Indices of abundance for Mutton Snapper (*Lutjanus analis*) using combined data from two fishery independent video surveys

Heather M. Christiansen, Kevin A. Thompson, Theodore S. Switzer, Sean F. Keenan, Christopher Gardner, Katherine E. Overly, Matt Campbell

SEDAR79-DW-21

23 August 2023

Please cite this document as:

Christiansen, Heather M., Kevin A. Thompson, Theodore S. Switzer, Sean F. Keenan, Christopher Gardner, Katherine E. Overly, Matt Campbell. 2023. Indices of abundance for Mutton Snapper (*Lutjanus analis*) using combined data from two fishery independent video surveys. SEDAR79-DW-21 SEDAR, North Charleston, SC. 17 pp.

Indices of abundance for Mutton Snapper (*Lutjanus analis*) using combined data from two fishery independent video surveys

Heather M. Christiansen[1], Kevin A. Thompson[2], Theodore S. Switzer[1], 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] Southeast Fisheries Science Center, Miami Laboratory, Miami, 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 of reef fishes were conducted in the northern Gulf of Mexico (GOM). Two of these surveys operated in the range of Mutton Snapper (*Lutjanus analis*); the NMFS SEAMAP reef fish video survey (SRFV), carried out by NMFS Mississippi Laboratory, has the longest running time series (1992-1997, 2002, and 2004+), followed by the Florida Fish and Wildlife Research Institute survey (FWRI, starting year 2008). While the surveys use standardized deployment, camera field of view, and fish abundance methods to quantify fish on reef or structured habitat, there were variations in survey design and habitat characteristics collected in addition to the time period and area sampled. Historically, independent indices were submitted for each respective survey. However, in most recent reef fish stock assessments, data from these video surveys have been combined to generate indices more representative of the total unit stock (Thompson et al. 2019a, 2019b, 2022a). Early efforts indicated that combining data from multiple surveys with varied spatial coverage through the use of a year only model can yield spurious conclusions regarding stock abundance (Campbell 2004; Ye et al. 2004). Accordingly, we used a habitat-based approach to combine relative abundance data for generating annual trends for Mutton Snapper throughout the eastern GOM (Thompson et al. 2022b).

Survey Comparisons

Survey design

The SRFV survey (1992 – 2019) primarily targeted high-relief topographic features along the continental shelf from south Texas to south Florida. Sites were 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 were selected at random from known reef areas identified through habitat mapping (multi-beam and side-scan sonar). This survey used the Mississippi river delta as a geographic feature separating the west and east regions of the GOM (Campbell et al. 2017); data from the western GOM were excluded due to the absence of Mutton Snapper in the region.

The FWRI survey (2008 – 2019) 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 subsequently expanded to include statistical zones 9 and 10 off the Florida Panhandle in 2014, and the remainder of the West Florida Shelf, encompassing statistical zones 2-10, in 2016. Sites were initially mapped using standardized side scan sonar surveys (Keenan et al. 2022), with video deployment sites randomly selected from identified natural reef features (Thompson et al. 2017).

Beginning in 2020, Gulf-wide video survey efforts were integrated under a single, novel survey design as the Gulf Fishery Independent Survey of Habitat and Ecosystem Resources (G-FISHER) program (Switzer et al. 2023). The G-FISHER survey utilizes a stratified-random sampling design, where sampling effort is randomly assigned to spatial (region and depth) and habitat strata (relief and spatial extent of reef features). Annual effort is allocated proportionally based on the product of habitat availability and managed species richness for each sampling stratum. Due to the Covid pandemic, sampling in 2020 was restricted to the eastern Gulf of Mexico only; full Gulf-wide effort was implemented beginning in 2021. Sites sampled under the new design (G-FISHER) are within the historical spatial footprint of the FWRI surveys, therefore all sites sampled in 2020 and 2021 were assigned to the FWRI survey.

Video annotation methodology

Both 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) to avoid duplicate counts. Measurements (fork length) are obtained at a single frame where the most individuals are measurable to avoid duplicate measurements. Habitat characteristics on video are also noted, including the percentage cover or presence/absence of abiotic and biotic habitat types (e.g. rock, sponge, algae, soft corals, hard corals), that may explain observed patterns of fish abundance, although some categories are not shared between the surveys (Campbell et al. 2017; Gardner et al. 2017; Thompson et al. 2017).

Data reduction

For both surveys, video reads were excluded if they were unreadable due to turbidity or deployment errors. For the SRFV, data included in this index are from 1993 and on, due to different counting methods in 1992. The FWRI data was limited to 2010 and on due to the lack of side-scan sonar habitat mapping efforts from which geoform, an important explanatory a variable, was determined. Mutton Snapper are rarely observed north of statistical zone 5; therefore, all data north of 28° N were excluded from subsequent analyses. Total number of samples used in subsequent analyses by lab and year can be found in Table 1, and spatial coverage is shown in Figure 1 and 2.

