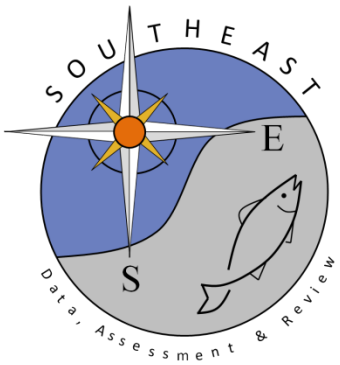
Investigation of Cobia Length Frequency Distributions and Potential for Differences Amongst Data Sets

Justin Yost, Joseph Ballenger, and Michael R. Denson

SEDAR58-DW05

Submitted: 11 January 2019
Revised: 4/23/19



Please cite this document as:

Yost, J., J. Ballenger, and M.R. Denson. 2019. Investigation of Cobia Length Frequency Distributions and Potential for Differences Amongst Data Sets. SEDAR58-DW05. SEDAR, North Charleston, SC. 18 pp.

Investigation of Cobia Length Frequency Distributions and Potential for Differences Amongst Data Sets

SEDAR 58 – Atlantic Cobia

Life History Working Group

Working Paper SEDAR58-DW05

Justin Yost, Joseph Ballenger, Michael R. Denson

South Carolina Department of Natural Resources

Charleston, SC

Introduction

Cobia are a coastal migratory species that inhabit warmer tropical and subtropical oceans around the world. Cobia typically arriving in South Carolina in late April or early May when water temperatures reach 20°C and remain until early fall when water temperatures begin to drop below that temperature. Since cobia travel in small groups, traditional fishery-independent sampling is less effective in capturing them requiring researchers to obtain the majority of life history samples from the recreational and charter boat fisheries in the form of carcass donations. A major problem with fishery-independent sampling is that some scientists and managers consider the data provided by fishermen as not to have been collected in a sufficiently rigorous manner and potentially to be biased (Hoare et al., 2011). This working paper compares the length frequency distribution of five fishery-dependent datasets, three data sets provided by South Carolina Department of Natural Resources (SCDNR) fishery-dependent sampling program beginning in 2007 (tournament, charter boat captain donations, and private recreational donations), two traditional NMFS fishery-dependent sampling efforts operating in the region (the Marine Recreational Information Program (MRIP) and the Southeast Region Headboat Survey (SRHS)) and one fishery-independent dataset collected by SCDNR staff from a Cooperative Research Program (CRP) funded grant to determine if SCDNR's carcass collection program is an accurate representation of the recreational fishery. Most fishery-dependent datasets (MRIP, SRHS) for cobia suffer from the same potential biases due to angler selectivity in the fishery which is not reflected in the sample even if a random design is followed since fish were collected and donated by recreational and for-hire fisherman and not scientific staff. All fork length measurements were made by trained biologists with the exception of genetic samples in the private data source which were measured by recreational fisherman.

Comparisons were made using a Kolmogorov-Smirnov test and a Kruskal-Wallis test to evaluate length frequency distributions between datasets. South Carolina fishery-dependent data sets were examined in more depth using the same statistical tests to evaluate year of sample, source of sample, sex of sample, and location of sample to determine if these variables influenced the size of animal reported in catch.

Data sources

I. South Carolina DNR fishery-dependent sampling programs (generally have associated location, sex, year, and source information) with samples collected from 2007-2016. No samples were collected in 2017 due to an inshore and federal closure of cobia.

a. Charter Boat: This current sampling program began in 2007 utilizing a shared dock in the Hilton Head area which includes approximately ten charter boat captains. Multiple coolers were provided for carcass donations and checked daily during the peak collection season. Captains were instructed to use a colored zip tie to indicate if a fish was captured inshore or offshore. SCDNR staff took measurements and removed otoliths when possible. Since this program relies on fishermen for samples, fish under the minimum size limit were not sampled. This dataset may also be biased to capturing fish in a particular location (inshore vs offshore) depending on regulations and availability of fish in inshore waters. These samples are also limited to fish captured in southern South Carolina waters which constitutes the bulk of the cobia fishery in South Carolina.
b. Private: These current sampling programs began in 2007 and 2008 and include freezer drop-off locations throughout the southern coastal region of South Carolina ranging from Charleston to Port Royal Sound and a genetic fin clip program covering the same geographical area. Participation in the freezer program was limited, thus all freezer locations were removed with the exception of one freezer on Hilton Head Island in 2009 and drop-off locations at SCDNR in Charleston and Bluffton. Recreational anglers where provided trash bags and data cards and instructed to place card and carcass into bag and into freezer. SCDNR staff checked freezers daily and measured and removed otoliths from any provided carcasses. Similar to the charter boat donations, this program relies on fishermen for samples restricting collections to fish over the legal size limit. This dataset may also be biased to capturing fish in a particular location (inshore vs offshore) depending on availability of fish in inshore waters and is also limited to fish captured in southern South Carolina waters. We are also aware of the potential for this program to miss cobia captured by anglers returning to private docks. The fin clip program provides recreational anglers with genetic vials to remove a portion of the anal fin for genetic evaluation and requires anglers to measure fish and provide a capture location and date. All undersized fish were removed from the analyses for comparison between other sampling methods and accounts for 175 of the private mode samples.
c. Tournament – Tournament sampling began in 2007 and continues to date. Due to a fishing closure in state waters which began in 2016 and public pressure, annual tournament numbers have decreased dramatically from six in 2008 down to one tournament in 2018. All fish weighed at the tournament are measured, capture location recorded, otoliths removed, and gonads preserved by trained fishery biologists from SCDNR. These fish were expected to be larger, given that they represent the fish weighed in during fishing tournaments (could be reason to exclude from analyses within state). These samples were also limited to the southern coastal region of the state.

II. South Carolina fishery-independent projects
a. CRP – This grant funded project began in 2016 to examine migration patterns of cobia using acoustic telemetry. To date fifty-one cobia have been tagged and measured by SCDNR biologists in South Carolina. This sampling project tended to catch the smallest Cobia because collection was not constrained by minimum length limits and undersized individuals were used for acoustic tagging. Fish were tagged throughout most of the region with the exception of the northern coast where cobia catches are minimal.

III. NMFS fishery-dependent monitoring programs – constrained to only include records from SC for the MRIP survey but includes SC/GA collections for the SRHS survey because more detailed capture locations were not available at the time
   a. MRIP – Marine Recreational Information Program is a federal program tasked with implementing surveys that measure how many trips recreational saltwater fisherman take and how many fish they catch and can be downloaded at https://www.st.nmfs.noaa.gov/recreational-fisheries/data-and-documentation/queries/index. Data collection begins in 1982 and includes fish captured in 2018 and were not separated into mode due to the lack of sample size.
   b. SRHS – Southeast Region Headboat Survey is a federal program focusing on monitoring and sampling the recreational Headboat fisheries in the Atlantic and Gulf of Mexico. Data was provided by Kelly Fitzpatrick from NOAA and encompasses data collected from 1978 until 2016. Exact landing information is limited so samples were restricted to South Carolina and Georgia. Similar biases to other fishery-dependent data sources can be expected since samples are captured by recreational anglers and size limits are enforced.

