Indices of abundance for Greater Amberjack (*Seriola dumerili*) using combined data from three independent video surveys

Kevin A. Thompson, Theodore S. Switzer, Mary C. Christman, Sean F. Keenan, Christopher Gardner, Katherine E. Overly, Matt Campbell

SEDAR70-WP-07

8 June 2020
Updated: 4 August 2020

Please cite this document as:

Thompson, Kevin A., Theodore S. Switzer, Mary C. Christman, Sean F. Keenan, Christopher Gardner, Katherine E. Overly, Matt Campbell. Indices of abundance for Greater Amberjack (*Seriola dumerili*) using combined data from three independent video survey. SEDAR70-WP-07. SEDAR, North Charleston, SC. 21 pp.

Indices of abundance for Greater Amberjack (*Seriola dumerili*) using combined data from three independent video surveys

Kevin A. Thompson[1], Theodore S. Switzer[1], Mary C. Christman[2], Sean F. Keenan[1], Christopher Gardner[3], Katherine E. Overly[3], Matt Campbell[4]

[1]Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute, St. Petersburg, FL

[2] MCC Statistical Consulting LLC, Gainesville, FL

[3] Southeast Fisheries Science Center, Panama City Laboratory, Panama City, FL

[4]Southeast Fisheries Science Center, Mississippi Laboratories, Pascagoula, MS

Introduction

Historically, three different stationary video surveys were conducted for reef fish in the northern Gulf of Mexico (GOM). The NMFS SEAMAP reef fish video survey, carried out by NMFS Mississippi Laboratory (Pascagoula), has the longest running time series (1993-1997, 2002, and 2004+), followed by the NMFS Panama City lab survey (PC; 2005+), with the most recent survey being the Florida Fish and Wildlife Research Institute video survey (FWRI, starting year 2100; Table 1). While the surveys use standardized deployment, camera field of view, and fish abundance methods to assess fish abundancies on reef or structured habitat, there are variations in survey design and habitat characteristics collected in addition to the time period and area sampled. Traditionally the surveys have submitted independent indices for each survey, however, combining indices across datasets likely increases predictive capabilities by allowing for the largest possible sample sizes in model fitting and encompassing a greater proportion of the distribution of the stock. Previous research has indicated that combining data across changing spatial areas and surveys and using a year only model, can yield spurious conclusions regarding stock abundance (Campbell 2004; Ye et al. 2004).

The standardization methods used in developing the index described in this paper have now been used in several SEDARs for the Gulf of Mexico (Thompson et al 2017; 2018; 2019). Specifically, for Greater Amberjack, the last benchmark assessment (SEDAR 2014) incorporated only the NMFS Pascagoula SEAMAP survey and the subsequent Update assessment (SEDAR 33 Update) combined the Pascagoula SEAMAP (Pascagoula), Panama City Laboratory (PC) surveys (SEDAR 2016). The SEDAR 33 Benchmark Assessment recommended to not use the FWRI survey at that time due to the short time series. The 2016 Update assessment specifically recommends investigation of methods to incorporate FWRI data as more years become available with the already used Pascagoula and PC datasets. As such, we used a habitat-based approach to combine relative abundance data for generating annual trends for Greater Amberjack (*Seriola dumerili*) throughout the GOM.

Survey Comparisons

Survey design

The Pascagoula survey primarily targets high-relief topographic features along the continental shelf from south Texas to south Florida (Table 2; Fig. 1). Sites are selected using a stratified, random design with strata determined by region and total proportion of reef area in a sampling block (10 minute latitude X 10 minute longitude blocks). Sites are selected at random from known reef areas identified through habitat mapping (multi-beam and side-scan sonar). This survey uses the Mississippi river delta as a geographic feature separating the west and east regions of the GOM (Campbell et al. 2017). Because of differences in spatial extent, habitat types and availability, and potential variation habitat association across regions, the east and west regions of this survey were treated as two surveys. This was done to yield more appropriate habitat models as well as appropriate weighting values in the final index values.

The Panama City video survey targets the inner shelf of the northeast GOM (5-60 m depth) ranging from NMFS, SEFSC statistical zone 6 through 10 (Table 2; Fig. 1). Survey design has changed through time, but since 2010 a two-stage unequal probability design has been used. Blocks are 5 minutes x 5 minutes in size with sites randomly, proportionally allocated by region, sub-region and depth. Two known reef sites, a minimum of 250 m apart within each selected block are randomly selected. This survey is broken up into eastern and western regions by Cape San Blas in the Florida Panhandle. Sites are described using side-scanning before video deployment (Gardner et al. 2017).

The FWRI survey initially focused on the regions offshore of Tampa Bay and Charlotte Harbor, FL (NMFS statistical zones 4 and 5) with habitats either inshore (10-36 m depth) or offshore (37-110 m depth). The survey has since expanded to also include NMFS, SEFSC statistical zones 9 and 10 off the Florida Panhandle in 2014 with additional sites added in 2016 to cover the entirety of the West Florida Shelf from statistical zones 2-10, although only data from statistical zones 4 and 5 are included in these analyses due to the short time series available (Table 2; Fig. 1). Sites are initially randomly selected and mapped using side scan sonar over a 2.1 km^2 area (Switzer et al. 2020). Video deployment sites are then randomly assigned proportionally across region and depth zones (Thompson et al. 2017).

Video reads

All three surveys use paired stereo-imaging cameras at each site. All videos are read to identify the maximum number of individuals of each species viewed in a single frame within a 20-minute time frame (i.e. MaxN, MinCount). Habitat characteristics on video are also noted with the percentage or presence/absence of abiotic and biotic habitat types that may contribute to fish biomass (e.g. sponge, algae, and corals), although some categories are not shared among all labs (Campbell et al. 2017; Gardner et al. 2017; Thompson et al. 2017).

Fish length measurement

Fish length measurements have varied through time for the surveys, starting with the Pascagoula survey in 1995 fish lengths were measured from video using lasers attached on the camera system with known geometry (Campbell et al. 2017). Panama City survey also used this laser-based approach from 2007 to 2009. However, the frequency of hitting targets with the laser is low and to increase sample size any measurable fish during the video read was measured (i.e. not just at the mincount), and fish could have potentially been measured twice. Subsequent years from (2008 in Pascagoula and 2010 in Panama City) used a stereo-video approach, which is the only method used in the entirety of the FWRI dataset. Vision Measurement System (VMS, Geometrics Inc.) was used to estimate size of fish up to 2014 for all three surveys and all switched to SeaGIS software (SeaGIS Pty. Ltd.) and have used them for the remainder of the timeseries.

Some species assessed with this combined approach have shown highly similar length compositions across the surveys (e.g. Red Grouper, Thompson et al. 2018). However, in some species, the ontogenetic shift from inshore to offshore is captured with variations in lengths by survey, with FWRI and Panama City capturing inshore, potentially younger and smaller individuals than the more offshore focused Pascagoula survey (Fig. 1; Carruthers et al. 2015; Switzer et al. 2015). In Greater Amberjack, this length variation across surveys is observed with higher frequency of smaller fish in Panama City and FWRI surveys (Fig. 2). As such, one previous assessment with a similarly disparate size compositions across surveys, for Vermilion Snapper used a multinomial approach to generate length compositions (Walter et al. 2020). The use of this method was initially investigated for Greater Amberjack, however the large size range of the species observed in the videos (Fig. 2) combined with the sample sizes of measurements and the bin size used in the assessment model yielded too many zero observations for models to be fit. Therefore, to account for variation in sample sizes in the surveys across the years of the index, the relative proportional contribution of each survey in terms of sample size, or number of video's analyzed, was used to adjust the overall length composition for this index over the time period from 1995-2018.

Data reduction

For all surveys, video reads were excluded if they were unreadable due to turbidity or deployment errors. For the Pascagoula survey, data included in this index are from 1993 and on, due to different counting methods in 1992. The entire spatial extent of the Panama City data was used from 2006 on with 2005 excluded because of an incomplete survey. For the FWRI data from prior to 2010 was excluded due to the earlier year's not including side-scan geoform as a variable which was determined to be potentially important as an explanatory variable in the analyses. FWRI data were spatially limited to zones 4 and 5 due to the other areas of the WFS not having enough years of sampling (Table 2).