Index Construction

Habitat models

To combine the data from both surveys into one estimate of annual relative abundance of Mutton Snapper, we first generated a common habitat quality variable (Good, Fair, or Poor) to account for changing effort and habitat allocation through time. To do so, we used classification and regression trees (CART) because they can account for correlations among variables and can include both continuous and categorical data. CART models have 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; Thompson et al. 2022b).

For 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 both labs. For FWRI data only, 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. Geoform was not included as a predictor variable for the analysis of SRFV data because their historical habitat mapping has primarily been conducted utilizing multibeam sonar, and comparable habitat classification protocols have not yet been developed. We first used a random forest approach to reduce the number of potential explanatory variables. For the random forest, each lab was modeled separately with the entirety of that lab's dataset. The random forest runs fit 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 dataset.

We retained approximately 50% of the potential variables for each lab given by the random forest importance values for a final CART model. The final model was created by fitting the presence of Mutton Snapper at each site to the independent variables for a training dataset representing 80% of the data. The remaining 20% of the data were retained in a test dataset to determine misclassification rates for each of the two models. The proportion of sites with positive Mutton Snapper catches at each terminal node were then evaluated to determine the habitat characteristics which were defined as Good, Fair or Poor habitat. Terminal nodes with double the overall proportion of positive catches for a dataset were assigned a Good habitat code. Poor sites were determined by proportion positives that were at most half 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 4.0.3 (R Core Team 2020) and the Party package for CART (Hothorn et al. 2006).

CART results varied by lab with respect to the final variables chosen, but both had latitude in the final model. The SRFV model included longitude, latitude, presence/absence of sponge, maximum relief and depth as explanatory variables (Fig. 4). The FWRI model included Geoform and latitude as explanatory variables (Fig. 5). This species has divergent proportions present across the surveys, with SRFV (19%) having moderate occurrence rates and FWRI (5%) having relatively low occurrence rates.

The site characteristics that define each node and habitat code were then used to create a habitat variable (hab: G, F, P) that was then back-applied to each site for each lab's dataset. The datasets were then combined for the index model. The final proportion of sites in the three habitat categories for each lab are shown in Table 2.

Index model fitting and diagnostics

Like the individual survey indices, the combined dataset did not conform to assumptions of normality. We initially evaluated zero-inflated and standard negative binomial models but given the low dispersion parameter we determined the negative binomial model to be most appropriate. The final index model was then:

*MaxN = Y*Hab*Lab*

Where Hab is the CART derived habitat code and Lab 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.

The index was fit in SAS using the Proc GLIMMX procedure. A weighted glm was fit to account for annual variability in sampling effort among surveys and habitats, while adjusting for (1) the total reef area within each survey, and (2) the proportion of Good, Fair, and Poor habitats within each survey. The known potential survey universe (entire sampling frame) for each survey was first multiplied by the proportion of habitat mapping grids within which reef habitat has been identified to provide an estimate of total reef area. Next, the proportion of historical sampling sites that fell within each habitat type (Good, Fair, Poor) was calculated, and this proportion was multiplied by total reef area to provide estimates of reef habitat coverage for each survey and habitat type. Annual weighting factors were then calculated as the percentage of total reef habitat, sampled within a given year, that fell within a particular survey – habitat combination. Area weighting factors are provided in Table 3. Weighted index values were then standardized to the grand mean following standard SEDAR protocols.

Length compositions

Length compositions data for this species were limited in terms of sample sizes relative to other species assessed with these methods. As such, length data was combined between the two surveys for a final total length composition (Fig. 6). More specific annual length compositions or ones modeled with potential survey effects were not possible given the low sample sizes.

Results and Discussion:

Annual standardized index values for Mutton Snapper in the Eastern Gulf of Mexico, including coefficients of variation, are presented in Table 4. The model CVs are generally higher in the SRFV only survey portion of the time series but generally decrease as additional survey data are added. Biomass trends for Mutton Snapper increase early in the time series, followed by a decrease from 2004-2009, then increase until 2012 and subsequent drop off to somewhat stable, low overall catch in recent years (Table 4; Fig. 7).

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., Rademacher, K.R., Hendon, M., Felts, P., Noble, B., Caillouet, R., Salisbury, J., and Moser, J. 2017. SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Grey Snapper. SEDAR51-DW-07. SEDAR, North Charleston, SC. 31 pp.

Hothorn, T., Hornik, K., and Zeileis, A. 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. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. URL: http://www.R-project.org/.

Switzer, T.S., Keenan, S.F., Thompson, K.A., Shea, C.P., Knapp, A.R., Campbell, M.D., Noble, B., Gardner, C., Christman, M.C. 2023. Integrating assemblage structure and habitat mapping data into the design of a multispecies reef fish survey. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science 15: e10245.

Thompson, K.A., Switzer, T.S., and Keenan, S.F. 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, K.A., Switzer, T.S. Christman, M.C., Keenan, S.F., Gardner, C., Overly, K.E., Campbell, M. 2019a. 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.