Investigation by data source

Sample Summary Table

Table 1: Summary statistics by data source. Provided is the sample size (n), mean (avg), standard deviation (SD), standard error (SE), median, and inter-quartile range of fork length (in mm) by data source. Table is sorted from the largest to smallest mean fork length of individual fish in the data source.

Source	n	Avg	SD	SE	Median	Inter-Quartile Range
Tournament	600	1055	124	5	1046	190
MRIP	157	1016	128	10	990	185
SRHS	110	1013	131	13	980	190
Charter boat	1292	1010	105	3	998	149
Private	570	994	104	4	979	156
CRP	51	886	137	19	880	182

Fork length box plot by data source

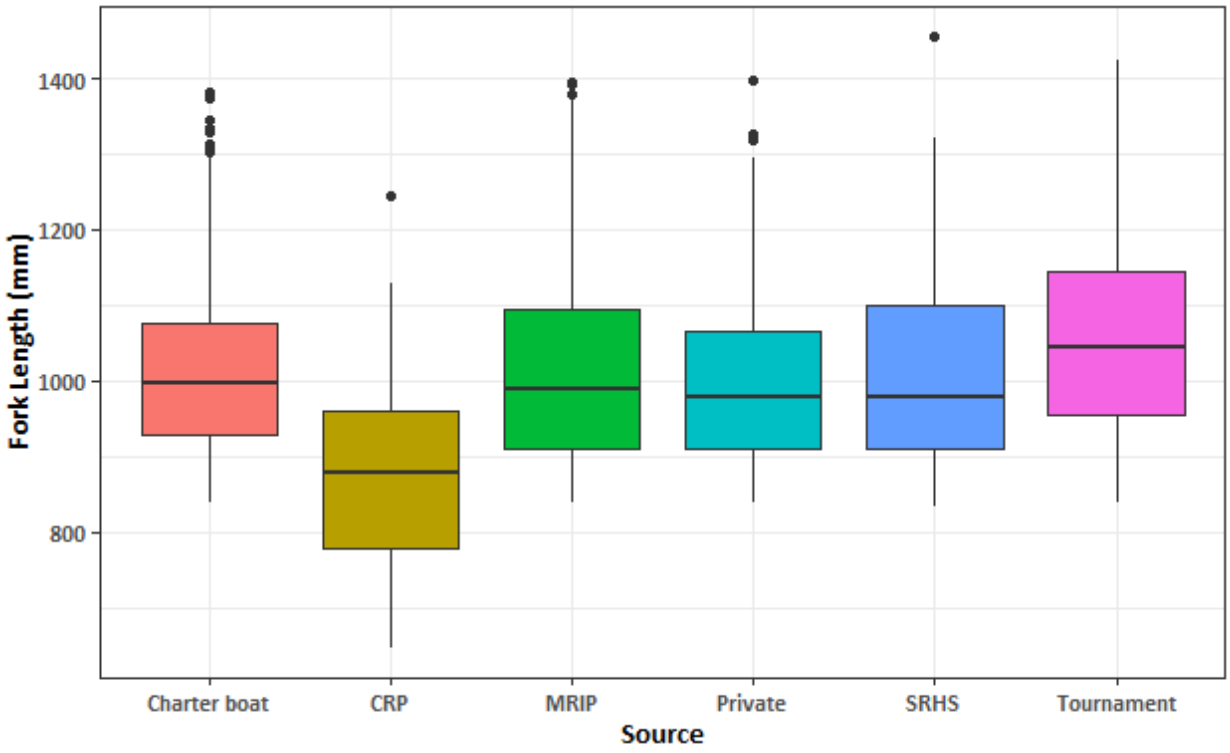


Figure 1: Box plot of fishery-dependent and independent data sets examining the fork length (mm) distribution of cobia captured using fishery-dependent sampling techniques.

Density plot by data source

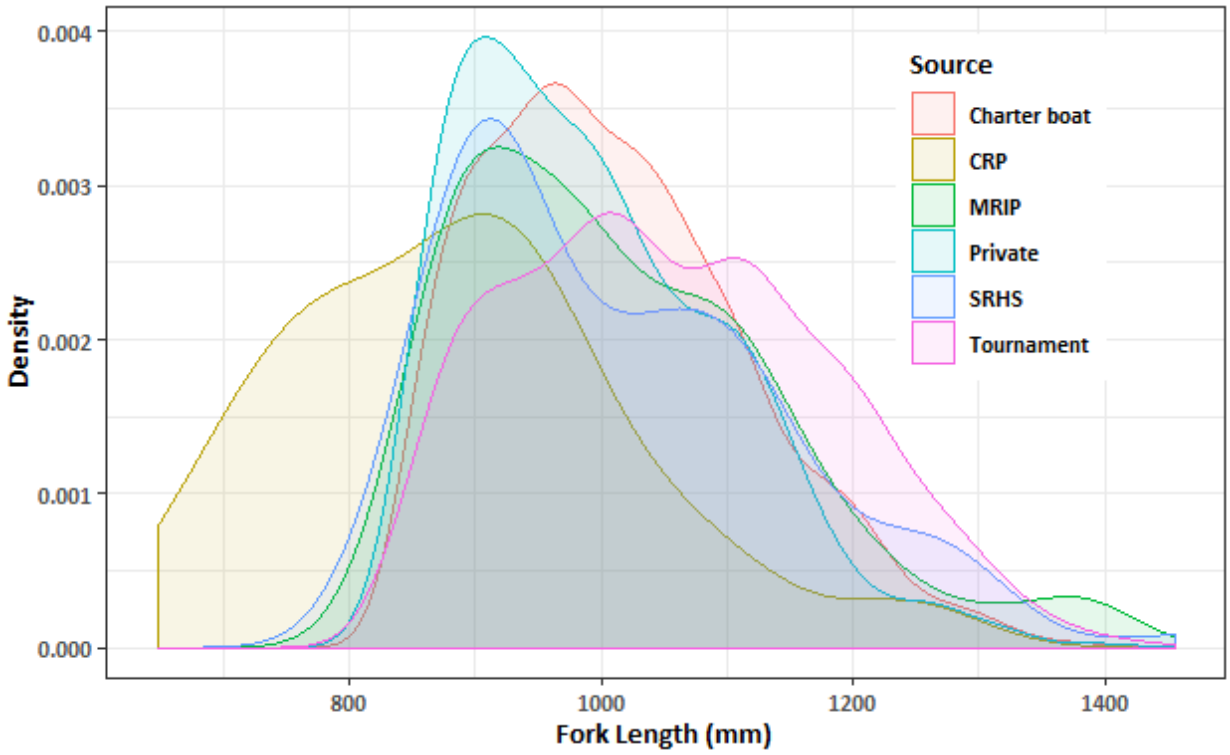


Figure 2: Density plot of the fork length comparisons in six cobia data sets

General overlap between SCDNR fishery-dependent charter boat, SCDNR fishery-dependent private, MRIP, and SRHS samples

