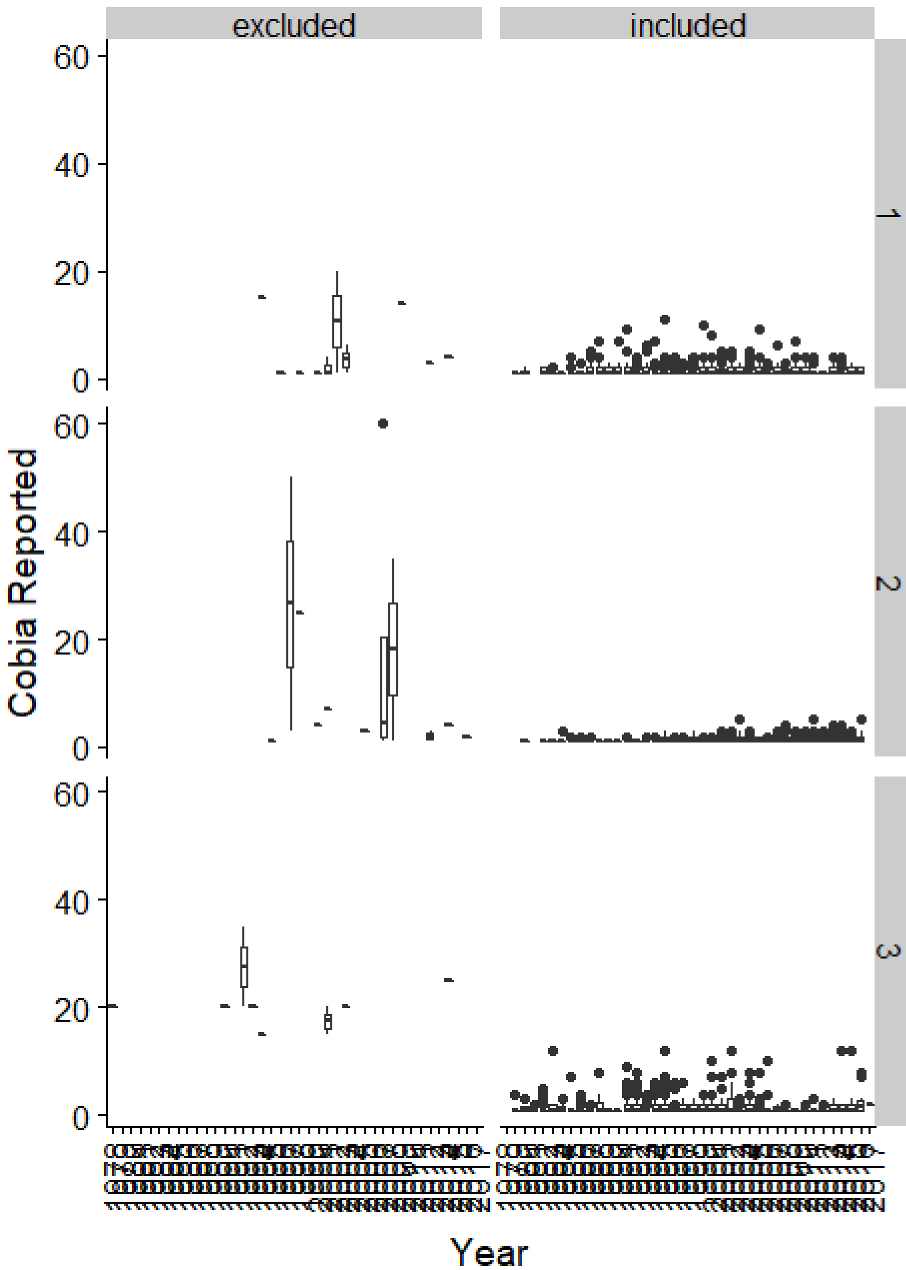

Figure 2. Records determined as outliers (excluded) based on removal of values above the 99.5th percentile for anglers and 99.9th percentile for number of fish caught.

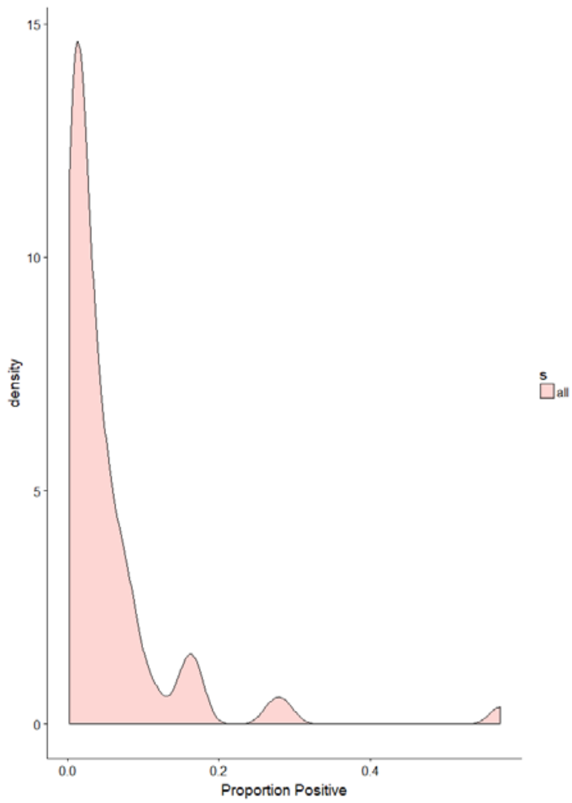

Figure 3. Density of proportion positive of positive cobia trips among vessels.