Index Construction

Habitat models

To develop a single index of abundance for Greater Amberjack the data from all three surveys was, a habitat variable was created that included each of the separate survey individual variables that could be applied to all the data. This was done so final index models can account for changing sampling effort and habitat allocation through time rather than limiting the model to be predicted only by year and survey. We first determined the percentage of sites that occurred on good, fair, or poor (G, F, P) habitats for each survey independently. For this we used a categorical regression tree approach (CART) because this method accounts for correlations among variables and allows both continuous and categorical data to be included. It has been previously demonstrated to be a useful tool in fisheries ecology and specifically in describing fish-habitat associations (De'Ath and Fabricus 2000; Yates et al. 2016).

For these initial analyses, MaxN for each site was reduced to a presence and absence variable and was used as the response variable for habitat designations. Predictor variables included the habitat metrics coded on the video reads (reduced to presence/absence), the latitude and longitude of each site and depth for all three survey sets. For FWRI and Panama City's data, side-scan geoform was also included as a landscape-level habitat variable, with values derived using a modified version of the Coastal and Marine Ecological Classification Standard (CMECS) classification approach (habitats used in these analyses are in Table 3). Geoform was not included as a predictor variable for the analysis of MS survey data because the habitat mapping for that survey has primarily been conducted utilizing multibeam sonar, and at present, comparable habitat classification is not possible using the MS survey multibeam data. We first used a random forest approach to reduce the number of potential variables to be selected from in the final model for each lab's dataset to reduce redundant or correlated variables used in the final indexing model. For the random forest analysis, each survey was modeled separately for the entirety of that dataset. The random forest analysis fitted 2000 CARTS to the data and then determined each variables importance, a scale-less number used to indicate the number of final models each variable occurred in and its significance therein. An example of output is given in Fig. 3 for the FWRI survey dataset.

From the random forest analysis, approximately 50% of the potential variables were retained for each survey given by the importance values for a final CART model. The final model was created by fitting the presence of Greater Amberjack at site to the independent variables for a training dataset of 80% of the data. The remaining 20% of the data were retained in a test dataset to determine misclassification rates for each of the three models. The proportion of sites with positive Greater Amberjack catches at each terminal node was then evaluated to determine the habitat characteristics defining good, fair or poor habitat. Terminal nodes with double (2X) the overall proportion of positive catches for a dataset were assigned a good habitat code. Poor sites were identified as those determined by proportion positives that were at least half (50%) of the overall proportion positive and were generally approaching zero. The remaining sites were deemed fair and included the range of the overall proportion positive. All analyses were carried out using R version 3.0.2 (R Core Team 2014) and the Party package for CART (Hothorn et al. 2006).

CART results varied by survey with respect to the final variables chosen. Greater Amberjack habitat models indicated less of an association with factors commonly attributed to reef or rugose habitats, including rock and relief, as seen previously in other species (Thompson et al 2017; 2018; 2019). Primarily, the predictor variables were spatial (longitude, and latitude), related to the landscape Geoform (as for the PC survey), but did include some site-specific habitats for the PC survey data and the two Pascagoula regions (sponge, unknown sessile organisms, and seawhips) (Figs. 4-7). Greater Amberjack were found to be in a relatively low proportion of sites for FWRI survey (5.9%), moderate occurrence rates for PC (19.15%) and Pascagoula east (14.7%), and the highest in the Pascagoula west data (23.0%). The FWRI and Pascagoula west habitat models yielded only Fair and Good habitats, with the occupancy of Greater Amberjack not varying enough to predict Poor habitats with the variables used. These patterns are likely related to Greater Amberjack not being as specifically reef-obligates in comparison to previously assessed snappers and groupers (see CART models in Thompson et al. 2017, 2018).

The site characteristics that define each node and habitat code were then used to create a habitat variable (i.e., 'hab' and coded as: G or F or P) that was then back applied to each site for each of the three survey datasets. The datasets were then combined for the index model. The final proportion of sites in the three habitat categories for each individual survey set and year are shown in Table 4.

Index model fitting and diagnostics

The final model used to index abundance was fit using a negative binomial distribution with the formula:

*MaxN = Y*Hab *Survey*

Where Hab is the CART derived habitat code and survey represents the survey that collected the data for each site. Backwards variable selection was used and indicated that the full model performed best, given by AIC, compared to models with only one or two of the potential variables.

Model diagnostics indicated no discernible patterns of association between Pearson residuals and fitted values or the fitted values and the original data (Fig. 9), indicating correspondence to underlying model assumptions (Zuur et al. 2009).

The index was fit in SAS using the Proc GLIMMX procedure. To account for the variation in survey area, differences in area mapped with known habitat, and the distribution of Fair, Good, and Poor habitats by survey by year, the estimated MaxN means provided by the GLM were adjusted. The known potential survey universe for each of the three was first multiplied by the proportion of habitat mapping grids that had reef habitat to provide an area weight. This was then multiplied by each year x Survey X hab combination (up to 12 for the final years with three surveys and three habitat levels), providing a weighting factor for each of the mean estimates. Area weighting factors are provided in Table 3. Weighted index values were then standardized to the grand mean.

Results and Discussion:

Annual standardized indices for Greater Amberjack in the Gulf of Mexico, including coefficients of variation, are presented in Table 5. The model CV's indicate a good model fit, with highest values in earlier years ~20-35%, but somewhat decreasing CV's as additional surveys are added and continue. However, CVs and confidence limits were found to be high in 1997 compared to other years. Trends in standardized abundance for Greater Amberjack in the GOM show a relatively stable trend through time with small to moderate variation year-to-year with peaks in 1994, 1997 and 2009. The last five years of data show or predict a slightly negative overall trend (Table 5; Fig. 10). Given the utility of fishery-independent data in assessment and the potential increase in weight in the final assessment model of combined indices versus individual indices with smaller sample sizes and potentially divergent trends, we believe this approach is the most appropriate use of these video data. Furthermore, it is the most potentially representative of the entire GOM region, a priority discussed in previous SEDAR 33 Update assessment for this species (SEDAR 2016).

References Cited:

Campbell, R.A. 2004. CPUE standardization and the construction of indices of stock abundance in a spatially varying fishery using general linear models. Fisheries Research 70: 209-227.

Campbell, M.D., Kevin R. Rademacher, Michael Hendon, Paul Felts, Brandi Noble, Ryan Caillouet, Joseph Salisbury, and John Moser. 2017. SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Grey Snapper. SEDAR51-DW-07. SEDAR, North Charleston, SC. 31 pp.

Carruthers, T.R., J.F. Walter, M., M.K. McAllister, and M.D. Bryan. 2015. Modelling age-dependent movement: an application to red and gag groupers in the Gulf of Mexico. Canadian Journal of Fisheries and Aquatic Sciences 7: 1159-1176.

Gardner, C.L., D.A. DeVries, K.E. Overly, and A.G. Pollack. 2017. Gray Snapper Lutjanus griseus Findings from the NMFS Panama City Laboratory Camera Fishery-Independent Survey 2005- 2015. SEDAR51-DW-05. SEDAR, North Charleston, SC. 25 pp.

Hothorn, T, K. Hornik, and A. Zeileis. 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. Journal of Computational and Graphical Statistics 15: 651-674.

Jackman, S. 2008. Pack: Classes and methods for R developed in the political science computational laboratory, Stanford University. Department of Political Science, Stanford University, Stanford, CA.

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. URL: http://www.R-project.org/.

SEDAR. 2014. SEDAR 33 – Gulf of Mexico Greater Amberjack Stock Assessment Report. SEDAR, North Charleston SC. 490 pp. Available online at http://www.sefsc.noaa.gov/sedar/Sedar_Workshops.jsp?WorkshopNum=33

SEDAR. 2016. SEDAR 33 Stock Assessment Update Report Gulf of Mexico Greater Amberjack (Seriola dumerili). SEDAR, North Charleston SC. 490 pp. Available online at http://sedarweb.org/docs/suar/GAJ_S33_2016%20Update_Final.pdf

Switzer, T.S., E.J. Chesney, and D.M. Baltz. 2015. Habitat Use by Juvenile Red Snapper in the Northern Gulf of Mexico: Ontogeny, Seasonality, and the Effects of Hypoxia. Transactions of the American Fisheries Society 144:300-314.