- SCDNR CRP samples are skewed to smaller sizes due to fishery biologist tagging fish under the legal size limit
- Tournament samples are skewed towards larger sized fish

Frequency curve by data source

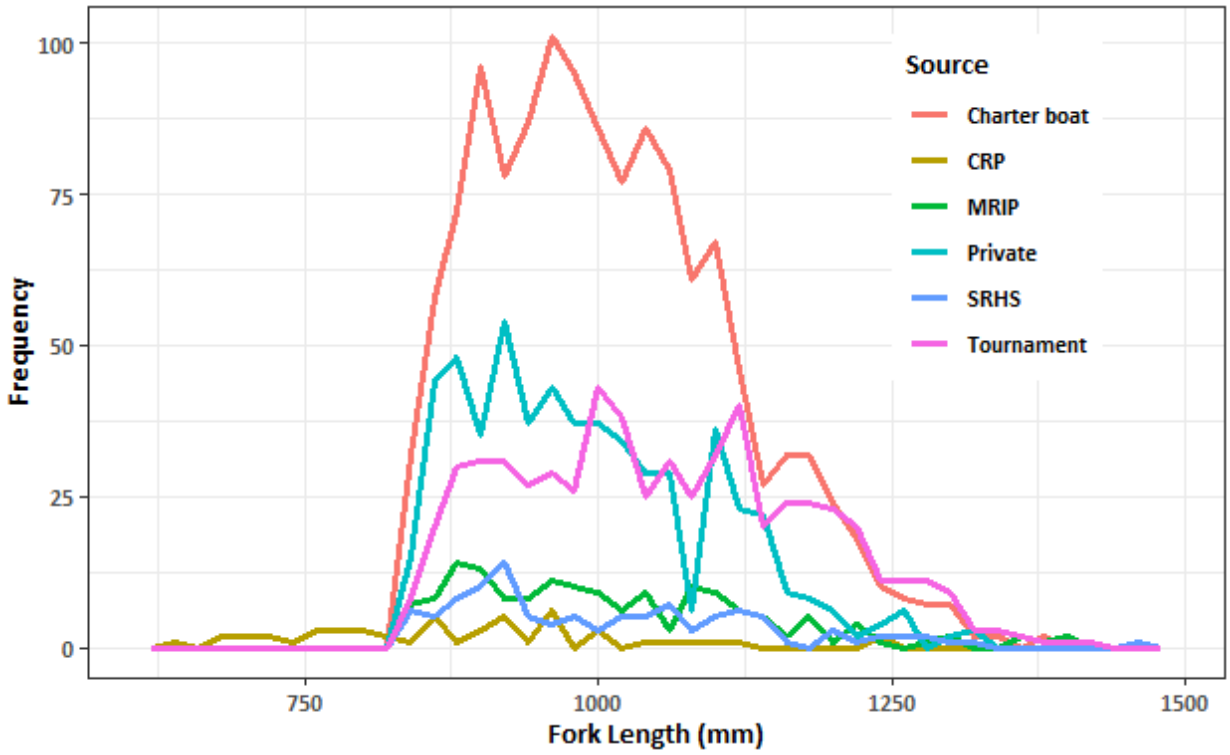


Figure 3: Frequency curve, by fork length (mm), from six cobia data collection sources.

South Carolina fishery-dependent, non-project samples (charter boat, private, and tournament) were the most common.

- NMFS fishery-dependent and SCDNR CRP sources are very few, in comparison (MRIP, SRHS, and CRP)

Kolmogorov-Smirnov Test
(nonparametric test of the equality of continuous, one-dimensional probability distributions; statistic quantifies a distance between the empirical distribution function of two samples, with the null being that the two samples are drawn from the same distribution; test is one of the most useful and general nonparametric methods for comparing two samples, as it is sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two samples)

- Significant Results, i.e., there is a difference between the compared distributions – significance based on Bonferroni adjusted p-value, adjusting for the 15 pairwise comparisons made.

Table 2: Results of significant Kolmogorov-Smirnov tests, sorted by most significant (smallest adjusted p-value based on Bonferroni correction) to largest adjusted p-value.

Group1	Group2	D	p.value	p.adjusted
Private	Tournament	0.2306	0.0000	0.0000
Charter boat	Tournament	0.1851	0.0000	0.0000
CRP	Tournament	0.5289	0.0000	0.0000
Charter boat	CRP	0.4329	0.0000	0.0000
CRP	Private	0.4025	0.0000	0.0000
CRP	MRIP	0.4074	0.0000	0.0001
CRP	SRHS	0.3916	0.0000	0.0007
SRHS	Tournament	0.2173	0.0003	0.0046
MRIP	Tournament	0.1864	0.0004	0.0053
Charter boat	Private	0.0963	0.0013	0.0196

Non-Significant Results – i.e., there is **NOT a DIFFERENCE** between the compared distributions – significance based on Bonferroni adjusted p-value adjusting for the 15 pairwise comparisons made.

Table 3: Results of non-significant Kolmogorov-Smirnov tests, sorted by least significant (largest adjusted p-value based on Bonferroni correction) to smallest adjusted p-value.

Group1	Group2	D	p.value	p.adjusted
Charter boat	MRIP	0.0790	0.3464	1.0000
MRIP	SRHS	0.0952	0.6010	1.0000
MRIP	Private	0.1003	0.1681	1.0000
Private	SRHS	0.1032	0.2800	1.0000
Charter boat	SRHS	0.1486	0.0227	0.3403

Kruskal-Wallis & Pairwise Comparisons of Group Medians using a Wilcoxon Rank Sum Test
(Kruskal-Wallis is a non-parametric one-way ANOVA; non-parametric method of testing whether samples originate from the same distribution which is used for comparing two or more samples of equal or different sample sizes; the null hypothesis is that the medians of all groups are equal, and the alternative is that at least one population median of one group is different from the population median of at least one other group)

(Wilcoxon Rank Sum Test is a non-parametric test to compare two related samples; the non-parametric equivalent to the t-test; used to identify which pairs of groups were different at the 95% confidence level after use of the Bonferroni multiple comparison correction)

Table 4: P-values from pairwise comparisons using Wilcoxon rank sum test, using the Bonferroni method for p-value adjustment. Shaded cells represent significant results. The significant and non-significant results are identical to the significant and non-significant results from the Kolmogorov-Smirnov pairwise comparisons.

	Charter boat	CRP	MRIP	Private	SRHS
CRP	0.0000	-	-	-	-
MRIP	1.0000	0.0000	-	-	-
Private	0.0110	0.0000	1.0000	-	-
SRHS	1.0000	0.0000	1.0000	1.0000	-
Tournament	0.0000	0.0000	0.0020	0.0000	0.0090

Conclusions

- The length comps from the CRP data set are different from the other data sources due to researchers capturing and acoustically tagging individual fish under the recreational size limit.
- SCDNR charter boat is not significantly different from the MRIP or SRHS data sources, but is significantly different from the other fishery-dependent, non-project data sources collected in South Carolina (SCDNR Private and SCDNR Tournament)
    - The result of the statistical power to detect differences amongst these data sets because they represent the sources with the largest sample sizes.
- As expected, the length frequencies from fish caught from tournaments in South Carolina are shifted to larger sizes relative to other data sources
- It is also comforting that there is no difference between MRIP and SRHS or Private data sources, or between Private and SRHS data sources.
- Overall, it is safe to assume that the MRIP, Charter boat, SRHS, and Private collections are all more or less sampling the same segment of the cobia population and could thus merge these data to develop length compositions for the South Carolina recreational fishery.