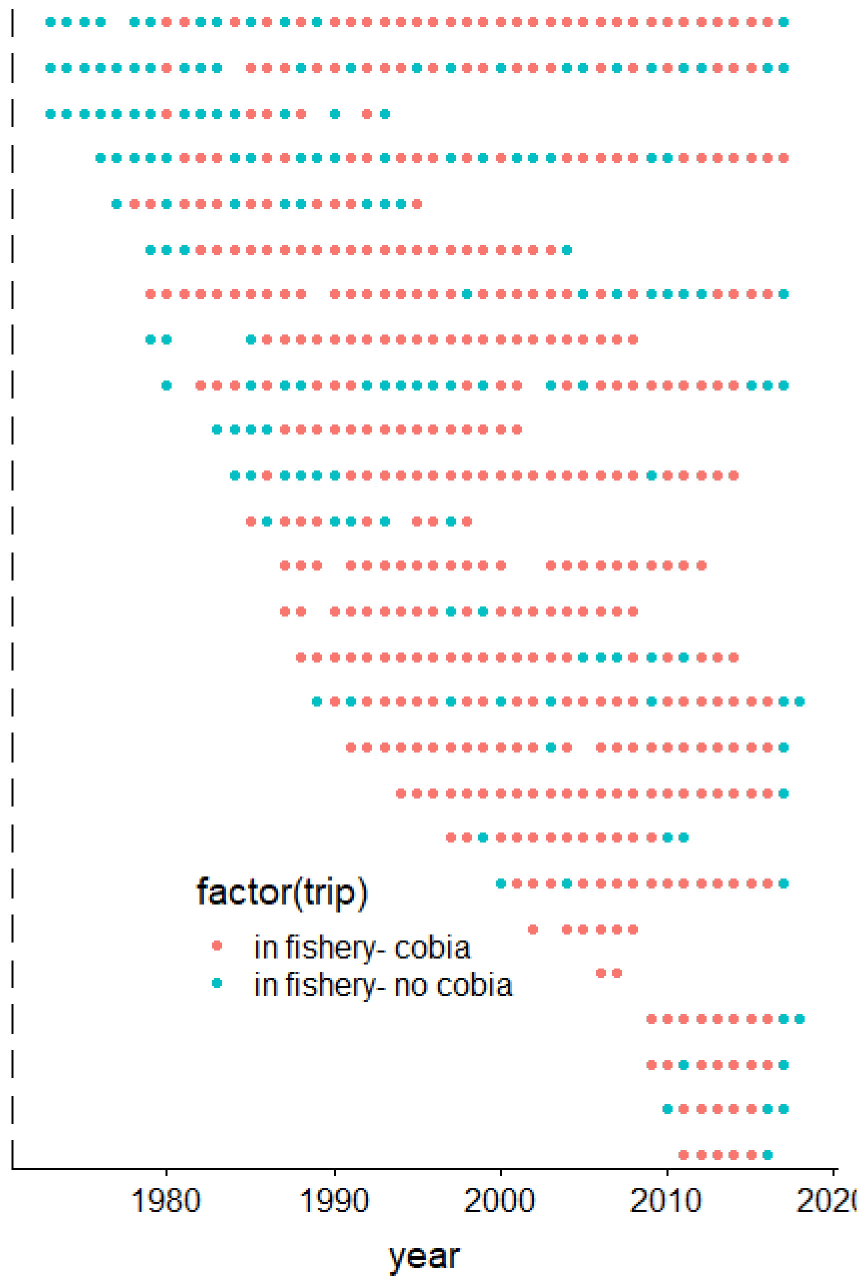

Figure 4. Each series of horizontal dots represents a ‘core’ vessel participation in the headboat fishery (blank=not active, red=in fishery-at least one positive cobia trips and blue=in fishery-no cobia.

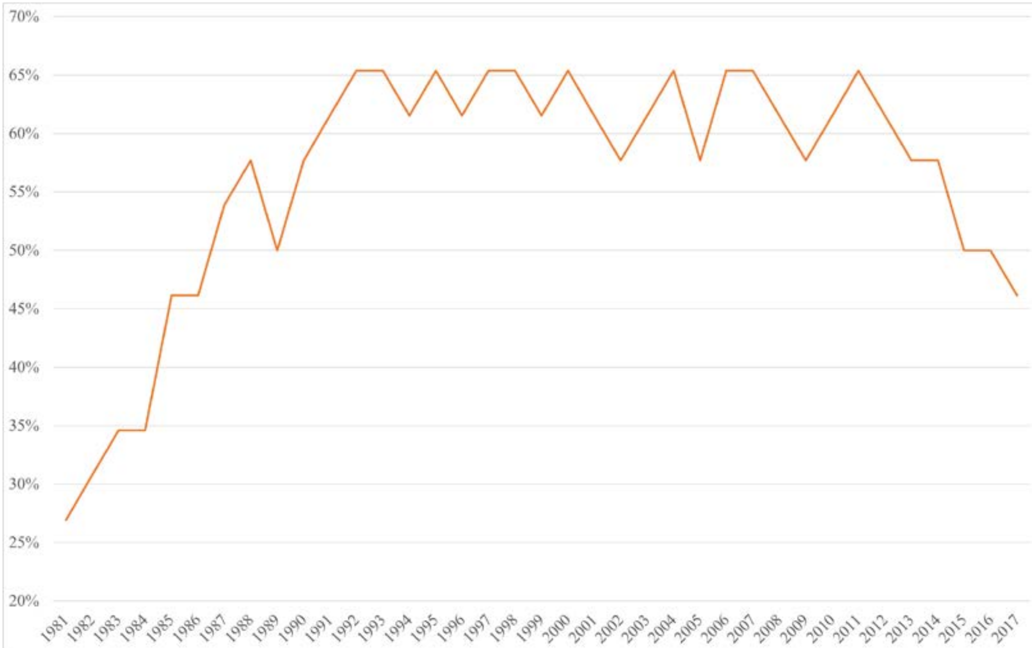

Figure 5. Proportion of active 'core' vessels participating in the headboat fishery by year.

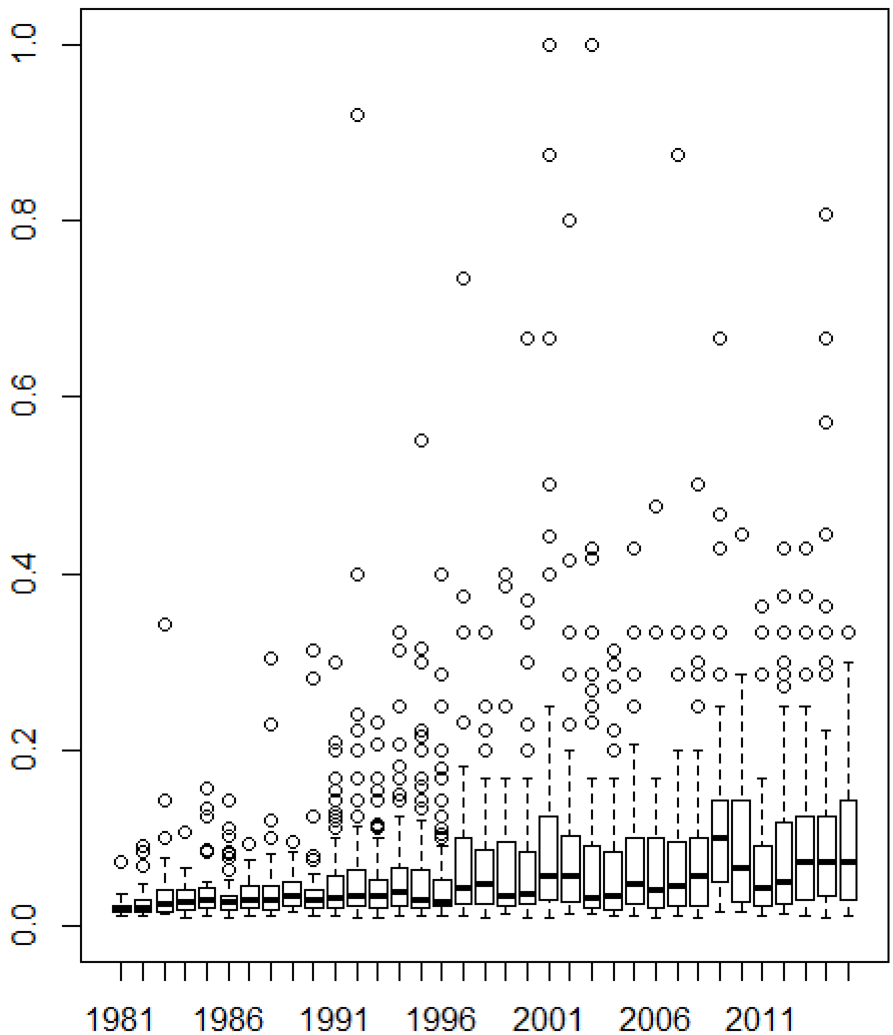

Figure 6. Box plot of cpue by year from headboat logbook records from 1981-2015 from NC to GA.