Thompson, K.A., Switzer, T.S, Christman, M.C, Keenan, S.F., Gardner, C., Overly, K.E., Campbell, M. 2019b. Indices of abundance for Gray Triggerfish (*Balistes capriscus*) using combined data from three independent video surveys. SEDAR62- WP-01. SEDAR, North Charleston, SC. 17 pp.

Thompson, K.A., Switzer, T.S., Christman, M.C., Keenan, S.F., Gardner, C., Overly, K.E., Campbell, M. 2022a. Indices of abundance for Red Snapper (*Lutjanus campechanus*) on natural reefs in the eastern Gulf of Mexico using combined data from three independent video surveys. SEDAR74-DW-23. SEDAR, North Charleston, SC. 25 pp.

Thompson, K.A., Switzer T.S., Christman, M.C., Keenan, S.F., Gardner C.L., Overly K.E., and Campbell M.D. 2022b. A novel habitat-based approach for combining indices of abundance from multiple fishery-independent video surveys. Fisheries Research 247: 106178.

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

Yates, K.L., Mellin, C., Caley, M.J., Radford, B.T., Meeuwig, J.J. 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., Ieno, E.N., Walkder, N.J., Saveliev, A.A., and Smith, G.M. 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 two included video surveys, Florida Fish and Wildlife Research Institute (FWRI) and NMFS SEAMAP (SFRV). No data were available or used from any survey from 1998-2001 and 2003.

Year	FWRI	SFRV	Total
1993		32	32
1994		31	31
1995		24	24
1996		42	42
1997		54	54
2002		48	48
2004		26	26
2005		78	78
2006		85	85
2007		110	110
2008		79	79
2009		80	80
2010	45	79	124
2011	205	102	307
2012	215	105	320
2013	184	44	228
2014	278	78	356
2015	220	35	255
2016	347	93	440
2017	286	125	411
2018	255	93	348
2019	346	116	462
2020	464		464
2021	547		547
total	3392	1559	4951

Table 2. Proportion of sites for each habitat level (**F**air**, G**ood**, P**oor) as determined by individual lab categorical regression trees (CARTs) for Mutton Snapper presence.

	SRFV	FWRI
Good	0.23	0.15
Fair	0.30	0.07
Poor	0.47	0.78

Table 3. Estimated amount of total reef habitat within each survey domain, and resultant survey-specific habitat weighting factors. These weights were multiplied by the total percentage of Good, Fair, and Poor habitats within each survey (Table 2) to define final habitat weights.

	Survey			
	SRFV	FWRI (2010-2015)	FWRI (2016-2019)	FWRI (2020-2021)
Total Universe Area (km^2)	13292.3	37290.0	71891.4	71891.4
Area x Proportion of mapped with reef	7417.1	7569.9	14809.6	14809.6
Time Period Weighting Values				
1993-2009	1			
2010-2015	0.49	0.51		
2016-2019	0.33		0.67	
2020-2021				1

Table 4. Number of stations sampled (N) by survey and year, proportion of positive sets, number of fish captured (N fish), standardized index, and CV for the annual FWRI Mutton Snapper video index of the eastern GOM.

Year	N	Prop pos	N fish	Std. Index	LCL	UCL	CV
1993	32	0.063	2	0.296	0.230	0.363	0.732
1994	31	0	0	0.000	0.000	0.000	0.302
1995	24	0.042	1	0.120	0.083	0.158	1.014
1996	42	0.214	10	2.959	2.370	3.548	0.652
1997	54	0.167	12	1.295	1.161	1.429	0.340
1998							
1999							
2000							
2001							
2002	48	0.250	19	1.802	1.643	1.962	0.290
2003							
2004	26	0.423	11	1.316	1.176	1.457	0.349
2005	78	0.167	21	1.389	1.286	1.492	0.243
2006	85	0.259	29	2.286	2.140	2.432	0.209
2007	110	0.236	27	1.482	1.387	1.579	0.212
2008	79	0.152	13	1.216	1.098	1.333	0.318
2009	80	0.138	13	0.8759	0.797	0.955	0.296
2010	124	0.153	20	0.592	0.548	0.636	0.245
2011	307	0.0814	47	1.306	1.238	1.374	0.171
2012	320	0.088	47	1.324	1.253	1.396	0.176
2013	228	0.022	5	0.642	0.528	0.756	0.582
2014	356	0.028	11	0.652	0.575	0.726	0.382
2015	255	0.000	0	0.000	0.000	0.000	0.312
2016	440	0.100	54	0.988	0.934	1.042	0.179
2017	411	0.054	30	0.500	0.467	0.533	0.215
2018	348	0.147	61	0.936	0.889	0.983	0.165
2019	462	0.123	82	1.274	1.218	1.330	0.144
2020	464	0.099	49	0.372	0.354	0.390	0.157
2021	547	0.077	48	0.376	0.357	0.395	0.169