South Carolina Fishery-Dependent, non-project related additional data analyses

The remaining comparisons utilized only SCDNR's fishery-dependent sampling for further analyses. For these data sets, we have additional information available for each measured fish, including year of collection, source, sex, and location (inshore and offshore). These variables are important in understanding why datasets may vary over time and what variables may more strongly influence the datasets.

It has been well established that cobia yearclass strength is highly variable (SEDAR 28) depending on recruitment. The 2007 yearclass being the largest yearclass reported and significantly impacted the catch in the years following recruitment.

In SC, there is a well-documented (SEDAR 28) inshore spawning aggregation (Distinct population segment, DPS) that remains spatially separate inshore from the larger offshore US South Atlantic population. This DPS was more easily accessed by anglers and was fished very heavy until abundance decreased and eventually the fishery was closed in state waters. Fish from this population were often the segment caught and reported at tournaments. As this distinct population segment was fished down size of fish decreased and anglers moved into offshore waters to catch larger and more abundant fish from a presumably much larger population.

Lastly cobia are sexually dimorphic and the size distribution of the catch can be influenced because regulations may influence the proportions of the faster growing females from the catch.

In effect, we have already investigated potential differences amongst the sources of the samples above. Here, we will further investigate the impact that the other available parameters may have on apparent length compositions.

Year

Sample Summary Information

Table 5: Summary statistics by year. Provided is the sample size (n), mean (avg), standard deviation (SD), standard error (SE), median, and inter-quartile range of fork length (in mm) by year.

Year	n	Avg	SD	SE	Median	Inter-quartile Range
2007	336	966	99	5	948	114
2008	283	1012	104	6	1000	144
2009	220	1027	113	8	1011	148
2010	208	1019	111	8	1012	174
2011	225	1031	131	9	1008	214
2012	223	1006	109	7	980	157
2013	299	1027	107	6	1021	154
2014	302	1036	104	6	1029	150
2015	184	1034	113	8	1032	183
2016	139	1035	101	9	1030	151

Box Plot by year

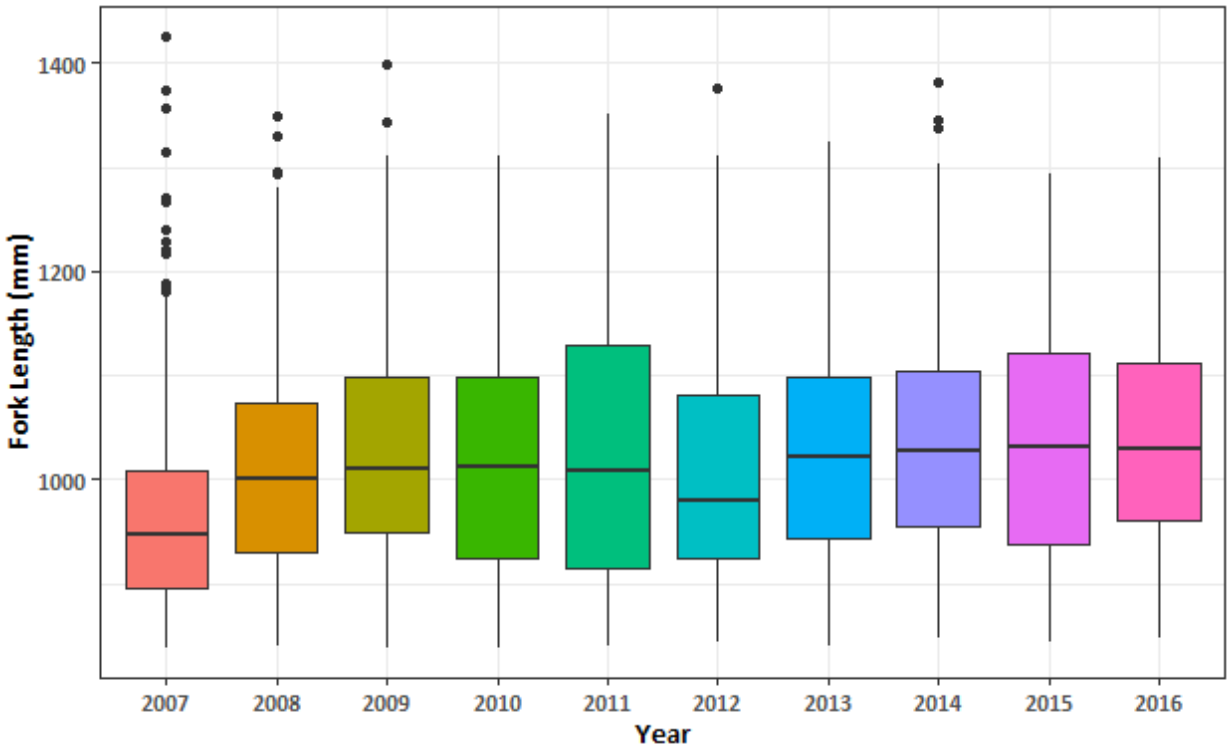


Figure 4: Box plot of South Carolina fishery-dependent samples examining the fork length (mm) distribution of cobia captured by year