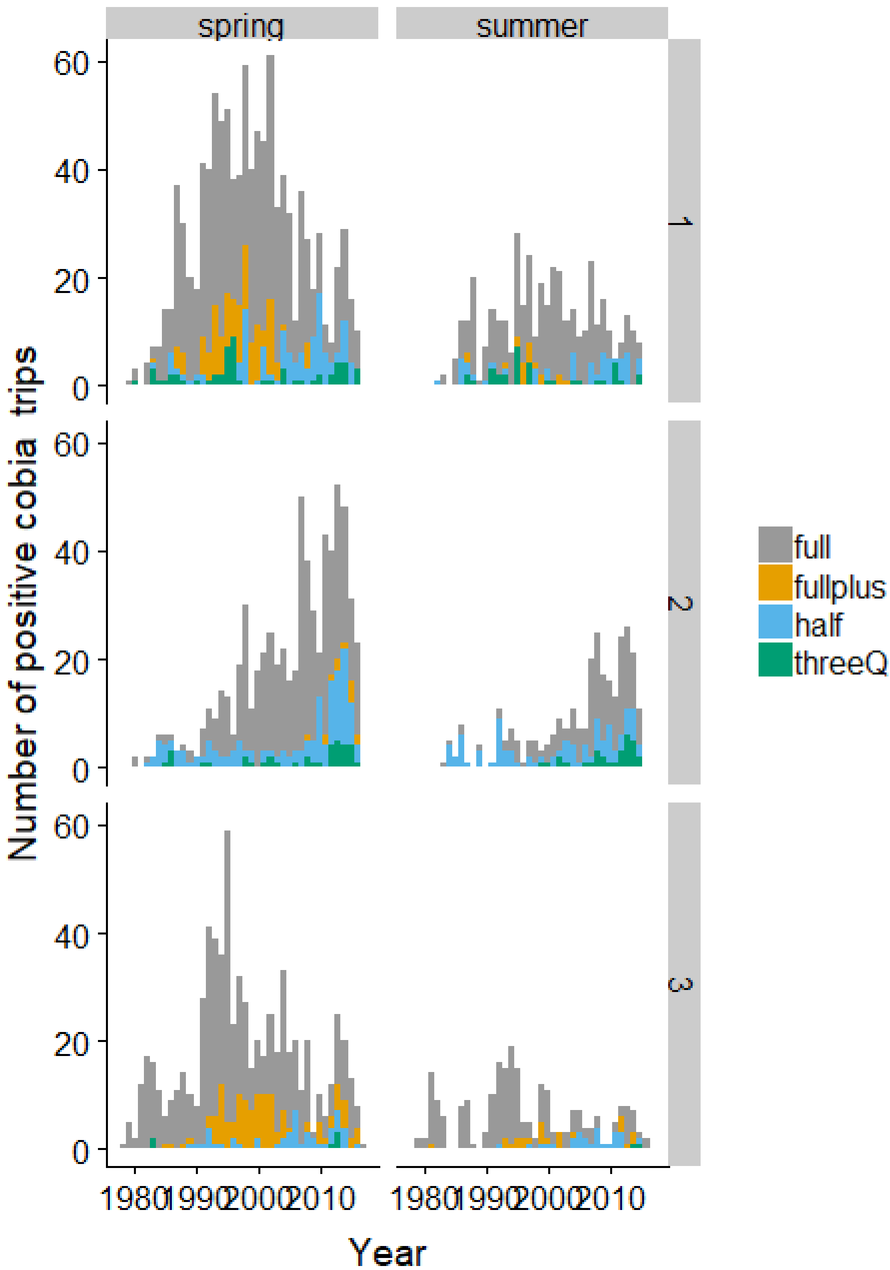

Figure 7. Unfiltered positive cobia trips by region and season prior to model input filtering ((St. Mary's GA.- Murrell's Inlet SC (1), Little River, SC – Carolina Beach NC(2), Masonboro Inlet NC – Oregon Inlet NC (3)).

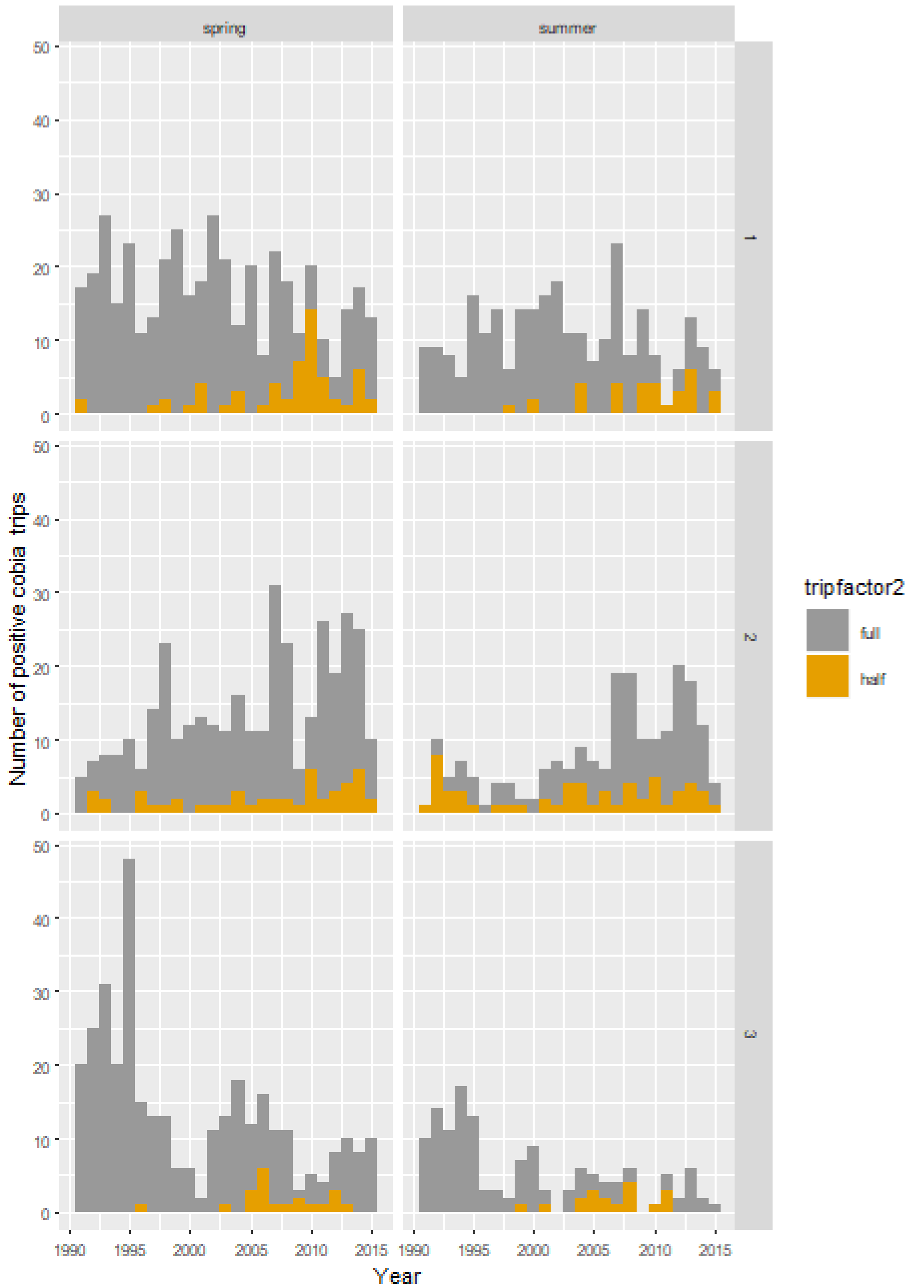

Figure 8. Model input of positive cobia trips by region and season (full and full plus trips only) ((St. Mary's GA.- Murrell's Inlet SC (1), Little River, SC – Carolina Beach NC(2), Masonboro Inlet NC – Oregon Inlet NC (3)).