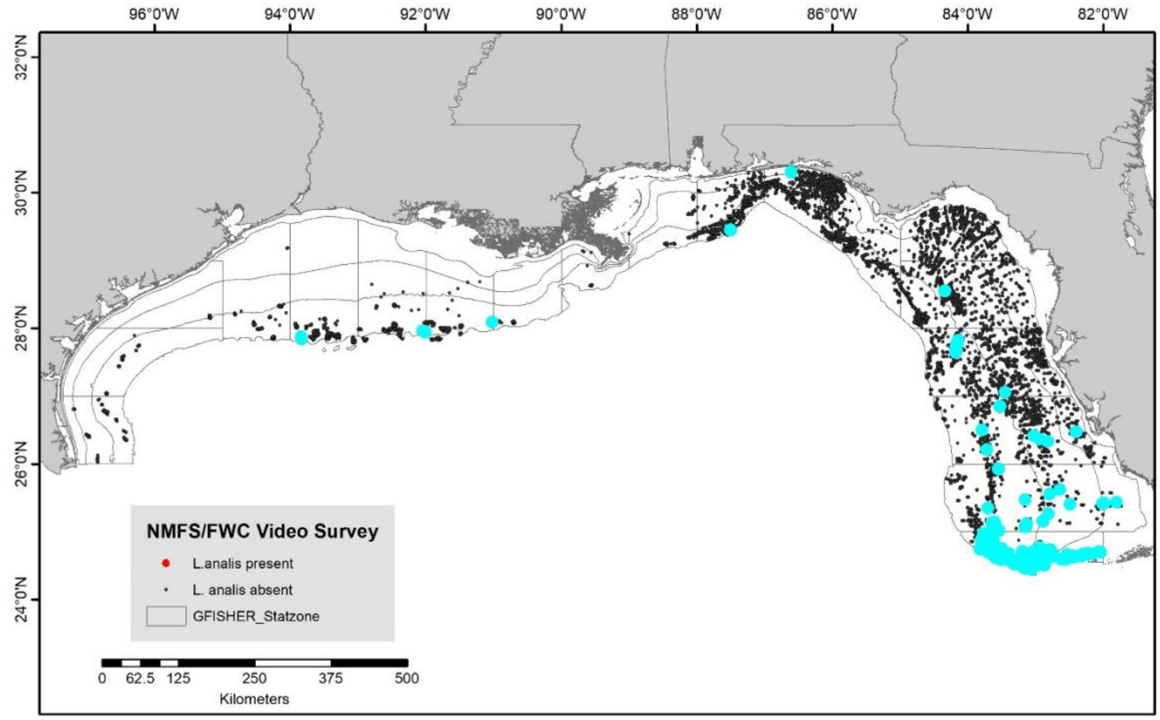

Figure 1. Map of the total video sites for each survey across all years 1993-2021.

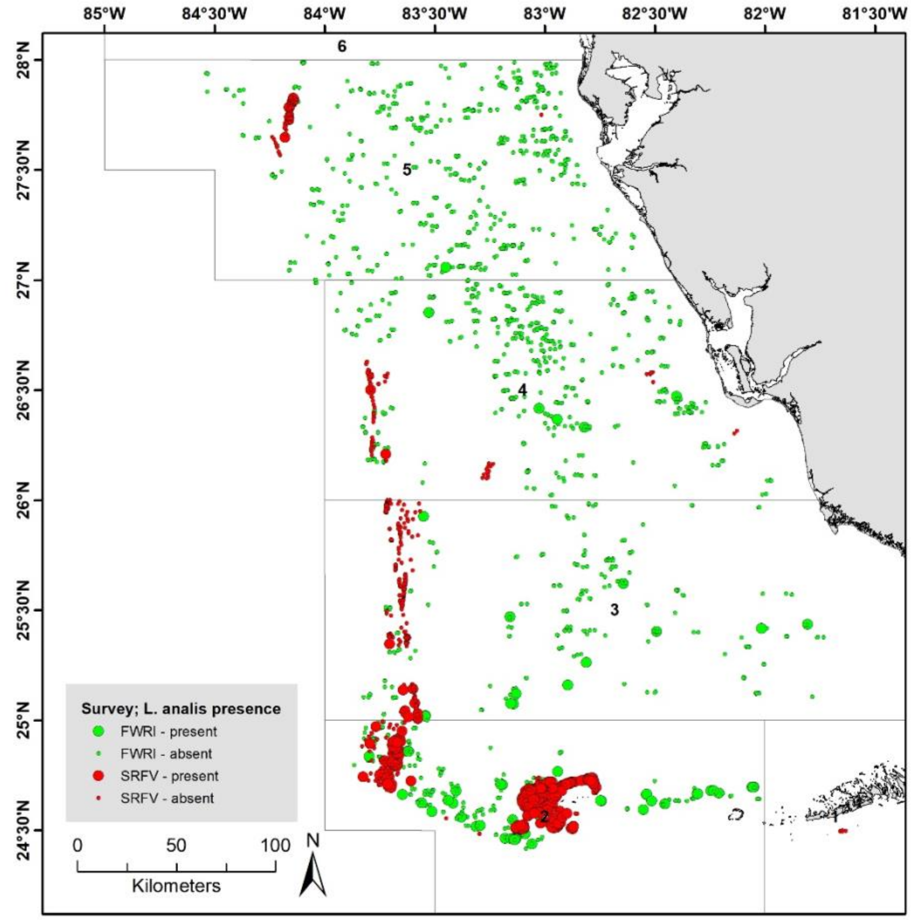

Figure 2. Map of the total video sites included in the index for each survey across all years 1993-2021.