Density Plot by Year

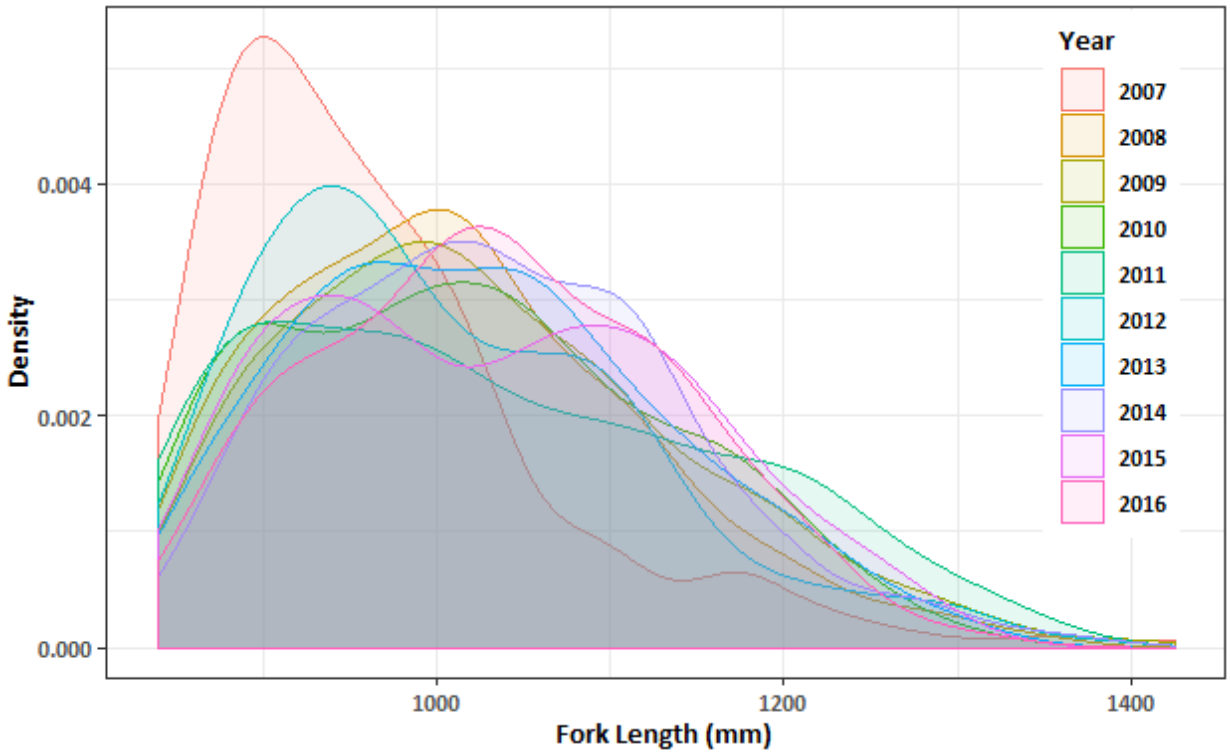


Figure 5: Density plot of the fork length comparisons in SCDNR's fishery-dependent samples by year

Kruskal-Wallis Test and sub-sequent Wilcoxon Test

Table 6: P-values from pairwise comparisons using Wilcoxon rank sum test, using the Bonferroni method for p-value adjustment. Shaded cells represent significant results.

	2007	2008	2009	2010	2011	2012	2013	2014	2015
2008	0.0000								
2009	0.0000	1.0000							
2010	0.0000	1.0000	1.0000						
2011	0.0000	1.0000	1.0000	1.0000					
2012	0.0001	1.0000	1.0000	1.0000	1.0000				
2013	0.0000	1.0000	1.0000	1.0000	1.0000	0.5449			
2014	0.0000	0.1544	1.0000	1.0000	1.0000	0.0147	1.0000		
2015	0.0000	1.0000	1.0000	1.0000	1.0000	0.4236	1.0000	1.0000	
2016	0.0000	0.5599	1.0000	1.0000	1.0000	0.1235	1.0000	1.0000	1.0000

Conclusions

Results suggest that the length composition observed in 2007 was different from other years, but otherwise seems to be pretty stable year to year. Thus it is probably safe to pool data across years when sample sizes are small. In addition, age data collected in 2007 (SCDNR pers. comm.) indicated a strong

2004year class entered the fishery as age 3 fish in 2007. This increase in younger fish drives the overall length frequency to smaller sizes in 2007.

Sex

Sample Summary

Table 7: Summary statistics by sex. Provided is the sample size (n), mean (avg), standard deviation (SD), standard error (SE), median, and inter-quartile range of fork length by sex.

Sex	n	Avg	SD	SE	Median	Inter-quartile Range
Female	1174	1064	110	3	1058	157
Male	970	959	82	3	948	117
Unknown	275	1016	107	6	1003	180

Box Plot by Sex

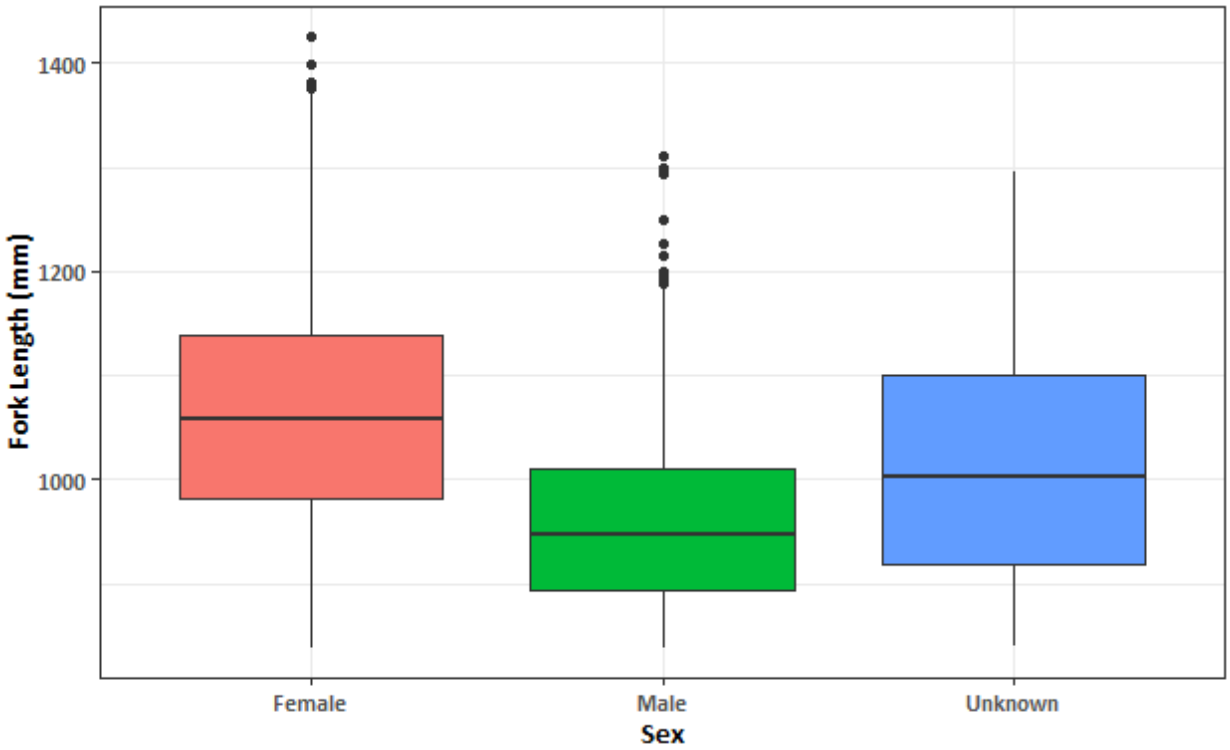


Figure 6: Box plot of South Carolina fishery-dependent samples examining the fork length (mm) distribution of cobia captured by sex