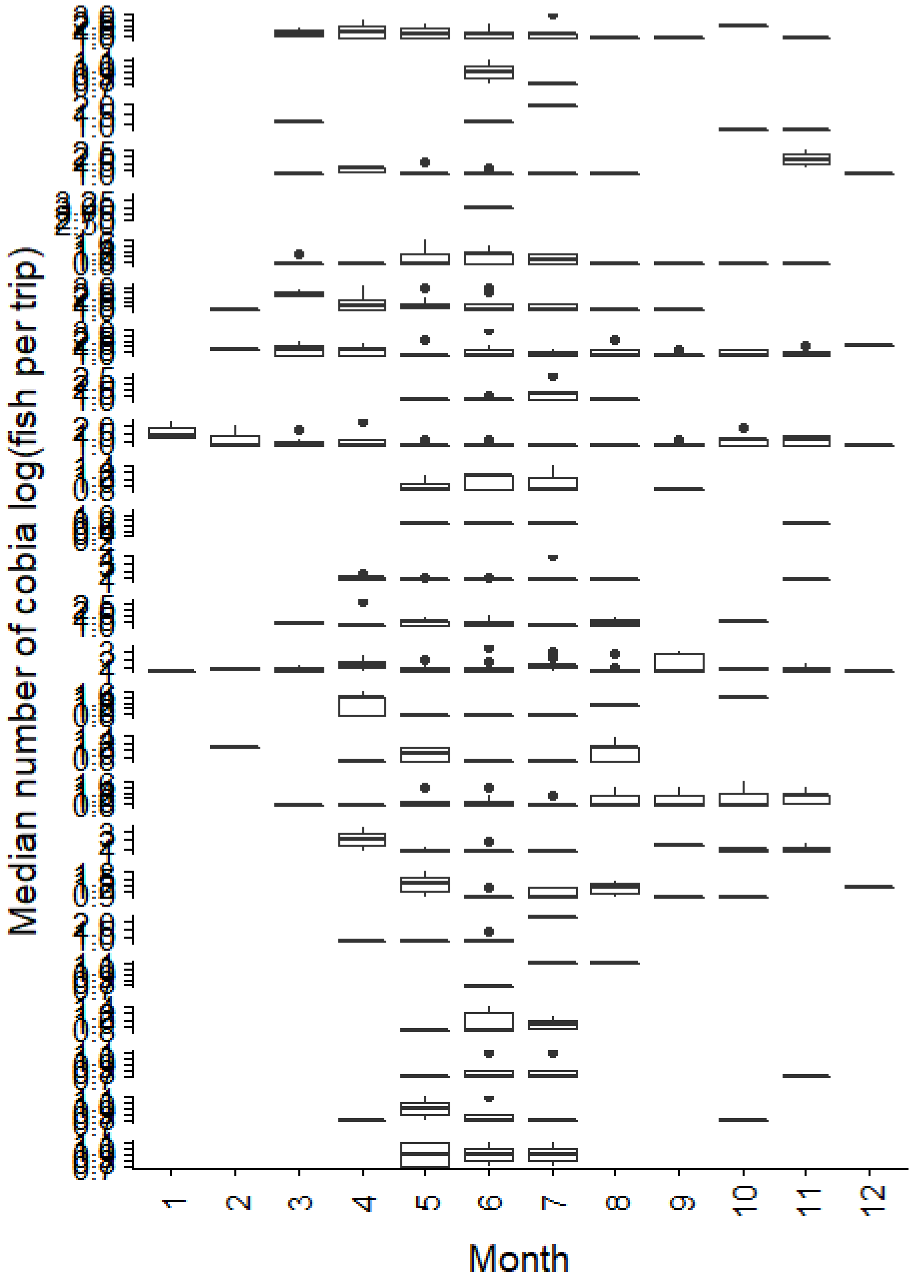

Figure 9. Boxplot of monthly mean cobia caught by ‘core’ headboats (prior to filtering months for model input) from Georgia to North Carolina.

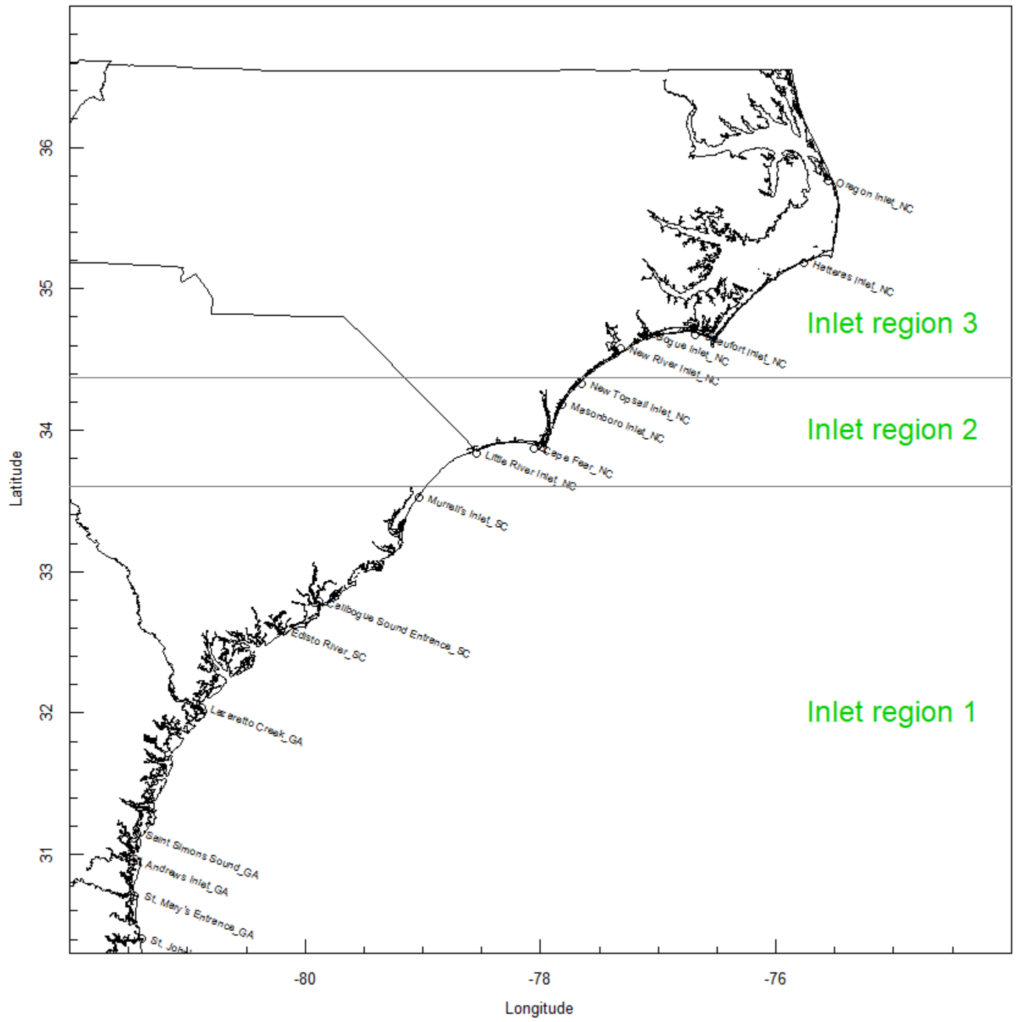

Figure 10. Illustration of individual inlets combined to form three regions for cobia index analysis (St. Mary's GA.- Murrell's Inlet SC (1), Little River, SC – Carolina Beach NC(2), Masonboro Inlet NC – Oregon Inlet NC (3)).