Switzer, T.S, A.J. Tyler-Jedlund, S.F. Keenan, and E.J. Weather. 2020. Benthic habitats, as derived from classification of side-scan sonar mapping data, are important determinants of reef-fish assemblage structure in the Eastern Gulf of Mexico. Marine and Coastal Fisheries. 12:21-32.

Thompson, K.A., Theodore S. Switzer, and Sean F. Keenan. 2017. Indices of abundance for Gray Snapper (Lutjanus griseus) from the Florida Fish and Wildlife Research Institute (FWRI) video survey on the West Florida Shelf. SEDAR51-DW-10. SEDAR, North Charleston, SC. 22 pp.

Thompson, Kevin, A., Theodore S. Switzer, Mary C. Christman, Sean F. Keenan, Christopher Gardner, Matt Campbell. 2018. Indices of abundance for Red Grouper (Epinephelus morio) using combined data from three independent video surveys. SEDAR61-WP-03. SEDAR, North Charleston, SC. 18 pp.

Thompson, Kevin A. Theodore S. Switzer, Mary C. Christman, Sean F. Keenan, Christopher Gardner, Katherine E. Overly, Matt Campbell. 2019. Indices of abundance for Vermilion Snapper (Rhomboplites aurorubens) using combined data from three independent video surveys. SEDAR67-WP-03. SEDAR, North Charleston, SC. 17 pp.

Walter, J., Kevin Thompson and Ted Switzer. 2020. Model-based size composition of vermilion snapper obtained from three visual surveys. SEDAR67-WP-16. SEDAR, North Charleston, SC. 15 pp.

Yates KL, Mellin C, Caley MJ, Radford BT, Meeuwig JJ (2016) Models of Marine Fish Biodiversity: Assessing Predictors from Three Habitat Classification Schemes. PLoS ONE 11(6): e0155634. https://doi.org/10.1371/journal.pone.0155634

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. Summary of sample sizes by year for each of the three included video surveys, Florida Fish and Wildlife Research Institute (FWRI), NMFS Pascagoula, East and West regions, and NMFS Panama City. No data were available or used from any survey from 1998-2001; 2003.

Year	FWRI	Panama City	Pascagoula-East	Pascagoula-West	Total
1993			120	57	177
1994			99	61	160
1995			69	56	125
1996			140	172	312
1997			162	134	296
2002			152	108	260
2004			149	51	200
2005			274	140	414
2006		95	288	162	545
2007		59	330	192	581
2008		86	208	131	425
2009		108	263	183	554
2010	158	144	223	114	639
2011	222	158	349	105	834
2012	237	150	283	202	872
2013	185	97	167	145	594
2014	287	163	235	113	798
2015	224	168	152	59	603
2016	195	171	206	178	750
2017	154	150	222	211	737
2018	127	101	214	201	643
Total	1789	1650	4305	2775	10519

Table 2. Summary of sample sizes for each survey by SEAMAP statistical zone by year.

	Pascagoula East										Pascagoula West										
	Statistical Zones										Statistical Zones										
Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1993		14	4		13	60		13	8	8						24	33				
1994		22		9		41		17	5	5	9					17	24	11			
1995	4	18			5	23		11	4	4	4				6	15	31				
1996		34			7	53		14	12	20					8	49	97			13	5
1997		33	8		13	67		29	12						23	19	58	12		22	
2002		34	6		8	58		24	7	15	16		7		26	21	34			2	2
2004		26				59		28	17	19					3	4	33	10		1	
2005		57	10		10	71		61	18	47					17	44	42	24		13	
2006		52	22		10	79		32	36	57	10				15	41	48	20		10	18
2007		80	20		10	50		50	40	80	15				20	60	70			17	10
2008		48	23		8	66		11	12	40	6				12	45	52	5		11	
2009		61	13		6	87		25	28	43	9				25	64	57	8		20	
2010		61	18			33		30	37	44	9				7	8	46	15		22	7
2011		67	25		10	77		54	46	70					12	26	59	8			
2012		69	30		6	73		38	36	31			10		28	55	65	17		25	2
2013		4	31		9	47		18	19	39					8	42	54	17		16	8
2014		30	38		10	60		20	29	48				7	16	29	44	17			
2015		2	16	8	9	57		10	16	34	3		4	6	8	5	17	8			8
2016		39	34	10	10	48		10	20	35			8	13	25	16	59	12		9	36
2017		56	24	33	12	38		17	12	30	2			12	24	35	53	31		22	32
2018		46	12	13	23	46			22	52	10			10	6	34	47	39	10	10	35

	FWRI										Panama City				
	Statistical Zones										Statistical Zones				
Year	1	2	3	4	5	6	7	8	9	10	6	7	8	9	10
2006											17	52	13	13	
2007											13	28	8	10	
2008											16	38	11	20	1
2009											11	63	12	22	
2010				82	76						15	59	27	43	
2011				80	142						20	74	25	38	1
2012				107	130						20	79	23	28	
2013				70	115						19	45	9	23	1
2014				142	152				68	22	23	82	13	44	1
2015				92	182				99	55	25	81	23	39	
2016		83	74	83	142	87	69	67	68	56	25	86	19	41	
2017		66	87	86	96	70	79	44	69	32	27	69	11	43	
2018		103	85	68	112	95	68	66	113	60	19	52	12	18	

Table 3. Proportion of sites for each habitat level (**F**air**, G**ood**, P**oor) as determined by individual survery categorical regression trees (CARTs) for Greater Amberjack presence. Note the gap in sampling for the Pascagoula lab (1998-2002 and 2003).

Pascagoula East				**Pascagoula West**		
Year	F	G	P	Year	F	G
1993	0.78	0.03	0.18	1993	1.00	0.00
1994	0.65	0.05	0.30	1994	0.82	0.18
1995	0.68	0.00	0.32	1995	0.93	0.07
1996	0.75	0.04	0.21	1996	0.94	0.06
1997	0.73	0.02	0.25	1997	0.81	0.19
2002	0.57	0.19	0.24	2002	0.74	0.26
2004	0.77	0.06	0.17	2004	0.82	0.18
2005	0.72	0.07	0.22	2005	0.73	0.27
2006	0.73	0.01	0.26	2006	0.80	0.20
2007	0.68	0.03	0.29	2007	0.78	0.22
2008	0.66	0.01	0.33	2008	0.81	0.19
2009	0.70	0.02	0.28	2009	0.89	0.11
2010	0.64	0.01	0.35	2010	0.89	0.11
2011	0.72	0.03	0.25	2011	0.77	0.23
2012	0.62	0.03	0.35	2012	0.76	0.24
2013	0.77	0.05	0.19	2013	0.81	0.19
2014	0.70	0.01	0.29	2014	0.77	0.23
2015	0.80	0.05	0.16	2015	0.76	0.24
2016	0.59	0.03	0.38	2016	0.79	0.21
2017	0.54	0.00	0.46	2017	0.80	0.20
2018	0.64	0.01	0.35	2018	0.77	0.23
Panama City				**FWRI**		
Year	F	G	P	Year	F	G
2006	0.29	0.44	0.26			
2007	0.41	0.54	0.05			
2008	0.28	0.47	0.26			
2009	0.45	0.53	0.02			
2010	0.52	0.46	0.02	2010	0.84	0.16
2011	0.34	0.40	0.26	2011	0.92	0.08
2012	0.59	0.39	0.02	2012	0.96	0.04
2013	0.01	0.99	0.00	2013	0.94	0.06
2014	0.25	0.52	0.24	2014	0.94	0.06
2015	0.33	0.40	0.26	2015	0.89	0.11
2016	0.32	0.46	0.23	2016	0.94	0.06
2017	0.28	0.34	0.38	2017	0.96	0.04
2018	0.31	0.02	0.67	2018	0.91	0.09

Table 4. The habitat weighting used with the annual distribution of Fair, Good, Poor habitats to adjust estimated model means to account for sampling variation across surveys.