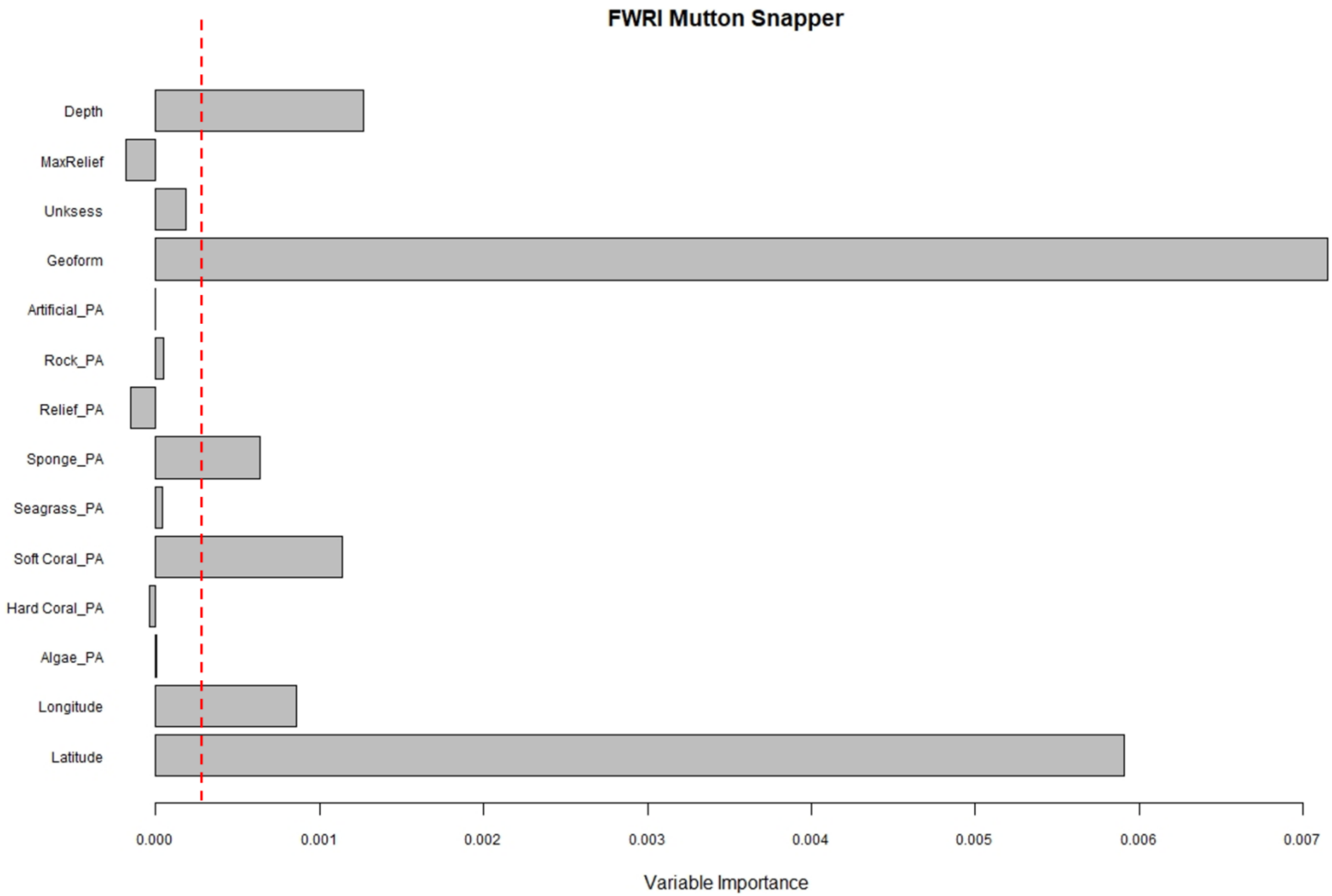

Figure 3. Random Forest generated variable importance for Mutton Snapper presence using FWRI survey data. Red dashed line indicates variables that were included in the analysis.