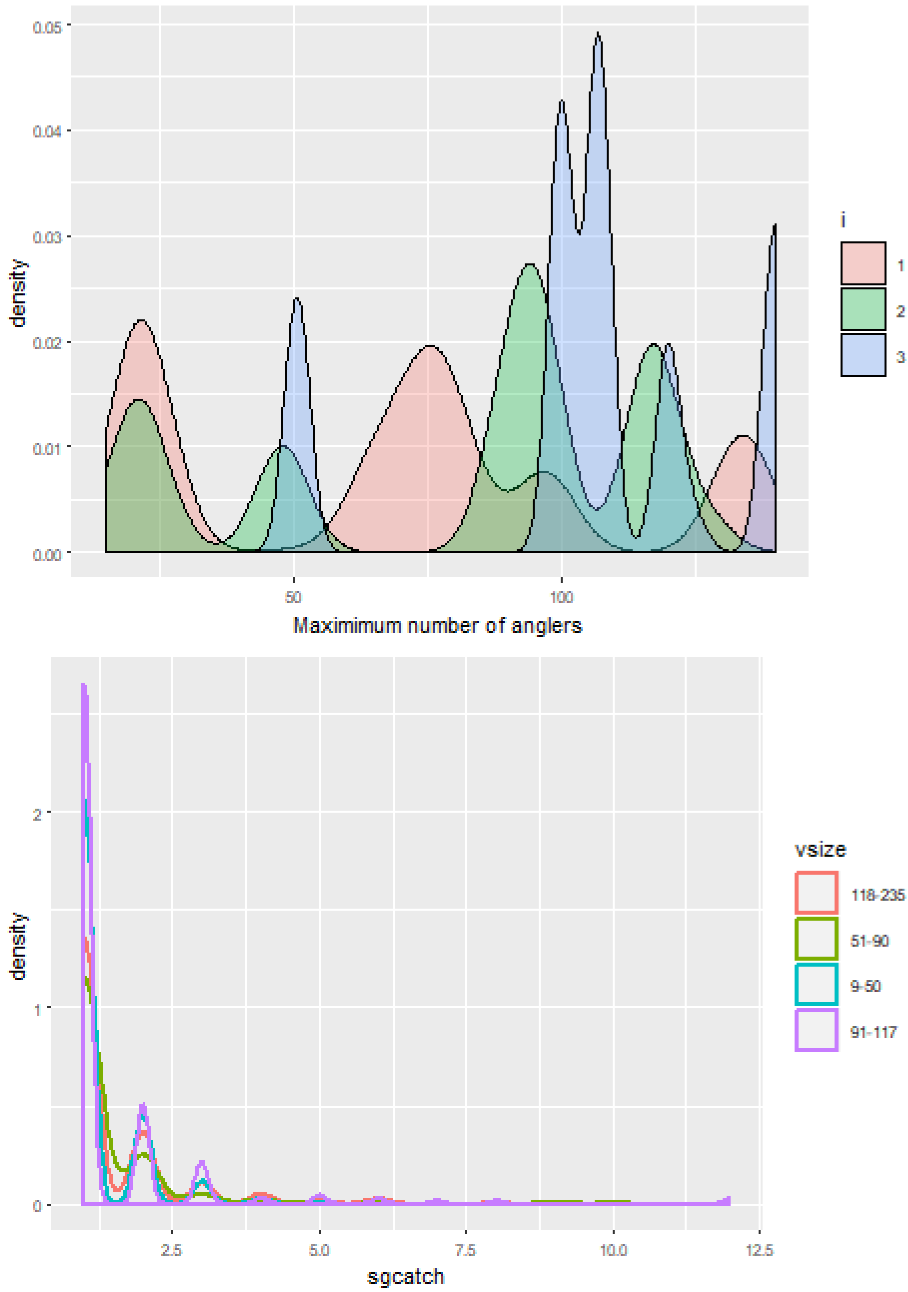

Figure 11. Density of maximum number of anglers across regions and cpue associated with the factors for maximum anglers as a proxy for vessel size. Due to convergence issues vessel size was simplified to 'small' and large'.

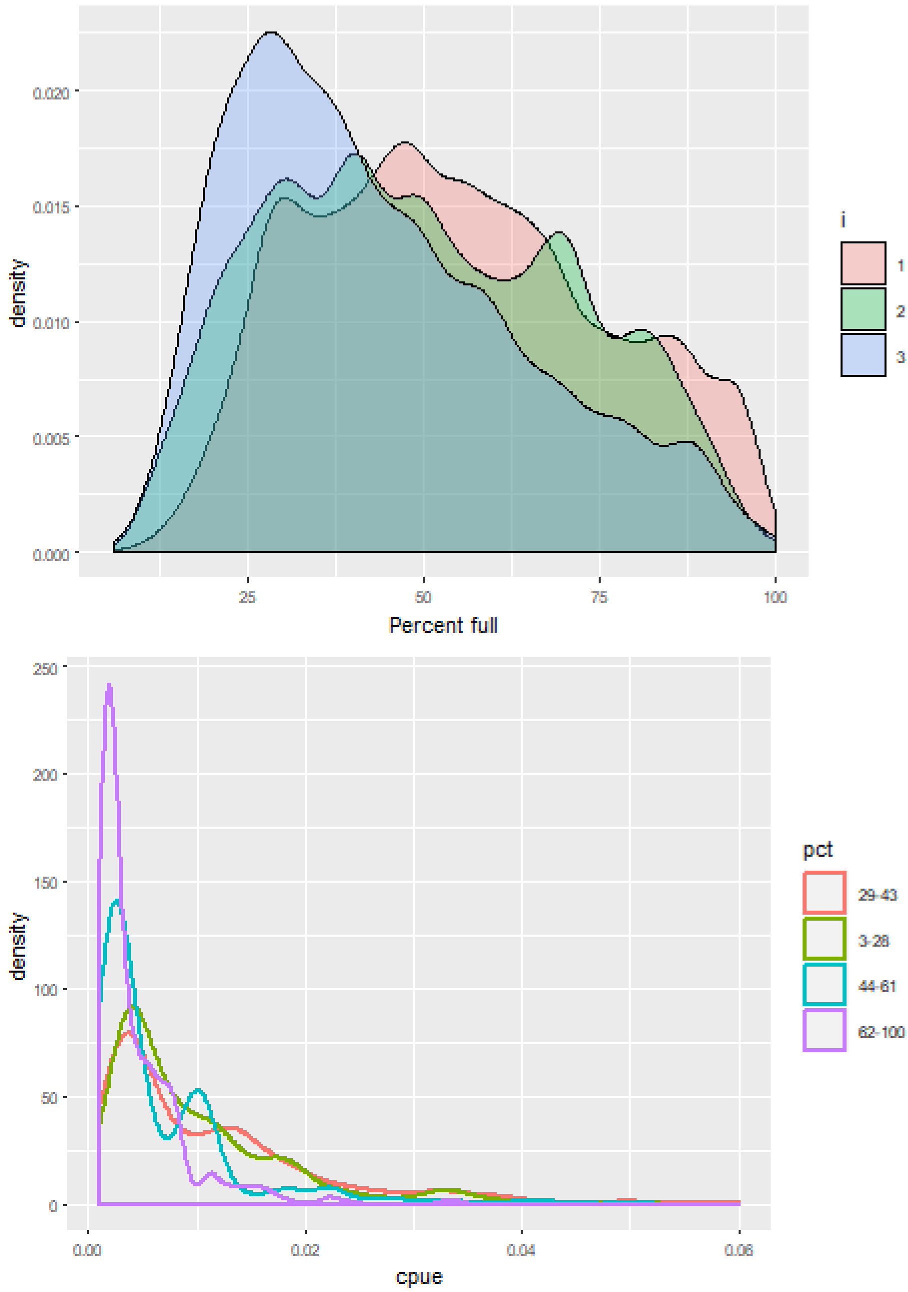

Figure 12. Density of percent full across regions and cpue associated with the factors for percent full.