Survey	Total Universe Area (km2)	Proportion of grids with habitat	Total Universe Area X Prop transects	Area Weighting values (1993-2005)	Area Weighting values (2006-2009)	Area Weighting values (2010-2017)
Pascagoula E	34490	0.81	27936.9	0.707	0.514	0.429
Pascagoula W	31258	0.37	11565.46	0.293	0.213	0.177
PC	22104	0.67	14860.9	0.000	0.273	0.228
FWRI	37290	0.29	10814.09	0.000	0.000	0.166

Table 5. Number of stations sampled (N) by survey and year, proportion of positive sets, standardized index, and CV for the annual FWRI Greater Amberjack video index of the Gulf of Mexico.

year	N	Proportion positives present	Std Nominal	Std Index	CV
1993	177	0.062	0.671	0.621	0.282
1994	160	0.206	1.113	1.200	0.593
1995	125	0.248	1.013	0.738	0.292
1996	312	0.160	0.704	0.642	0.240
1997	296	0.142	0.842	1.011	0.731
2002	260	0.338	2.329	2.295	0.220
2004	200	0.180	0.900	0.788	0.243
2005	414	0.210	1.109	1.097	0.227
2006	545	0.125	0.890	0.744	0.192
2007	581	0.160	1.005	0.972	0.220
2008	425	0.160	0.973	0.922	0.252
2009	554	0.227	1.364	1.336	0.168
2010	639	0.156	0.771	0.756	0.215
2011	834	0.158	0.909	0.919	0.191
2012	872	0.211	1.119	1.150	0.155
2013	594	0.189	0.999	1.103	0.201
2014	798	0.153	1.118	1.183	0.229
2015	603	0.159	1.208	0.974	0.152
2016	750	0.156	0.839	1.037	0.381
2017	737	0.140	0.811	0.765	0.160
2018	643	0.137	0.742	0.745	0.323

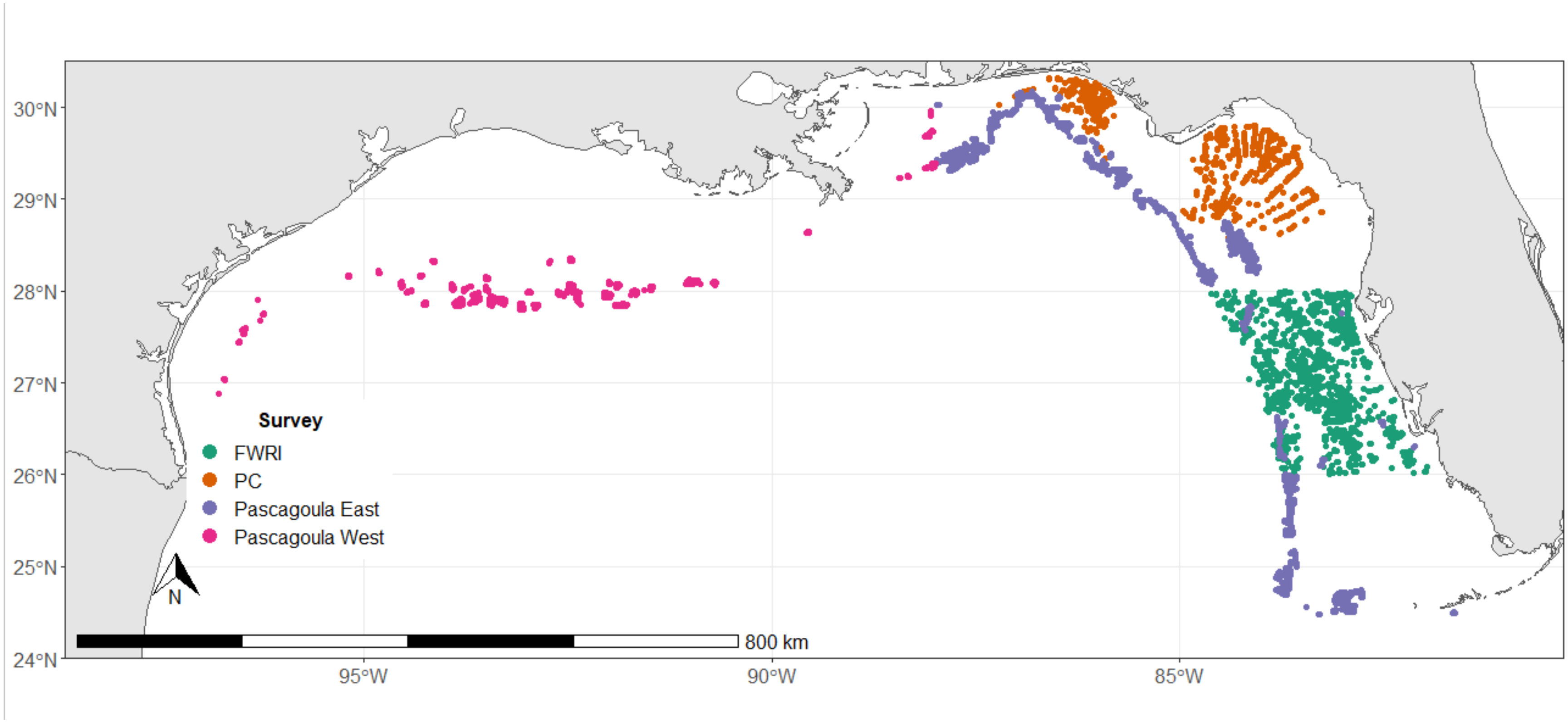

Figure 1. Map of all video sites included in the index for each survey (by laboratory) across all years 1993-2018.

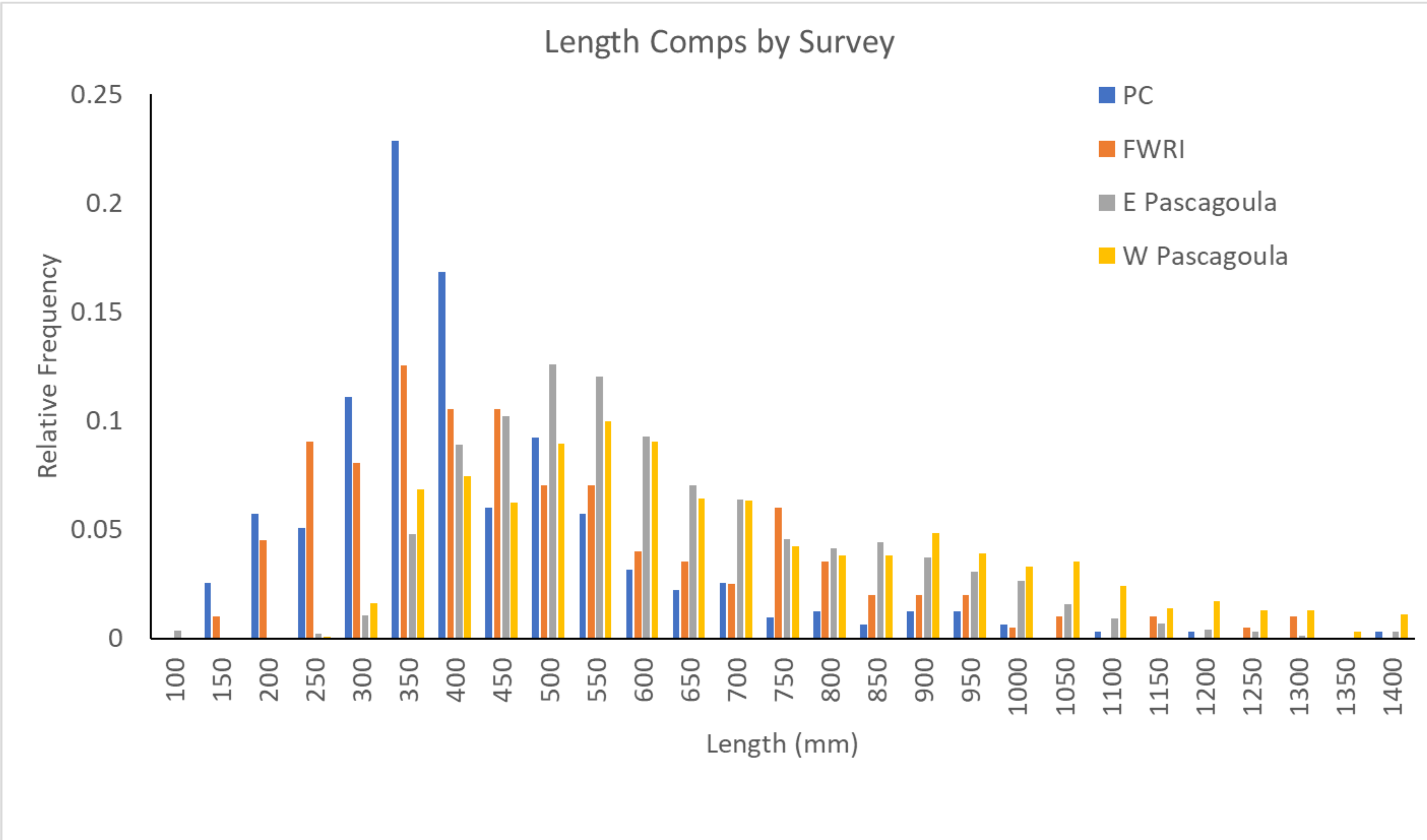

Figure 2. Nominal length compositions of the three surveys used in the combined index for Greater Amberjack.