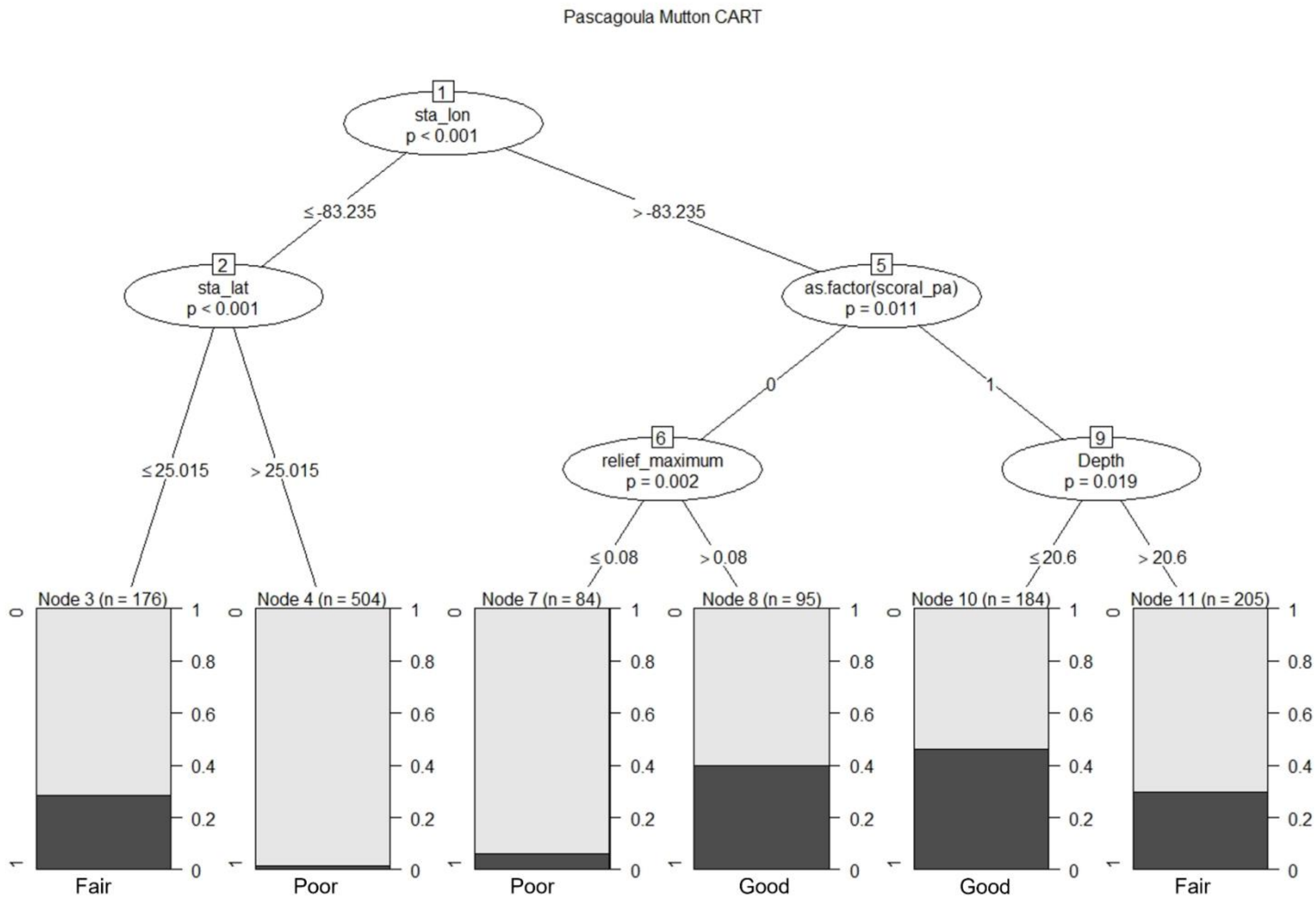

Figure 4. CART results for Mutton Snapper for SFRV survey. Shaded portion of the plots indicate proportion of sites given by a node where Mutton Snapper were observed (20% of sites had Mutton Snapper present overall; 19.6% misclassification rate).