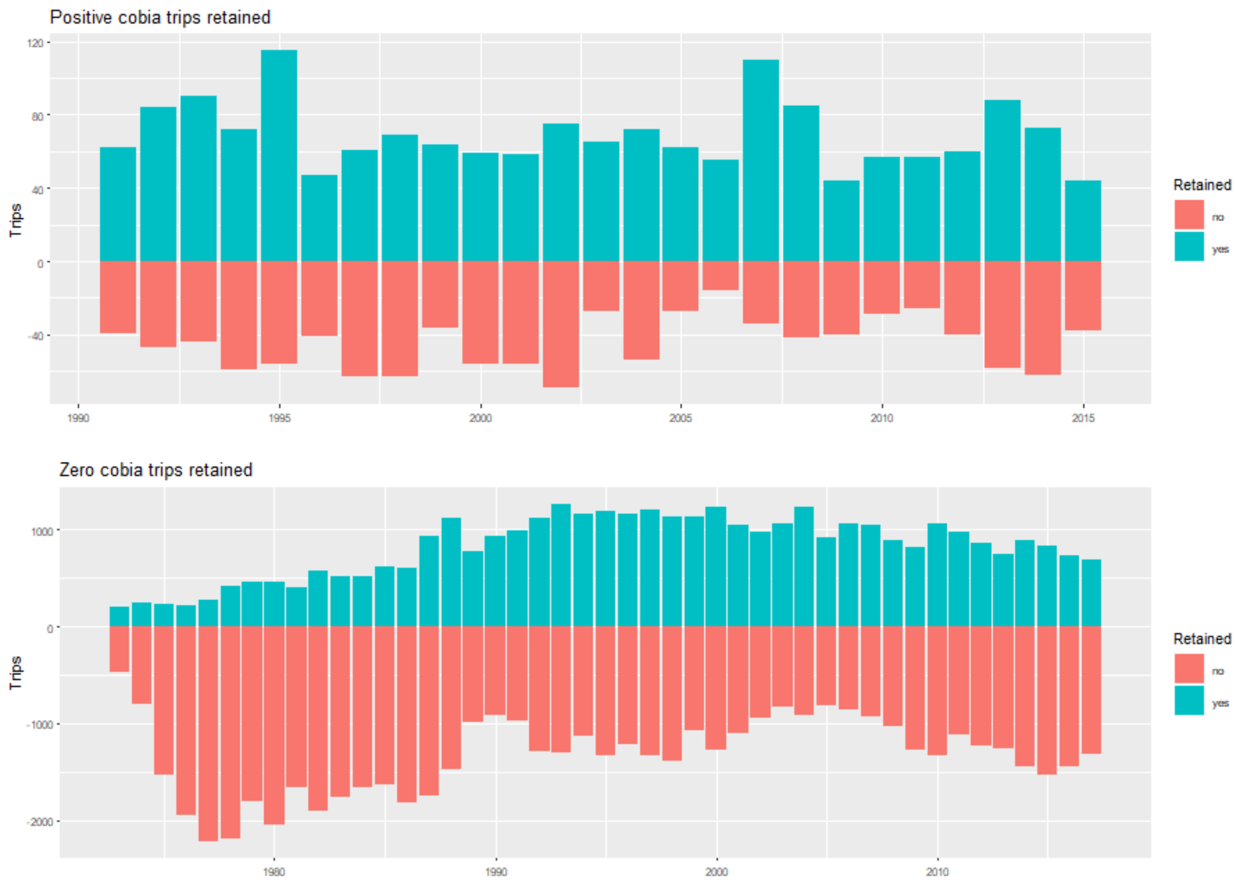

Figure 13. Positive and zero cobia trips retained after subsetting using ‘core’ vessels by year

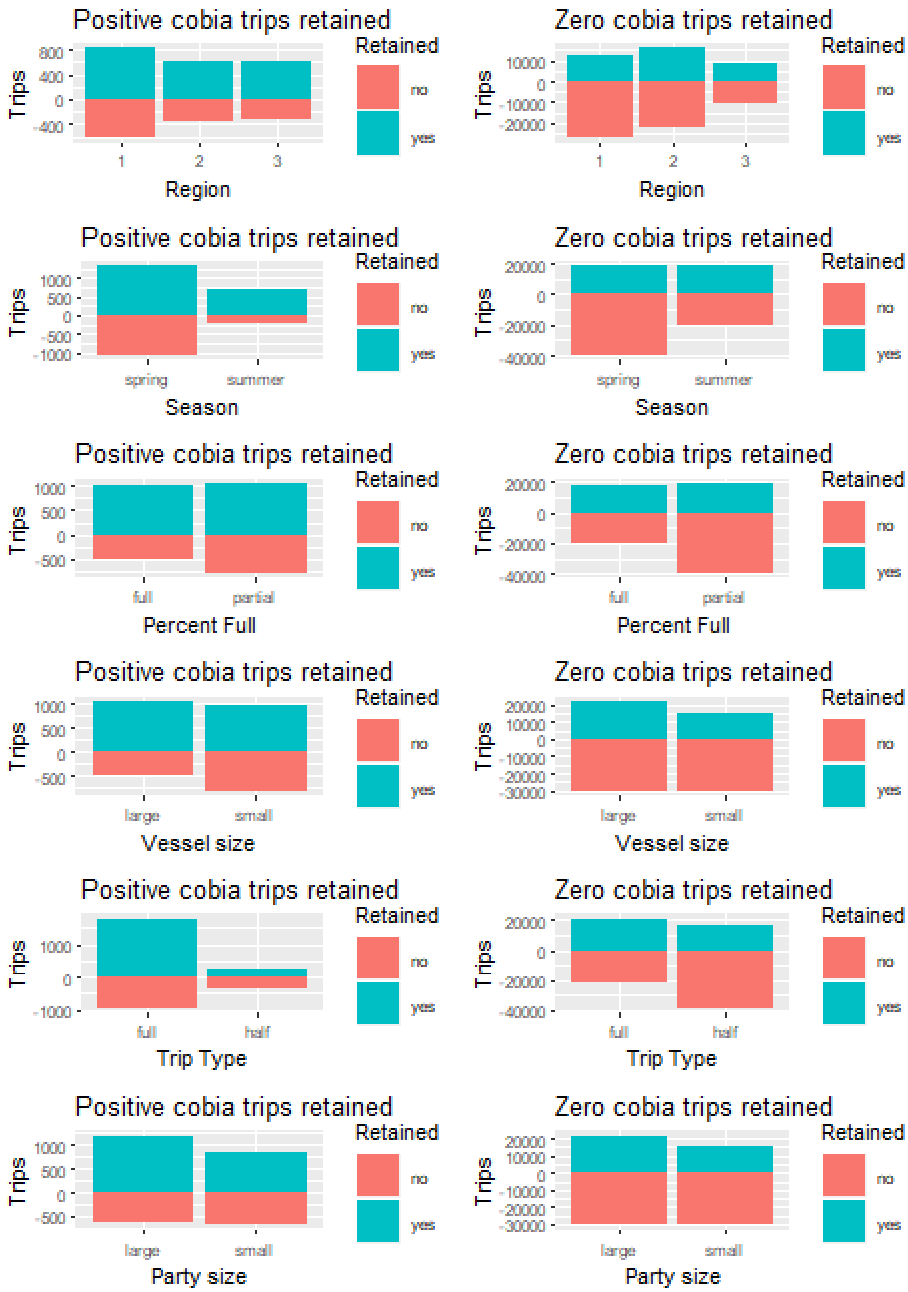

Figure 14. Positive and zero cobia trips retained after subsetting using 'core' vessels by factor. Note: Due to low proportion positive for half day trips they were filtered and trip type was removed as covariate in model.