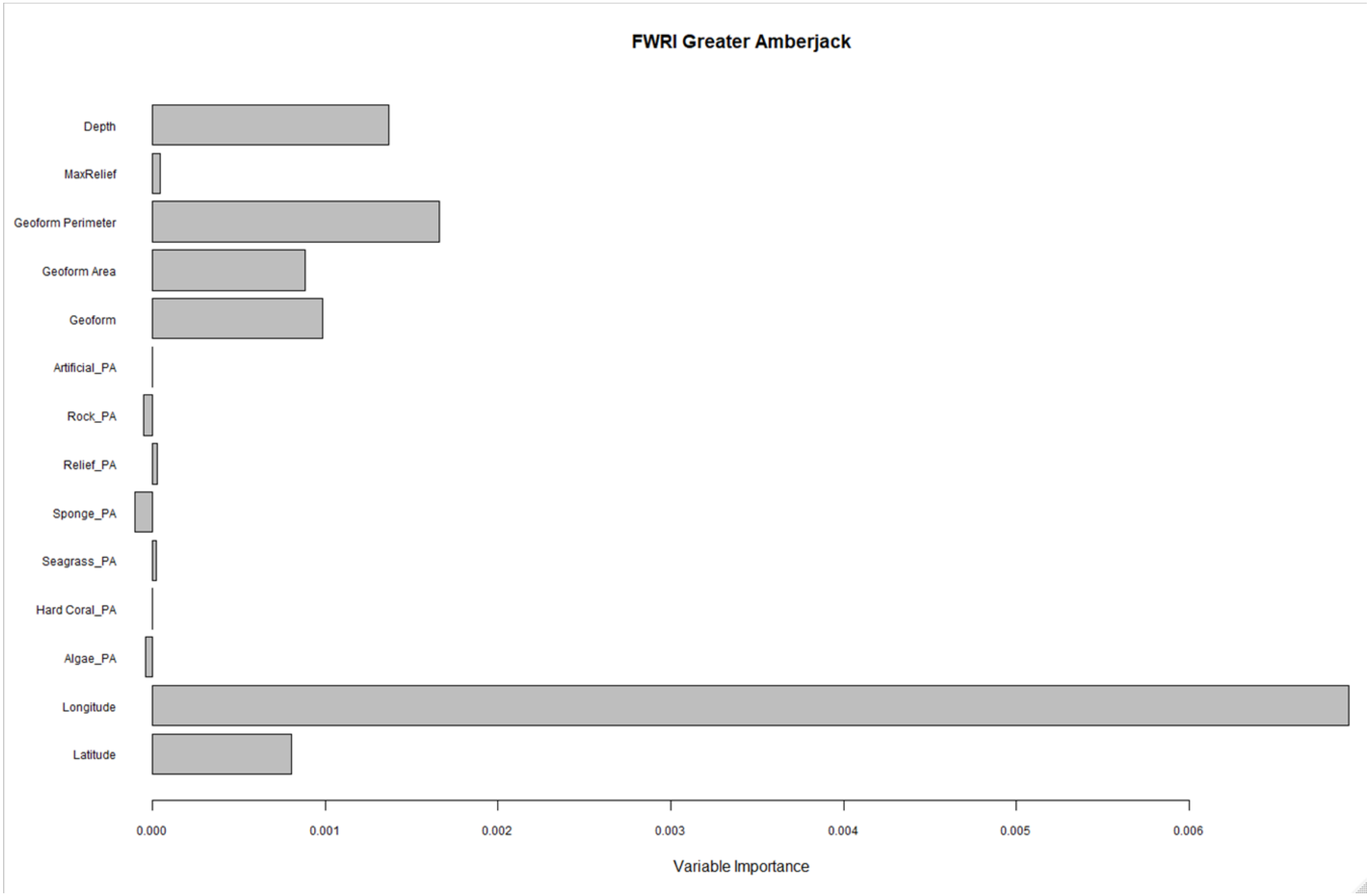

Figure 3. Random Forest generated variable importance for Greater Amberjack presence using FWRI survey data.

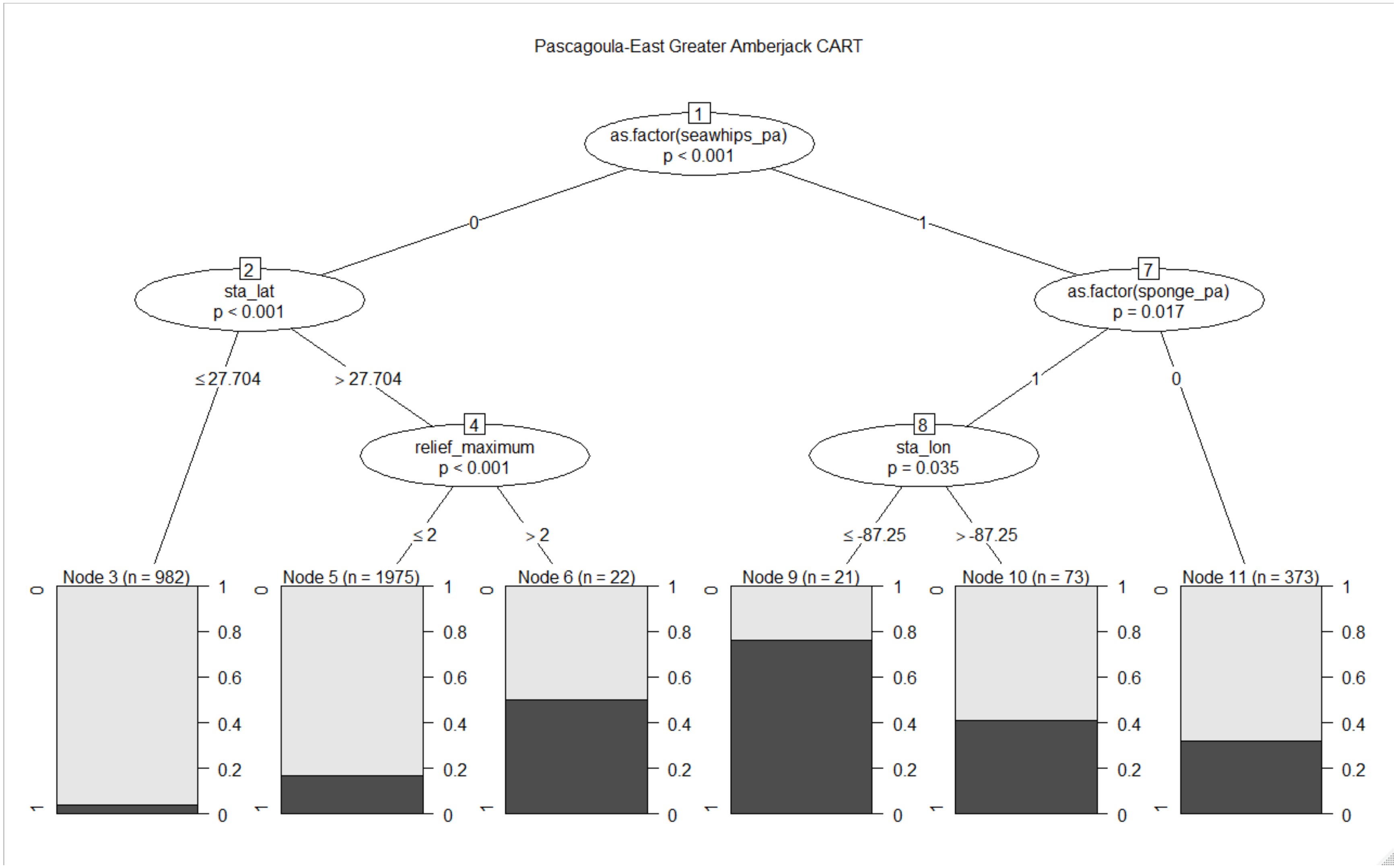

Figure 4. CART results for Greater Amberjack for Pascagoula's video survey for the eastern Gulf region. Shaded portion of the plots indicate proportion of sites given by a node where Greater Amberjack were observed (16.3% of sites had Greater Amberjack present overall).