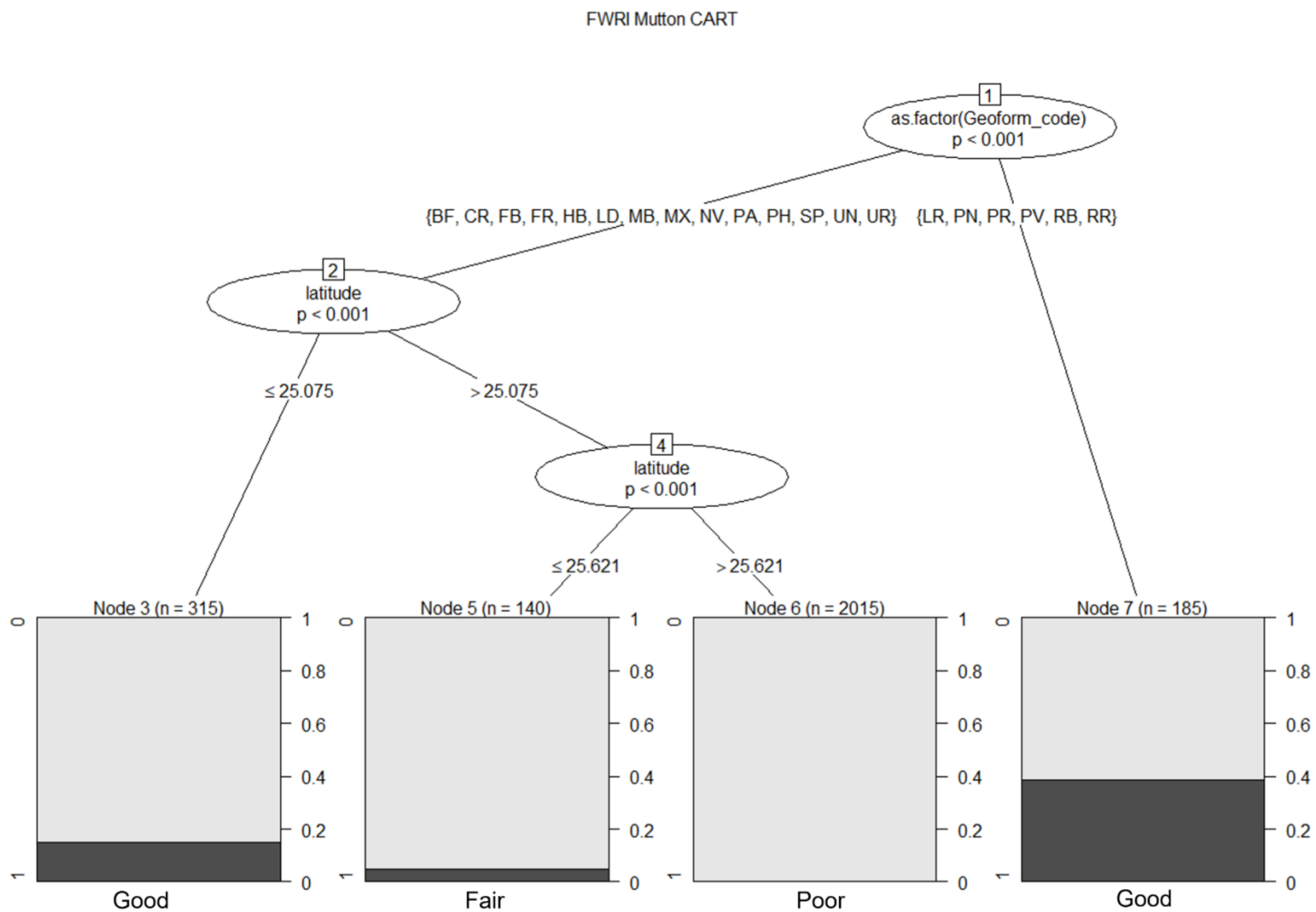

Figure 5. CART results for Mutton Snapper for FWRI's video survey. Shaded portion of the plots indicate proportion of sites given by a node where Mutton Snapper were observed (5% of sites had Mutton Snapper present overall; 5% misclassification rate).