Density Plot by Sex

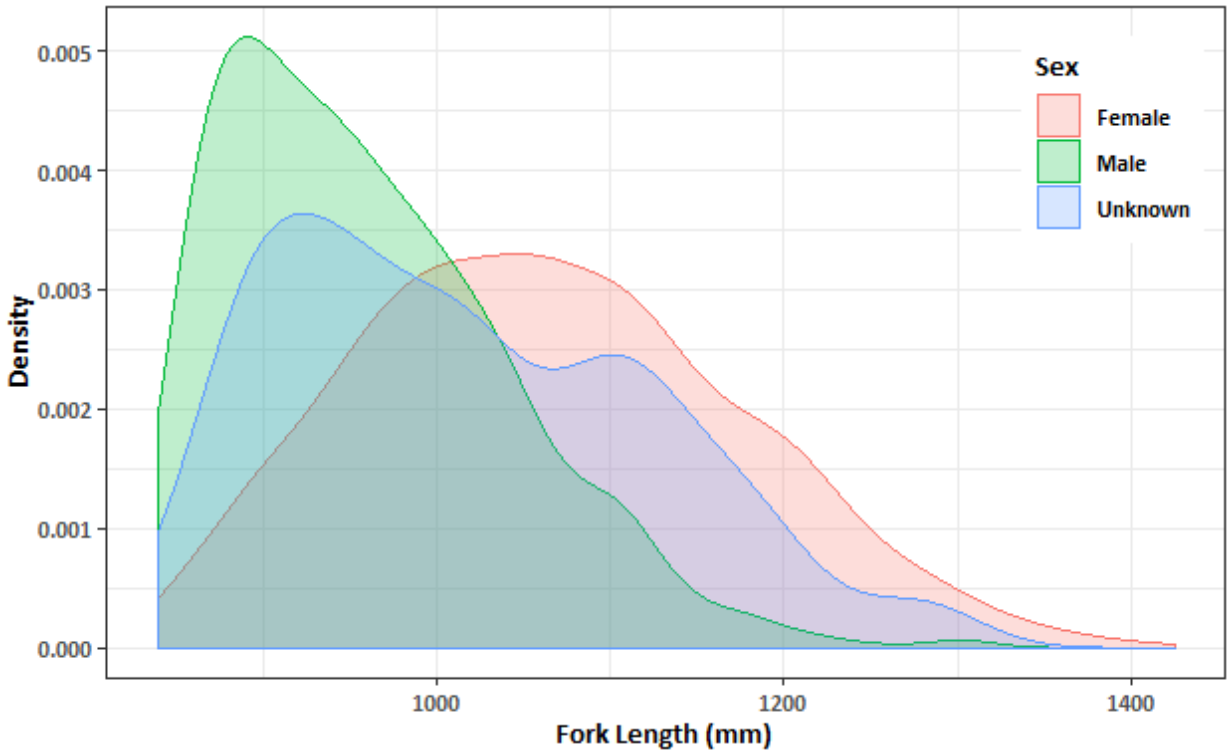


Figure 7: Density plot of the fork length comparisons in SCDNR's fishery-dependent samples by sex

Kruskal-Wallis and sub-sequent Wilcoxon Test

Table 8: P-values from pairwise comparisons using Wilcoxon rank sum test, using the Bonferroni method for p-value adjustment. Shaded cells represent significant results.

	Female	Male
Male	0.0000	
Unknown	0.0000	0.0000

Conclusions

There is a difference between the length distribution of male vs. female cobia in the SCDNR fishery-dependent collections, with females being larger. This is expected given the dimorphic growth of the species. With that said, it is unlikely we will ever have sex information allowing us to split the total catch-at-length or catch-at-age information by sex. Thus, while interesting biologically, we can't really use this information in the assessment model.

Location

Summary Statistics

Table 9: Summary statistics by location. Provided is the sample size (n), mean (avg), standard deviation (SD), standard error (SE), median, and inter-quartile range of fork length (in mm) by location.

Location	n	Avg	SD	SE	Median	Inter-quartile Range
Inshore	648	1000	112	4	984	167
Offshore	1225	1038	110	3	1028	161
Unknown	546	988	103	4	972	136

Box Plot by Location

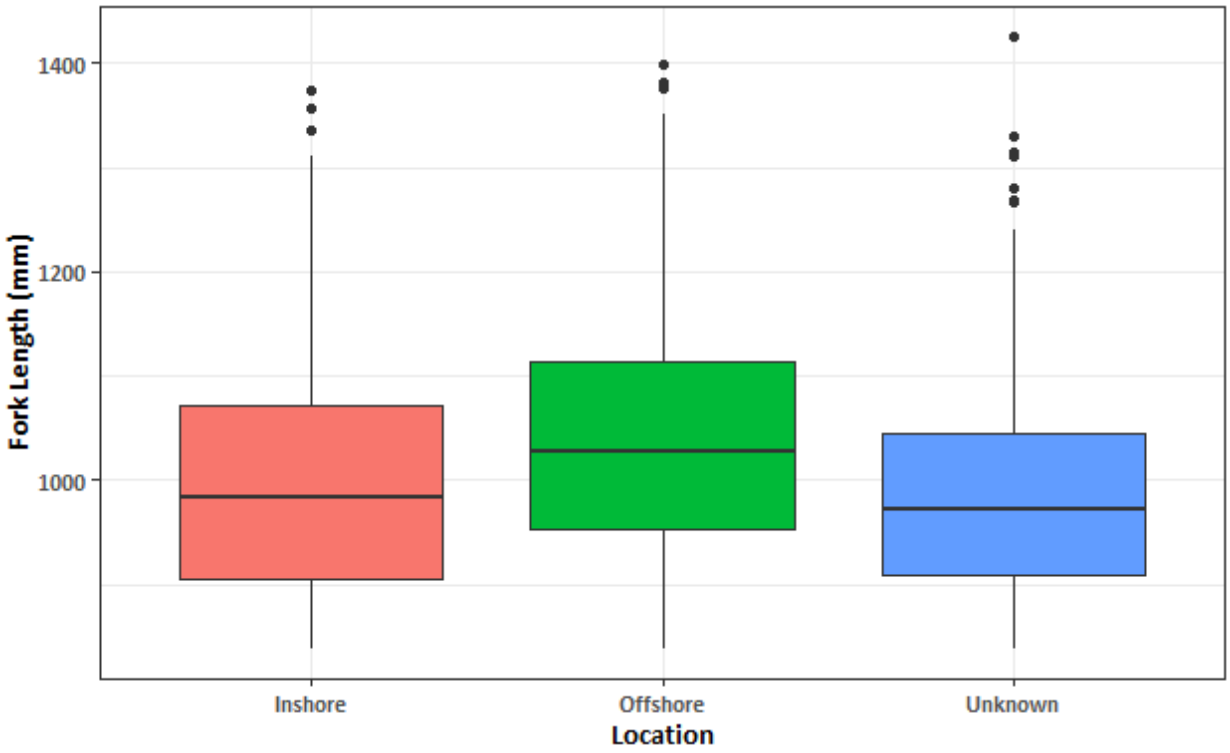


Figure 8: Box plot of South Carolina fishery-dependent samples examining the fork length (mm) distribution of cobia by capture location (inshore vs offshore)