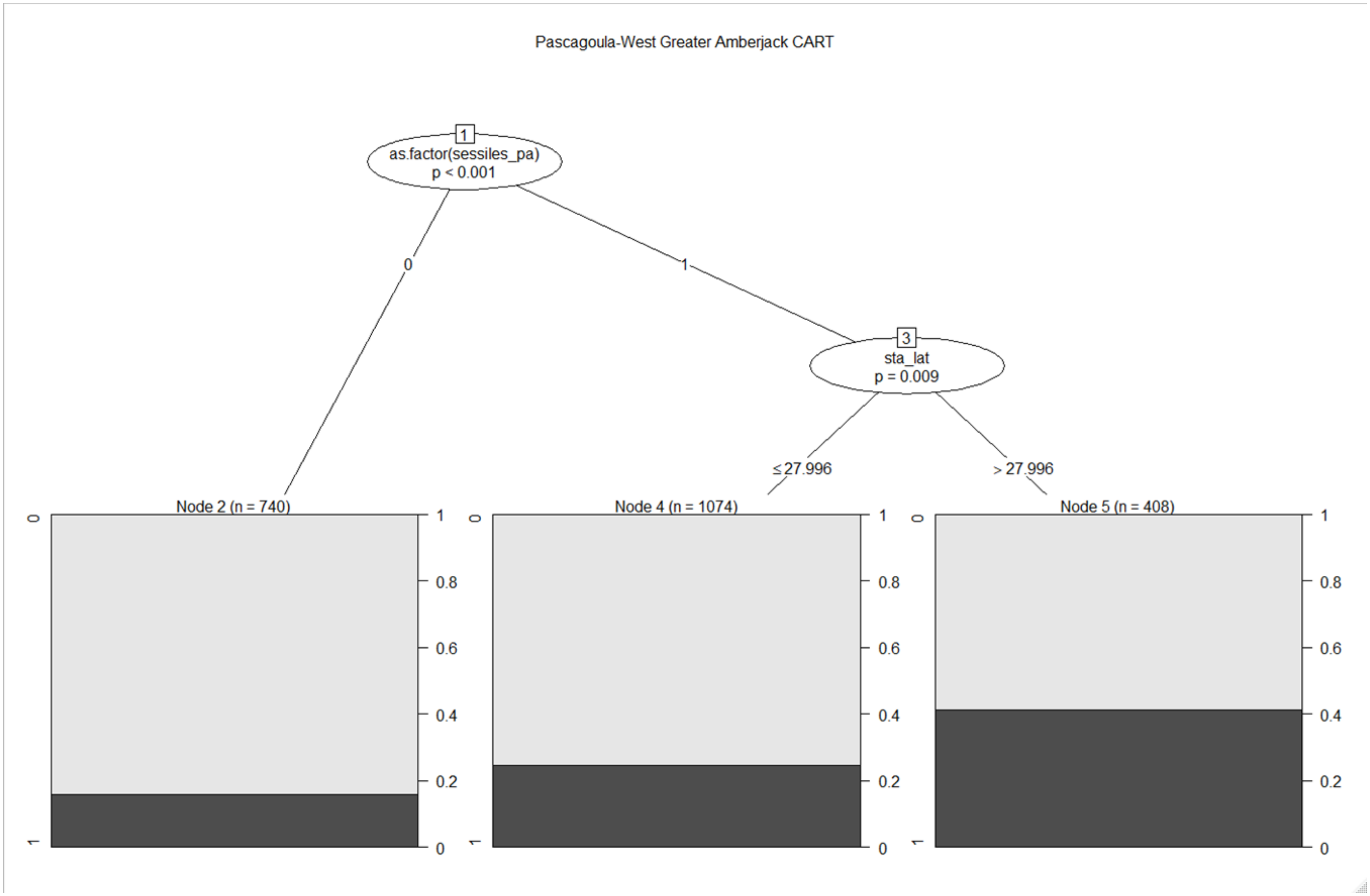

Figure 5. CART results for Greater Amberjack for Pascagoula's video survey for the western Gulf region. Shaded portion of the plots indicate proportion of sites given by a node where Greater Amberjack were observed (23.0% of sites had Greater Amberjack present overall).

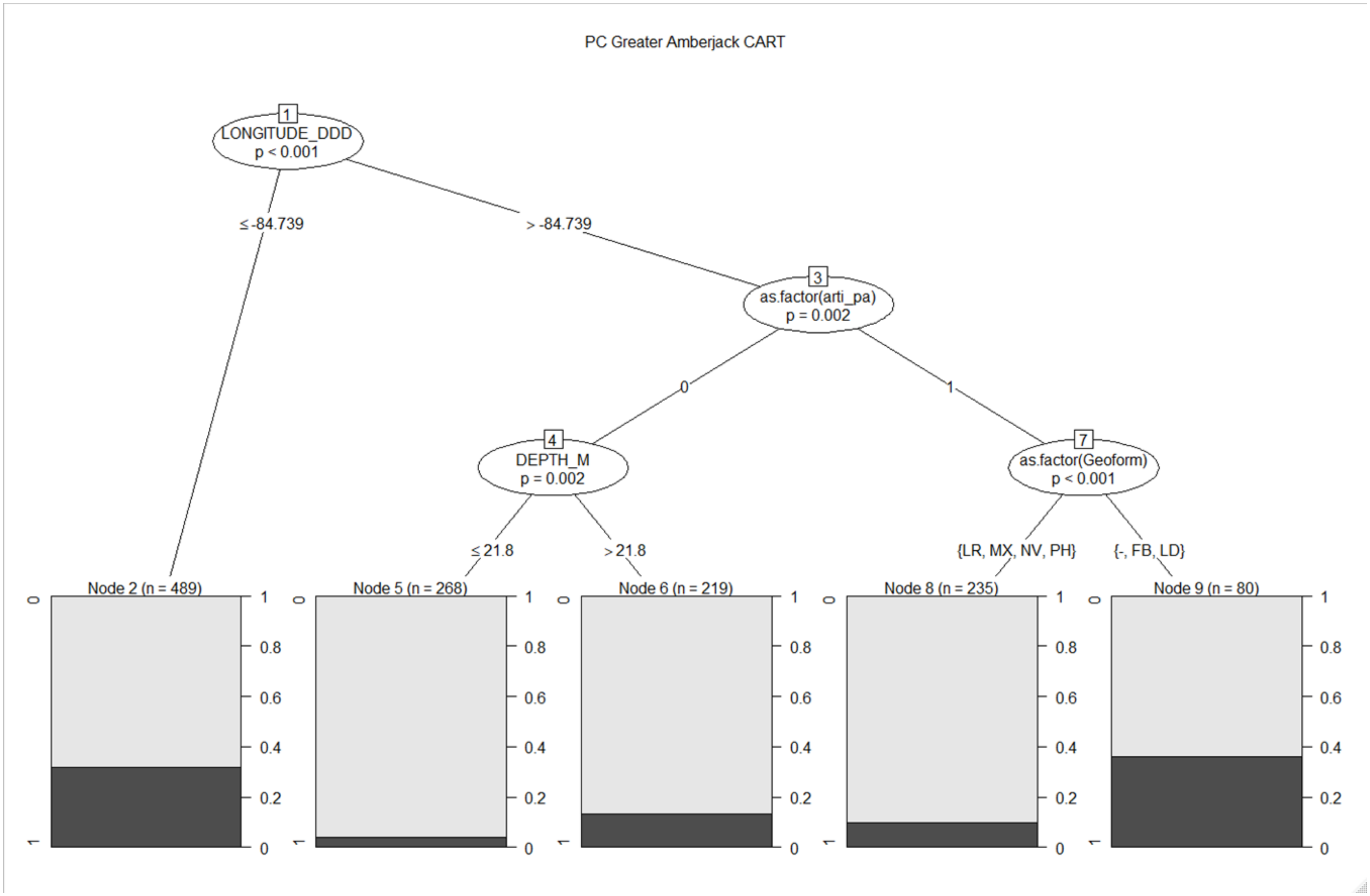

Figure 6. CART results for Greater Amberjack for Panama City's video survey. Shaded portion of the plots indicate proportion of sites given by a node where Greater Amberjack were observed (19.2% of sites had Greater Amberjack present overall)