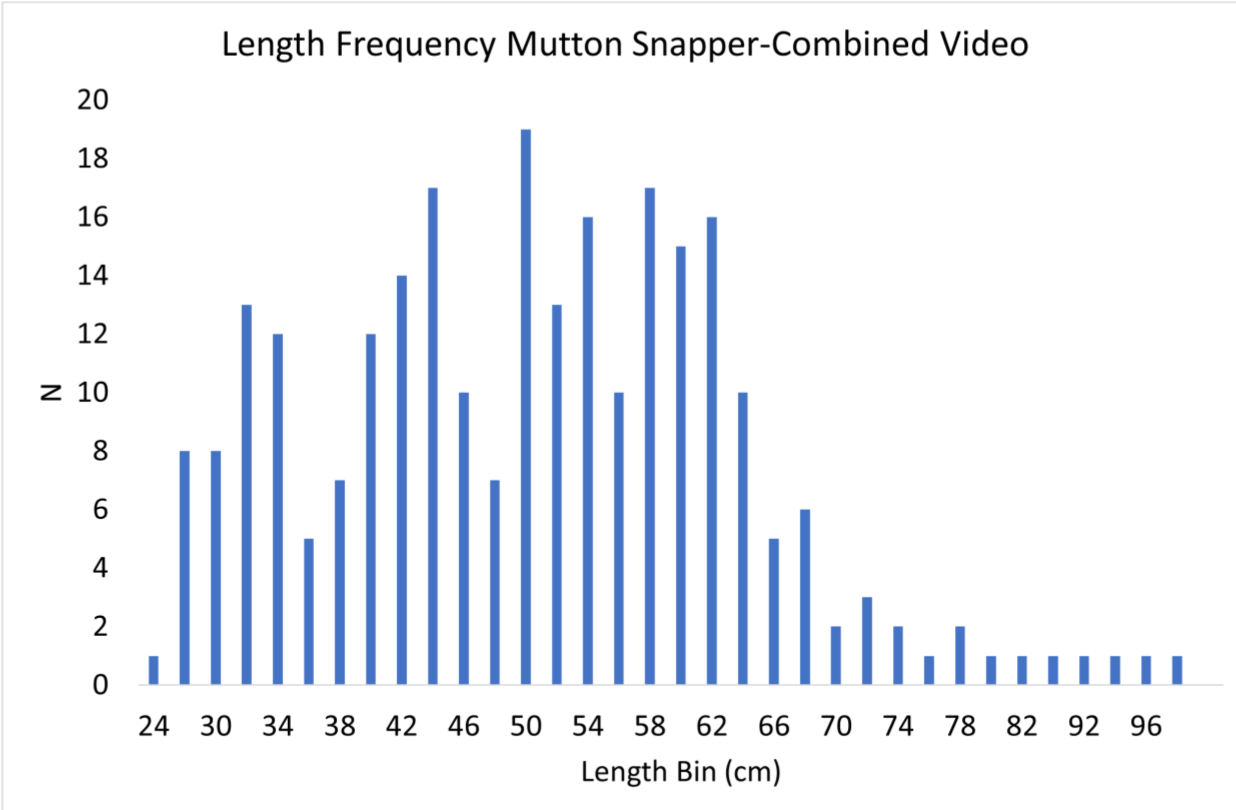

Figure 6. Length frequency used in the combined video index (n=257, 1993-2021).

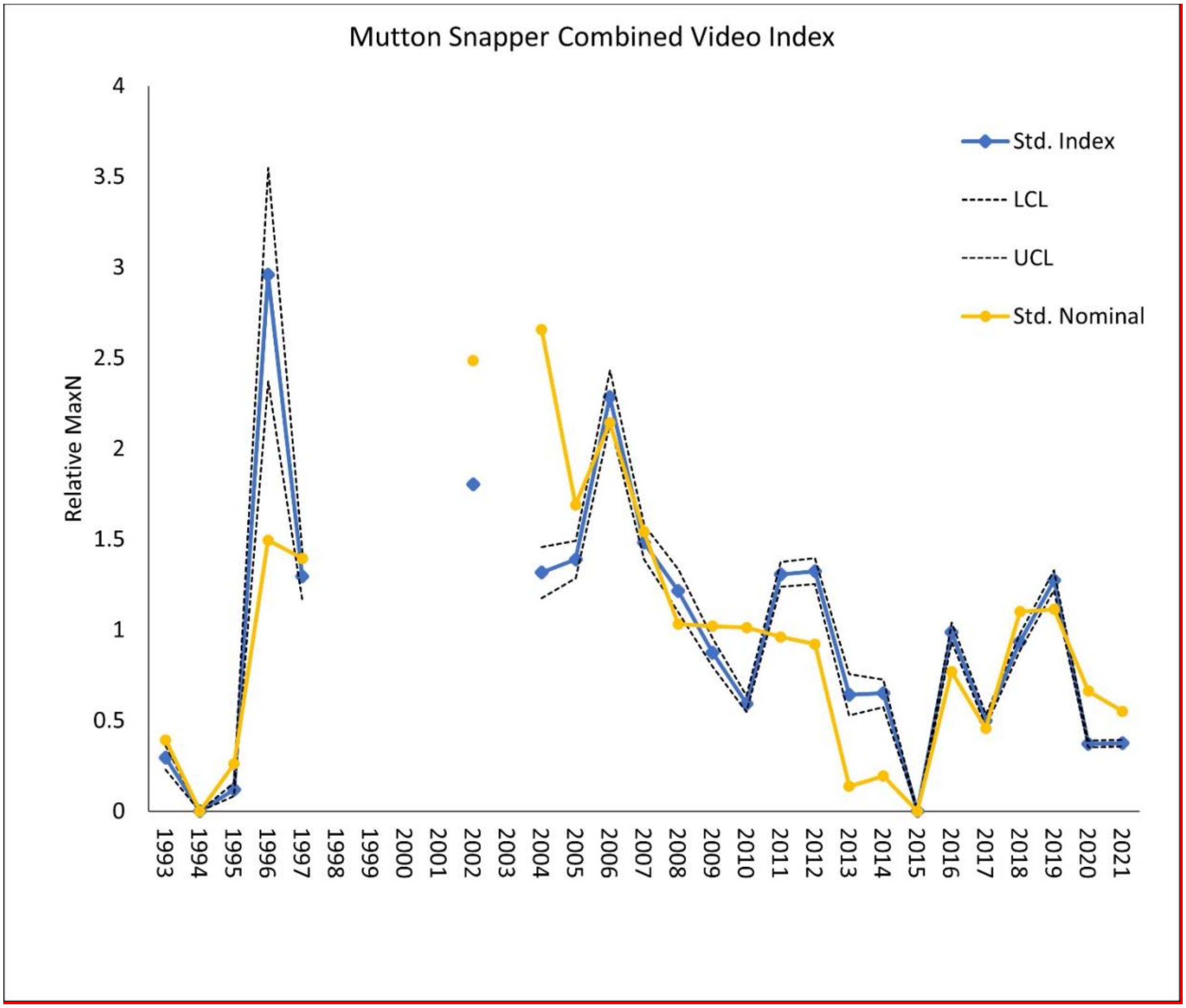

Figure 7. Relative standardized index with 2.5% and 97.5% confidence intervals (black dotted lines) and relative nominal index for Mutton Snapper CPUE (MaxN) using the integrated eastern GOM data.