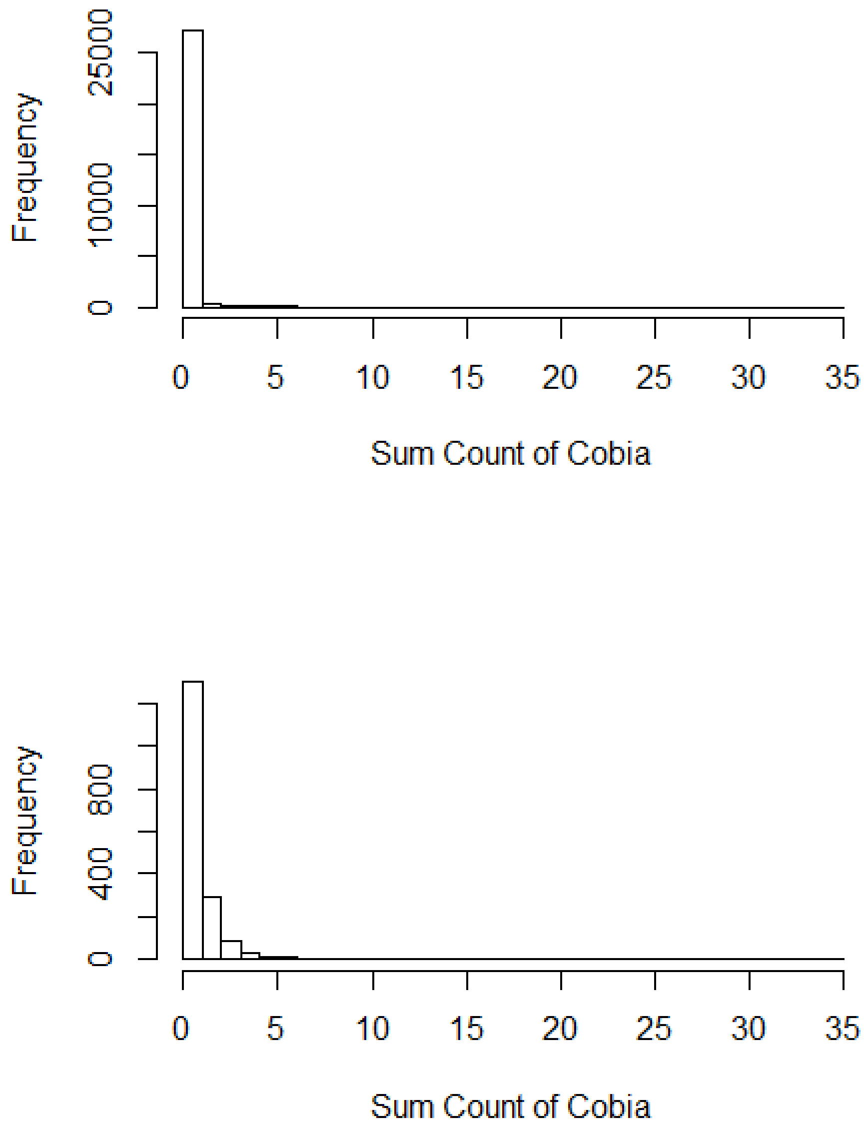

Figure 15. Count distribution of all ‘core’ cobia trips (top) and positive ‘core’ cobia trips (bottom) in the South Atlantic US headboat fishery (NC-GA).

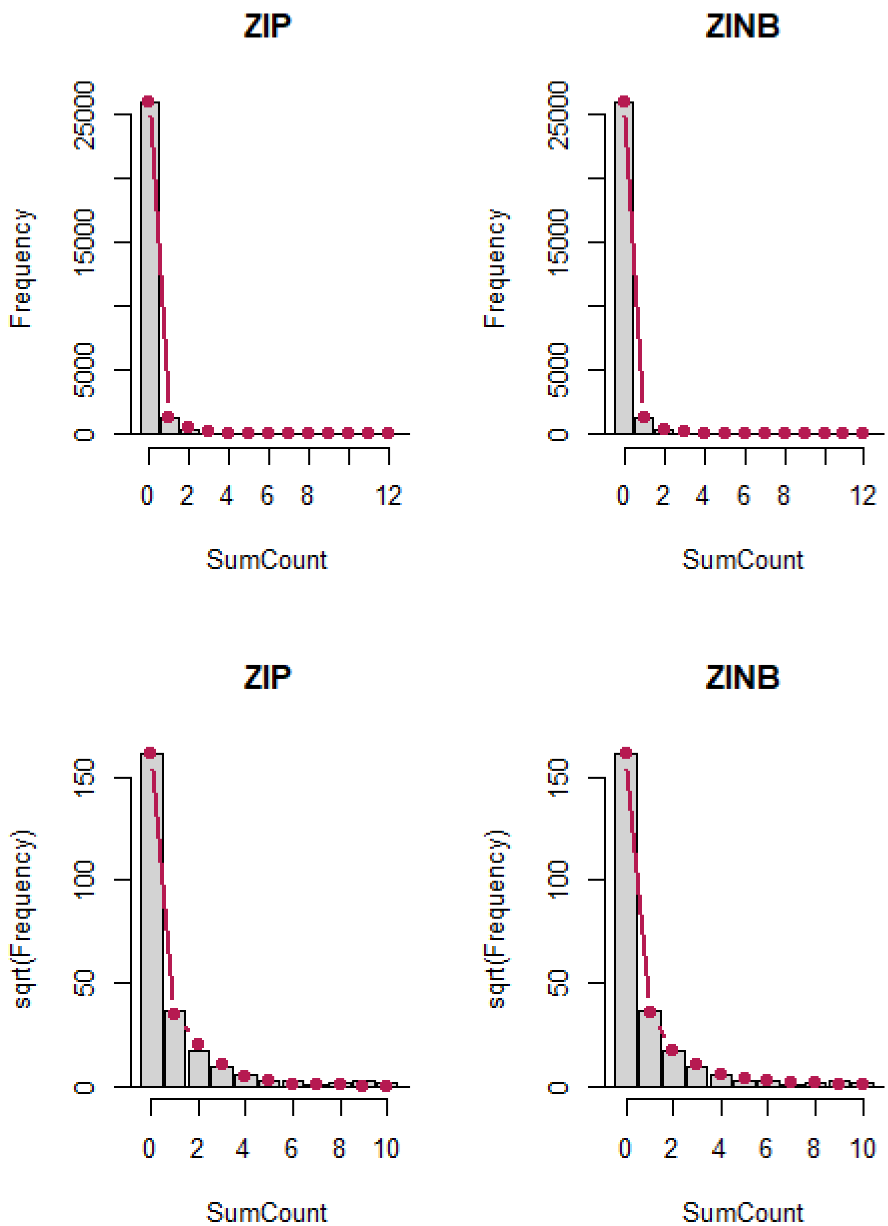

Figure 16. Model formulation comparison, with ZIP (left) and ZINB (right) fitted values plotted against the original data distribution with all covariates included. The lower panels are square root transformed and truncated at 20 fish for inspection of goodness of fit over the range of values for the bulk of the data.