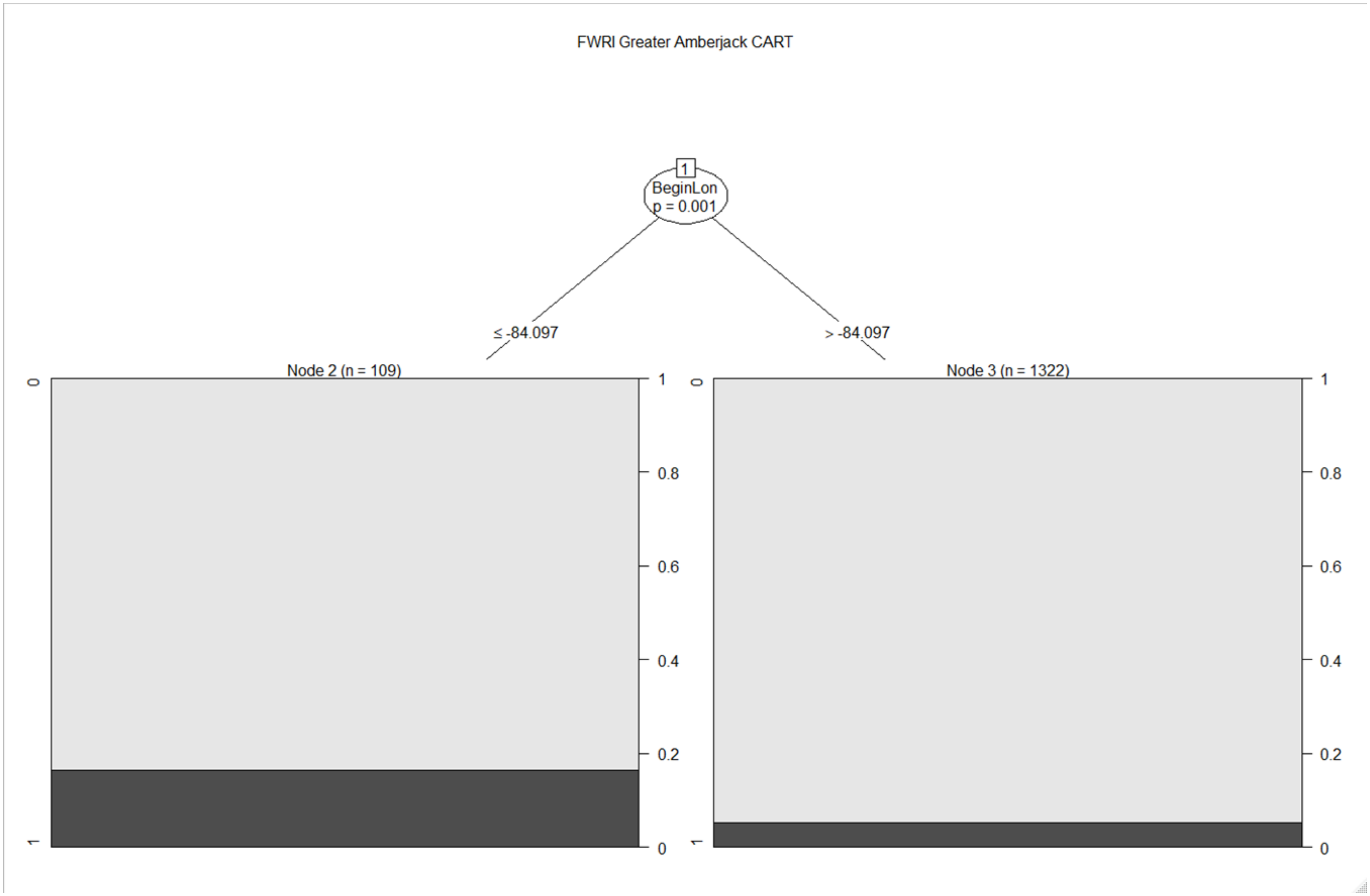

Figure 7. CART results for Greater Amberjack for FWRI's video survey. Shaded portion of the plots indicate proportion of sites given by a node where Greater Amberjack were observed (13.4% of sites had Greater Amberjack present overall).

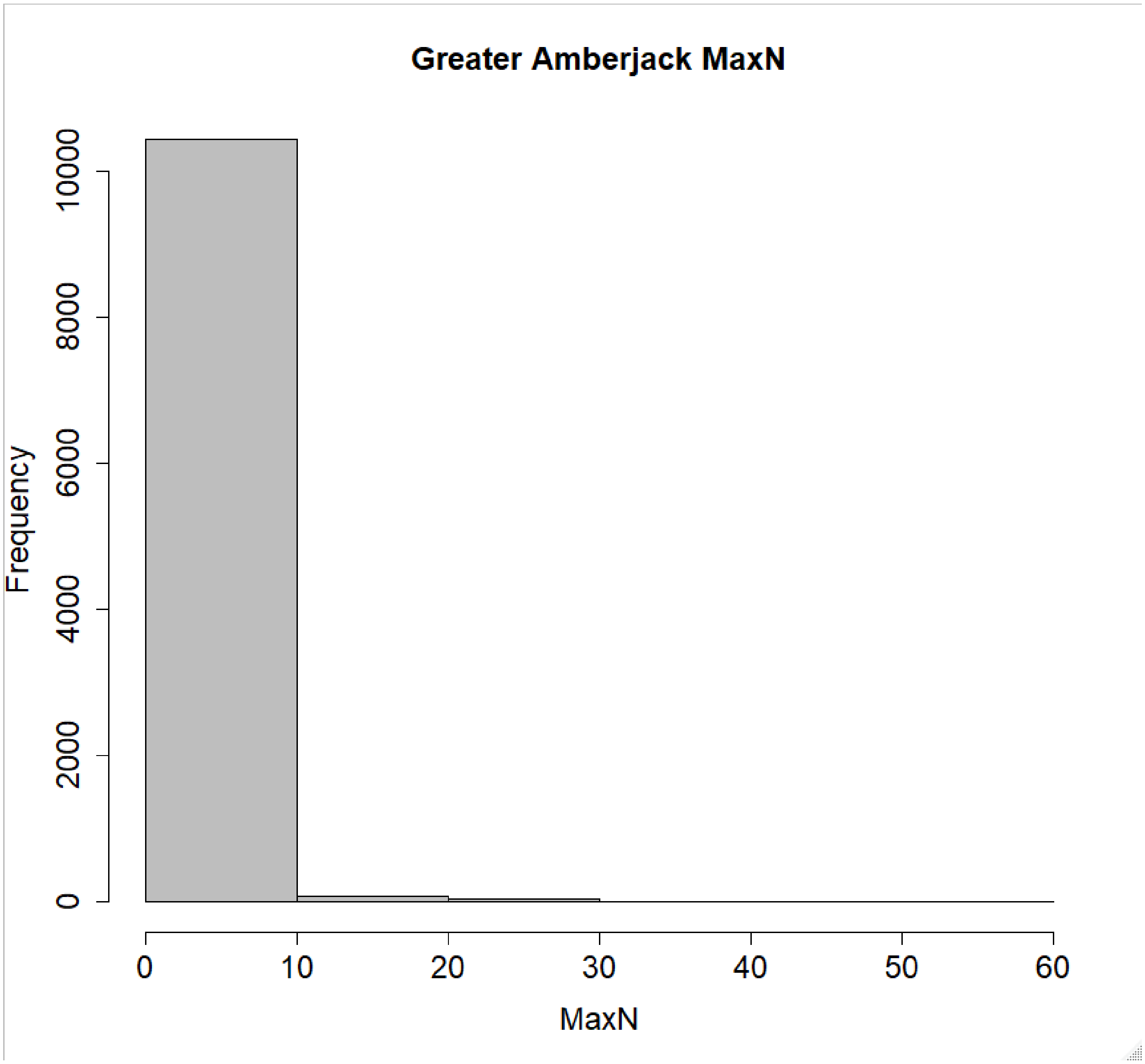

Figure 8. MaxN count distribution for Greater Amberjack observed in all three video surveys on the West Florida Shelf used for the combined index.