Density Plot by Location

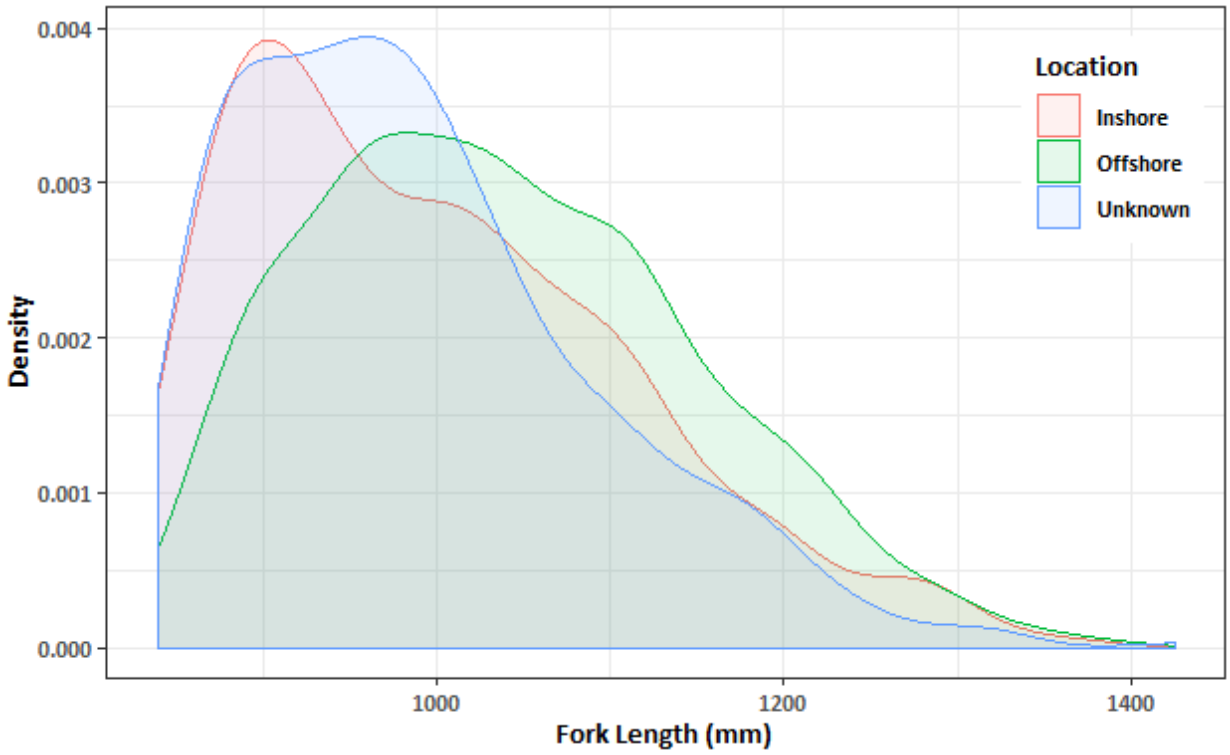


Figure 9: Density plot of the fork length comparisons in SCDNR's fishery-dependent samples by capture location (inshore vs offshore)

Kruskal-Wallis Test and sub-sequent Wilcoxon Test

Table 10: P-values from pairwise comparisons using Wilcoxon rank sum test, using the Bonferroni method for p-value adjustment. Shaded cells represent significant results.

	Inshore	Offshore
Offshore	0.0000	
Unknown	0.4622	0.0000

Conclusion

There is a difference in the length distribution of inshore vs offshore caught fish, with the offshore fish being larger than the inshore captured fish. This potentially could be explained by the distinct population segment (DPS) our genetic evaluation has determined between the inshore and offshore captured fish. This distinct inshore population segment is more accessible to recreational anglers and appears to be overfished due to the limited inshore catch and the smaller individuals captured in recent years. Based on these results, it appears that the "unknown" capture location fish likely represents some inshore and some offshore fish, though the median is not significantly different from the inshore captured fish indicating the majority are likely from inshore waters.

Source and Location (because of the generally highly significant effect of fishery-dependent, non-project source and capture location in the above analyses)

Summary Statistics

Table 11: Summary statistics by source and location. Provided is the sample size (n), mean (avg), standard deviation (SD), standard error (SE), median, and inter-quartile range of fork length (in mm) by data source. Table is sorted from the largest to smallest mean fork length of individual fish by data source and location.

Source	Location	n	Avg	SD	SE	Median	Inter-quartile Range
Tournament	Offshore	285	1105	111	7	1109	159
Tournament	Inshore	191	1024	120	9	1009	185
Private	Offshore	171	1020	96	7	1014	152
Charter	Offshore	769	1017	102	4	1005	143
Charter	Unknown	225	1004	104	7	993	144
Charter	Inshore	298	997	110	6	981	167
Tournament	Unknown	124	987	114	10	958	126
Private	Inshore	159	974	99	8	940	146
Private	Unknown	197	972	90	6	960	135

Box Plot by Source and Location

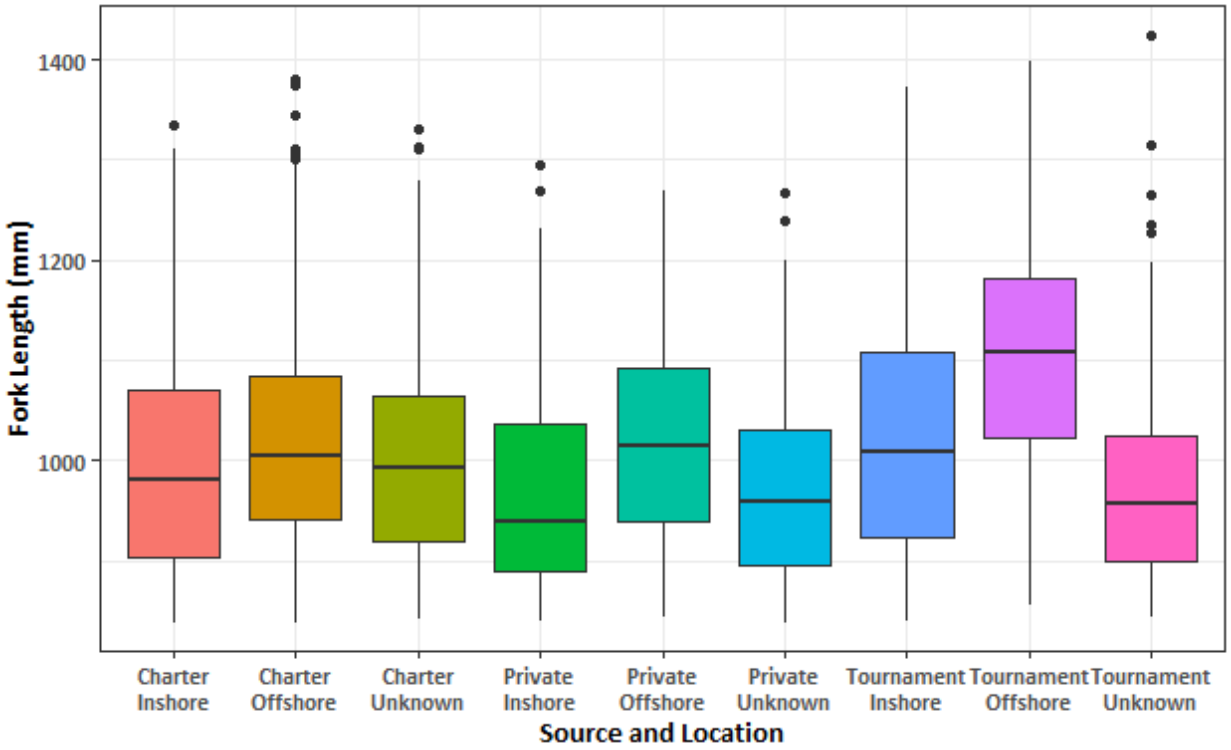


Figure 10: Box plot of South Carolina fishery-dependent samples examining the fork length (mm) distribution of cobia by SCDNR sample source (charter, private and tournament) and capture location (inshore vs offshore)