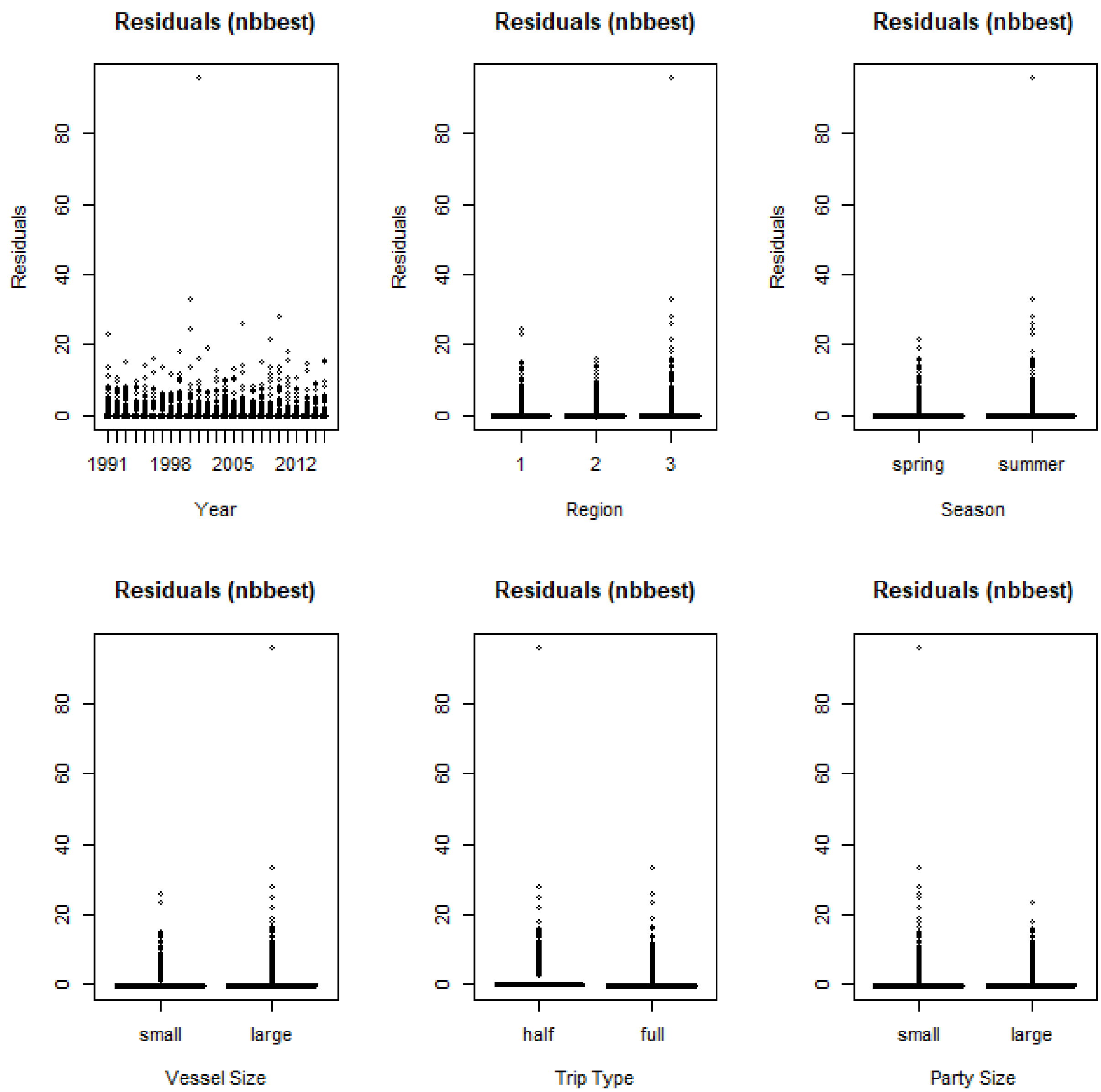

Figure 17. Model diagnostic plot showing residuals from final model (ZINB).

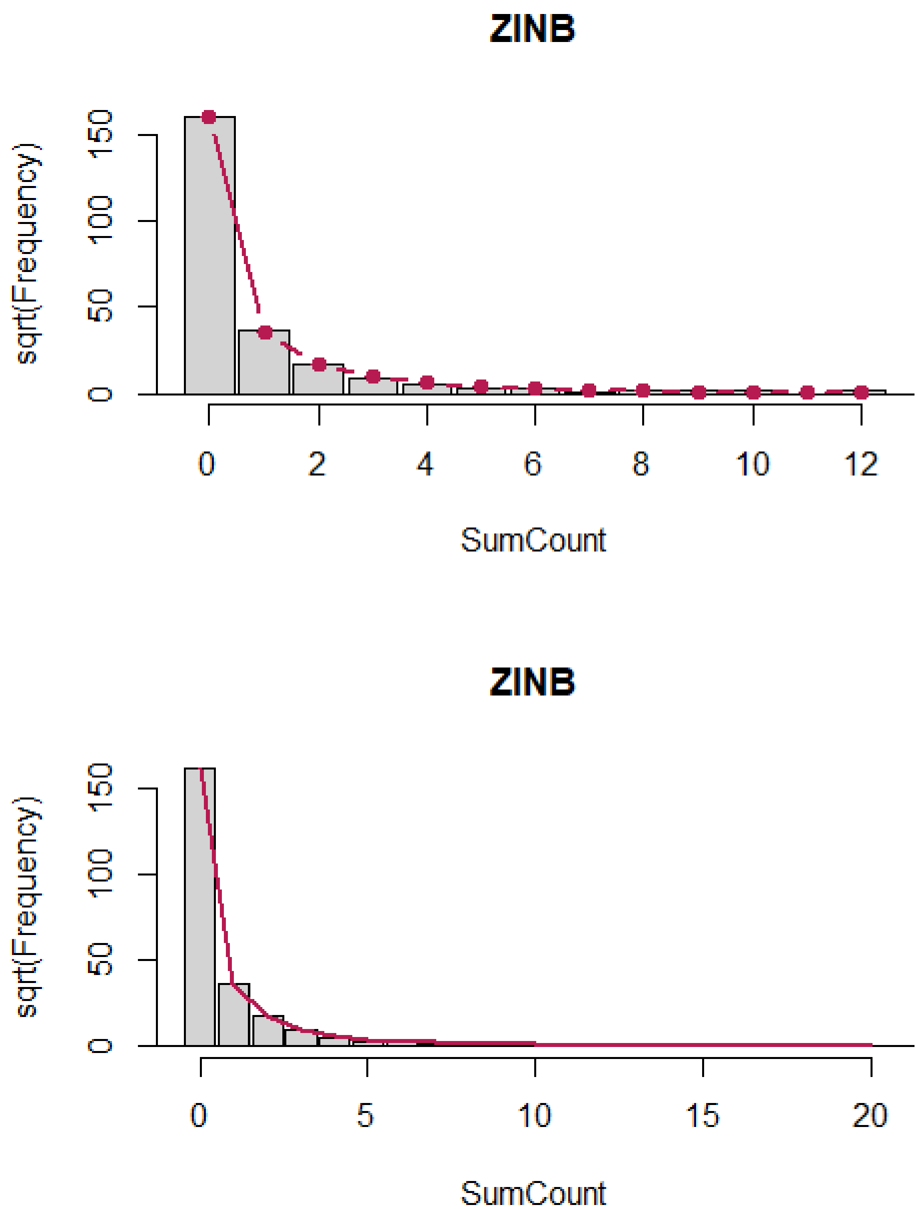

Figure 18. Model diagnostic plots of fitted model values (red line) against the original data distribution for the preferred model.

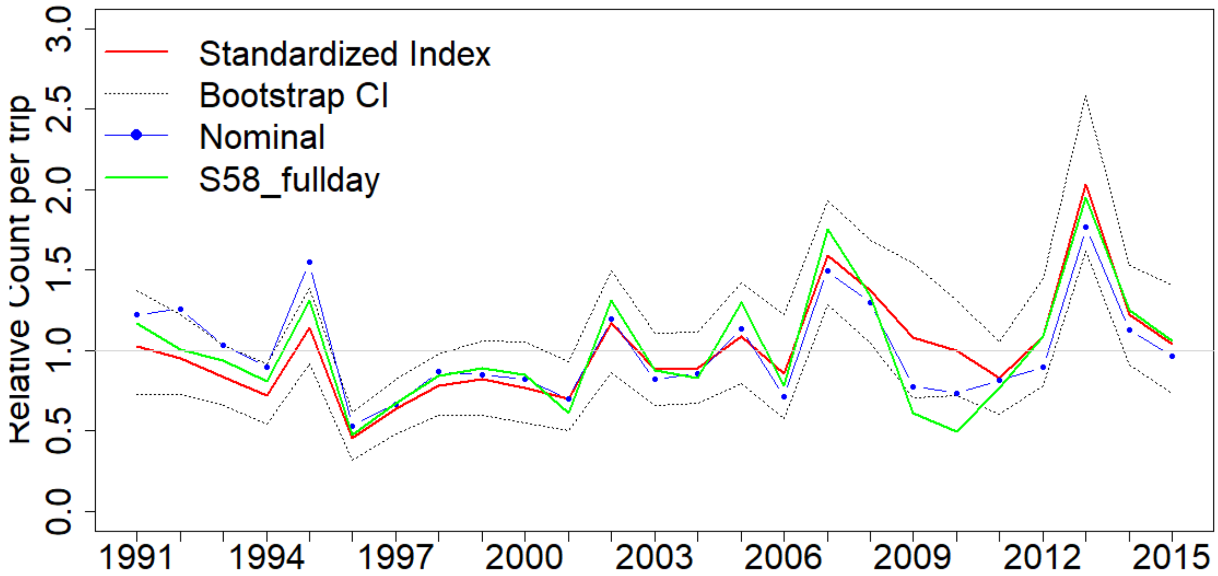

Figure 19. Relative standardized index (solid red line) with 2.5% and 97.5% confidence intervals (dashed lines) and the relative nominal index (blue) for cobia in the SRHS headboat logbook data.