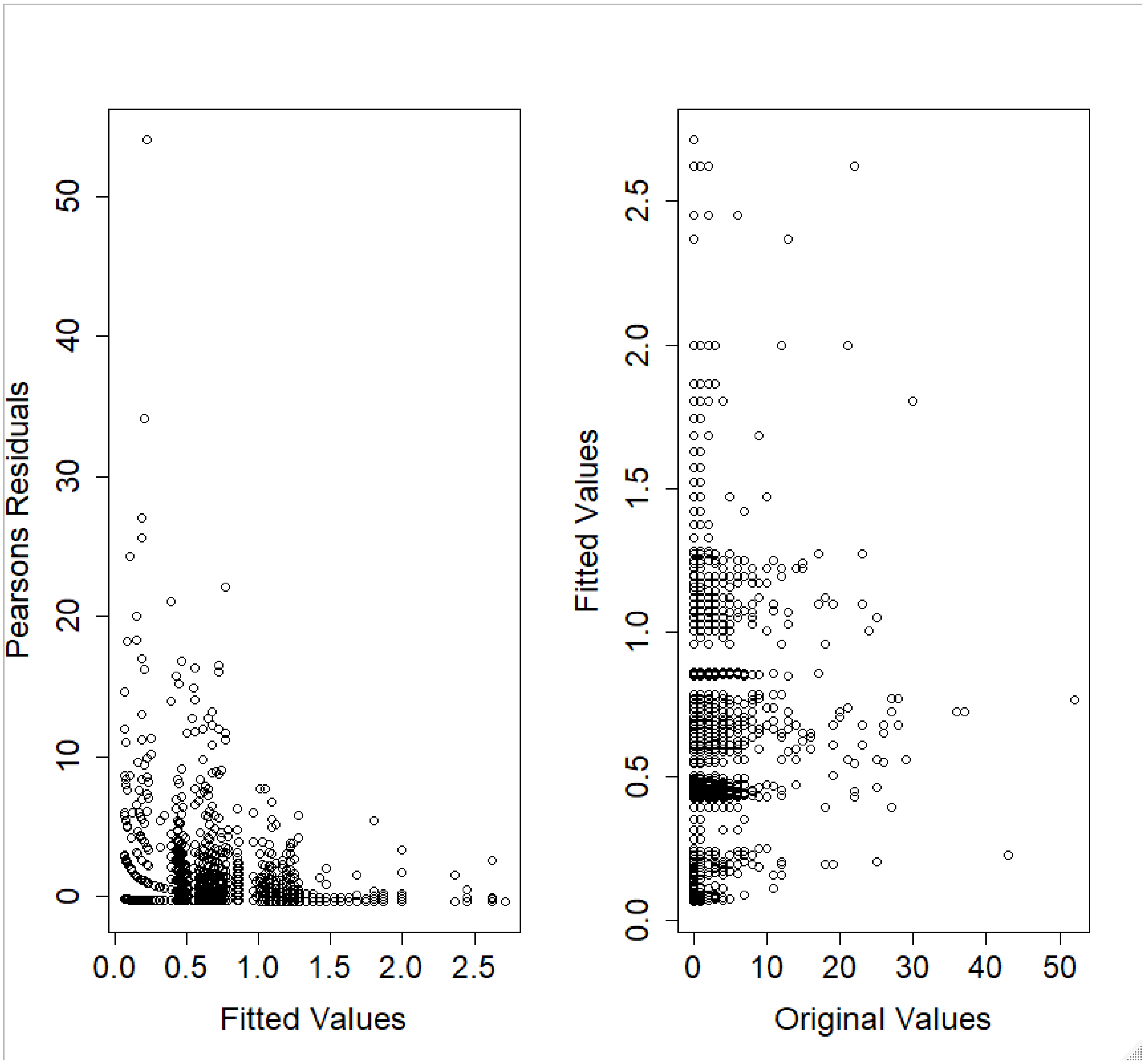

Figure 9. Model diagnostic plots showing fitted best model values against Pearson residuals (left panel) and fitted values plotted against original data values (right panel).

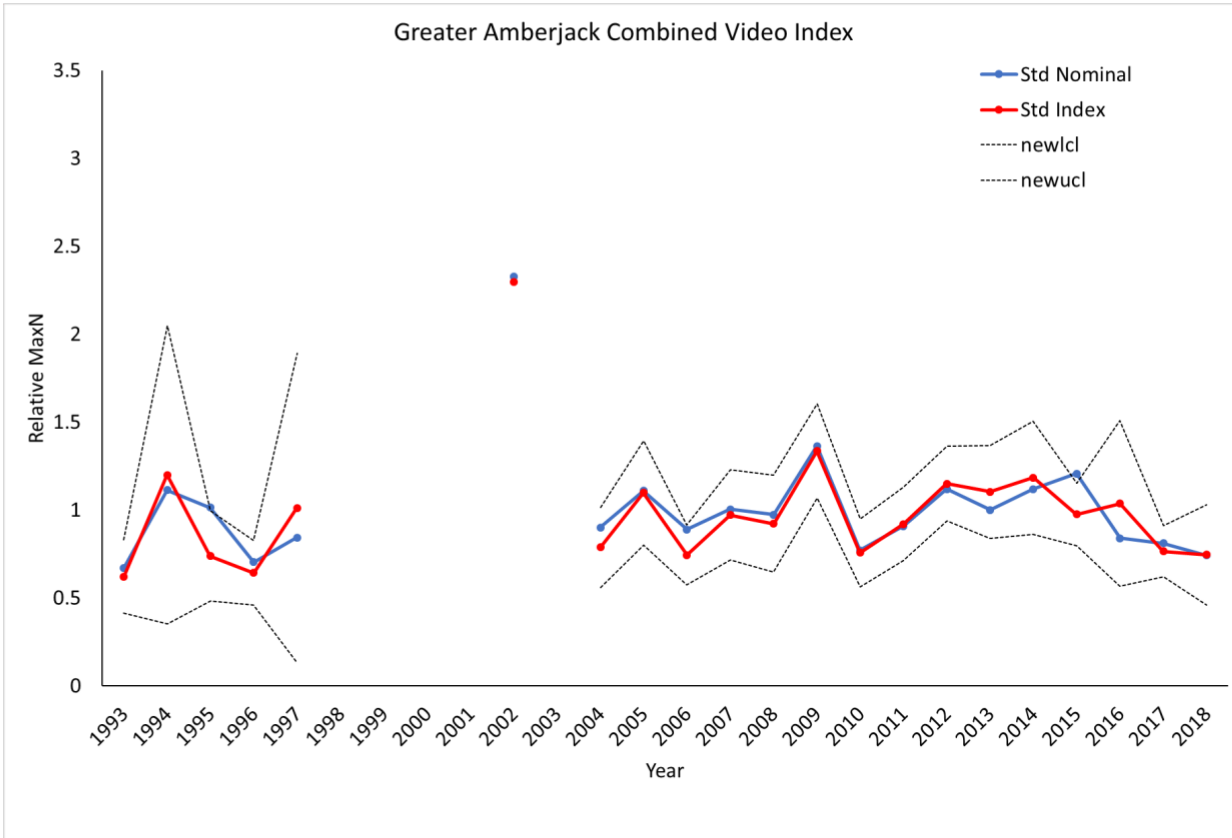

Figure 10. Standardized index (solid red line) with 2.5% and 97.5% confidence intervals (black dotted lines) and nominal index (solid blue line) for Greater Amberjack CPUE (MaxN) using the integrated West Florida Shelf video data.