Density Plot by Source and Location

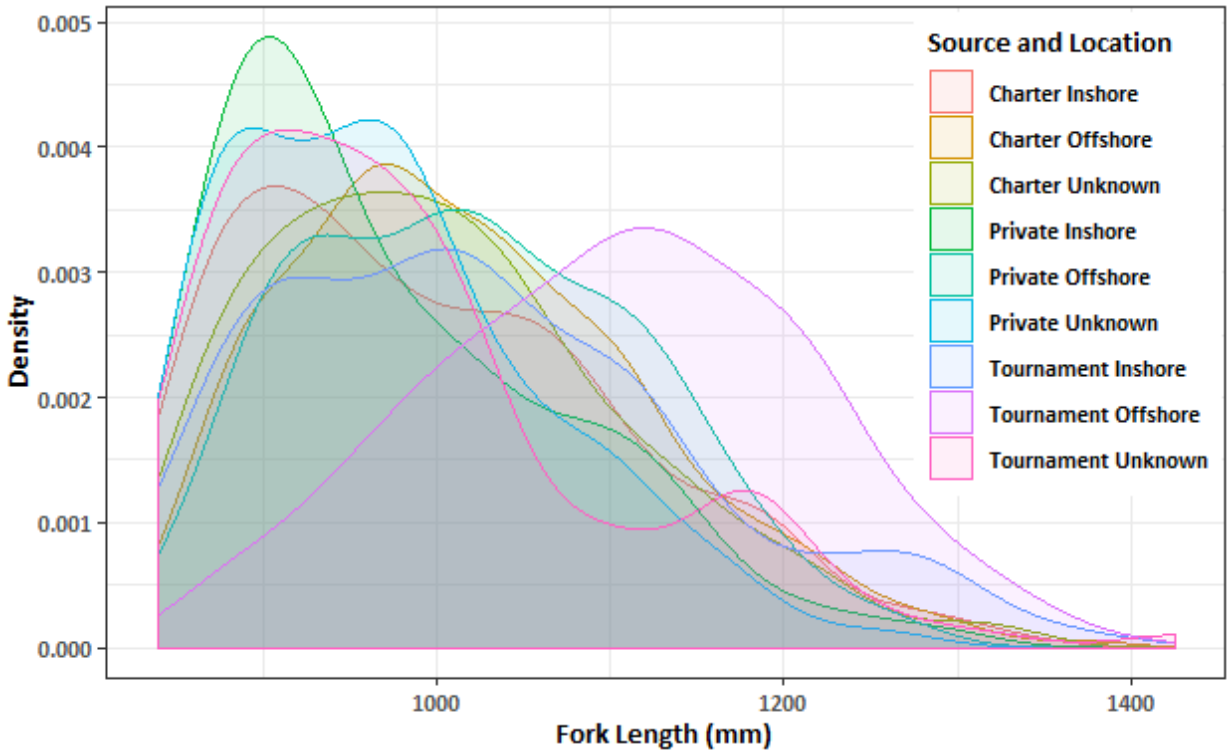


Figure 11: Density plot of the fork length comparisons in SCDNR's fishery-dependent samples by sample source (charter, private and tournament) and capture location (inshore vs offshore)

Kruskal-Wallis Test and sub-sequent Wilcoxon Test

Table 12: P-values from pairwise comparisons using Wilcoxon rank sum test, using the Bonferroni method for p-value adjustment. Shaded cells represent significant results.

		Charter Boat			Private			Tournament	
		Inshore	Offshore	Unknown	Inshore	Offshore	Unknown	Inshore	Offshore
Charter Boat	Offshore	0.0398							
	Unknown	1.0000	1.0000						
Private	Inshore	1.0000	0.0000	0.1119					
	Offshore	0.2052	1.0000	1.0000	0.0002				
	Unknown	1.0000	0.0000	0.0616	1.0000	0.0000			
Tournament	Inshore	0.5728	1.0000	1.0000	0.0026	1.0000	0.0004		
	Offshore	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
	Unknown	1.0000	0.0053	1.0000	1.0000	0.0219	1.0000	0.0745	0.0000

Conclusion

In this analysis, regardless of data source, fish identified as being caught offshore within the three SCDNR data sets (Charter boat, Private, Tournament) are significantly larger than inshore captured fish,

with the tournament captured offshore fish significantly larger than fish in any other data type. Charter boat inshore captured fish tend to not be significantly different from other data sources, with the exception of offshore captured charter boat and tournament captured fish. There is no difference in size of private and charter boat captured offshore fish, between tournament captured inshore fish and charter boat captured offshore fish, nor between private captured offshore fish and tournament captured inshore fish. There is a significant difference between charter boat captured offshore fish and private inshore fish and between private inshore fish and tournament offshore fish. These results agree with the previous analyses suggesting the offshore captured fish are significantly larger than inshore captures and that the tournament captured fish are significantly larger than any other SCDNR's fishery-dependent data sources.

Summary

Summarizing the SCDNR fishery-dependent datasets, it appears that size of fish can be influenced by year class strength with strong year classes influencing the catch for a number of years after they enter the fishery. There is a difference between the length distribution of male vs. female cobia, with females being larger, which was expected given the dimorphic growth of the species. Location of capture also has an effect on the apparent length distribution with larger fish being captured offshore compared to inshore catches within South Carolina's estuaries which is likely due to the genetic determination of an inshore DPS and its accessibility to recreational anglers. There seems to be a bias (not unexpected) towards larger fish being reported from tournaments where fish are mostly caught in offshore waters.

Comparisons between the NMFS fishery-dependent sources and SCDNR fishery-dependent sources suggest the CRP fishery-dependent samples were significantly smaller than all other data sources due to biologist acoustically tagging undersized fish. South Carolina charter boat samples are not significantly different from the MRIP or SRHS data sources, but is significantly different from the other fishery-dependent, non-project data sources collected in South Carolina (Private and Tournament); this is likely the result of the statistical power to detect differences amongst these data sets because they represent the sources with the largest sample sizes. As expected, the length frequencies from fish caught from tournaments in South Carolina are shifted to larger sizes relative to other data sources. No differences between MRIP and SRHS or Private data sources, or between Private and SRHS data sources were determined.

Finally, it is safe to assume that the MRIP, Charter boat, SRHS, and Private collections are all sampling the same segment of the cobia population and therefore these data could be pooled.

References

Hoare, D., Graham, N. & Schon, P. (2011). The Irish Sea data-enhancement project: comparison of self-sampling and national data-collection programmes – results and experiences. ICES Journal of Marine Science 68, 1778–1784. doi: 10.1093/icesjms/fsr100