Cluster Analysis for the Puerto Rico Island Region
NOAA Fisheries
Southeast Regional Office
St. Petersburg, FL
March 7, 2016
SERO-LAPP-2016-03

Introduction

The U.S. Caribbean is composed of the Commonwealth of Puerto Rico and the Territory of the U.S. Virgin Islands. The Caribbean Fishery Management Council (Council) is currently developing island-based fishery management plans (FMPs) specific for each island region (Puerto Rico, St. Thomas/St. John, and St. Croix). These FMPs would incorporate and replace the current Caribbean-wide FMPs for Corals and Reef Associated Plants and Invertebrates, Reef fish, Queen Conch, and Spiny lobster of Puerto Rico and the U.S Virgin Islands. The purpose of the individual island FMPs is to tailor management to the uniqueness of each island region with respect to their biological, ecological, economic, and cultural resources.

As part of the development of these plans and their associated environmental impact statements, the Council is proposing a list of species to be managed on each of the islands. The Council will then evaluate if these species should be managed at the individual level and/or in groups. For this purpose, the FMP's interdisciplinary planning team requested a cluster analysis that evaluates different grouping scenarios and that could potentially be used for grouping the species for each island region. The purpose of this report is to conduct a cluster analysis to determine potential species groupings for species proposed to be managed in Puerto Rico.

Methods

The Southeast Fisheries Science Center provided Puerto Rico commercial landings data on November 4, 2015 and recreational landings data on February 2, 2016 to the NOAA Fisheries Southeast Regional Office. The data contained landings for every individual commercial and recreational trip. The commercial data had information on species landed, depth, and gear. The recreational data had information on species landed and gear. The data was filtered so only data from 2010 through 2014 remained. These years were chosen since they are the most recent years of complete data and will likely reflect current and future catches. In August of 2015 the Council chose a list of species for management for each island region. The Council chose 66 species of finfish and also spiny lobster and queen conch for the St. Croix region. In December of 2015 the Caribbean Council's Scientific and Statistics Committee recommended removing spiny lobster and queen conch from the cluster analysis. Also in December of 2015 the Council removed guaguanche from the Puerto Rico list of species. Therefore, the data were filtered so only data for the 64 species of finfish remained.

The data was converted to a presence/absence format. Specifically, if a fish was caught on a trip it was given a 1 and if it was not caught on a trip it was given a 0. This format was chosen to allow each fish species to have equal weight. The cluster analysis was done with R software

(www.r-project.org) with the hclust function. The "AVERAGE" clustering method was used which is a hierarchical method that calculates the distance between clusters by taking the average of all pairwise differences between the points within each cluster.

The cluster analysis for the commercial data was first run for all of the 64 species of finfish. However, the fisher's catch report form only lists 23 of the 64 species chosen for management by the Council. Therefore 36% of the species chosen for management are not listed on the catch form. Any landings for the species not listed on the form were reported in the write-in spaces provided. Unfortunately, it's unknown if the fishers report all of the write-in species every time they report landings or if the fishers only occasionally list the landings for these species not listed on the form. An additional cluster analysis was run that excluded all of the write-in species and only included the species that were listed on the catch form.

Additional cluster analyzes of the commercial data were done by separating the data into the gear reported for the trip. The gears were diving, hook & line, trap, and nets.

Cluster analysis on the commercial data were also done by depth, however not all of the trips recorded depth. Any trips without depth information were removed from the cluster analysis by depth. The remaining trips were separated into depth bins. An examination of the distribution of depths per trip showed two primary depth bins: 1) trips that recorded a depth of 100 feet or less, and 2) trips that recorded a depth greater than 100 feet. These two depth bins each had more than 300 trips and were used to create additional cluster analysis results.

Cluster analyzes of the recreational data were not done for each gear because there were not enough samples for each gear type. The majority of the recreational trips from 2010 through 2014 reported the gear type of hook & line (95%, n = 943). The other trips reported the gears of nets (n = 6) and spear (n = 39) and did not have an adequate number of trips to do the cluster analysis.

A cluster analysis was also done combining both commercial and recreational data. First, a similarity matrix was calculated for both commercial and recreational trips. Then the matrixes were filtered so only species caught on both commercial and recreational trips remained. This was required to combine both the commercial and recreational matrices. The matrices were combined using two methods. The first method gave equal weight to each sector and calculated the average similarity for each species. The second method used a weighted average where the weight came from the percent of landings the sector contributed in the past 5 years. The commercial sector contributed 88% of the landings for the 64 relevant species from 2010-2014, where the recreational sector contributed 12% of the landings for the same time period. Therefore, the commercial sector similarity matrix was given more weight in the second method.

Results

The data was filtered for the years of 2010 through 2014. Then the data was filtered to only contain records for the 64 species of finfish chosen for management by the Council. There were a total of 27,320 commercial trips, and cluster analysis results for these trips are shown in

Figure 1. The commercial data had 33 species that were chosen by the Council that did not have any landings from 2010-2014. These species were Nassau grouper, goliath grouper, black grouper, red grouper, tiger grouper, blue parrotfish, midnight parrotfish, rainbow parrotfish, dog snapper, graysby, rock hind, queen parrotfish, princess parrotfish, redtail parrotfish, redband parrotfish, striped parrotfish, blue tang, ocean surgeonfish, doctorfish, ocean triggerfish, gray triggerfish, puddingwife, Spanish hogfish, queen angelfish, gray angelfish, French angelfish, tripletail, manta, spotted eagle ray, wenchman, schoolmaster, crevalle jack, and pompano dolphin. Since no landings were available for these species from 2010-2014 they were not included in the cluster analysis results for Figure 1.

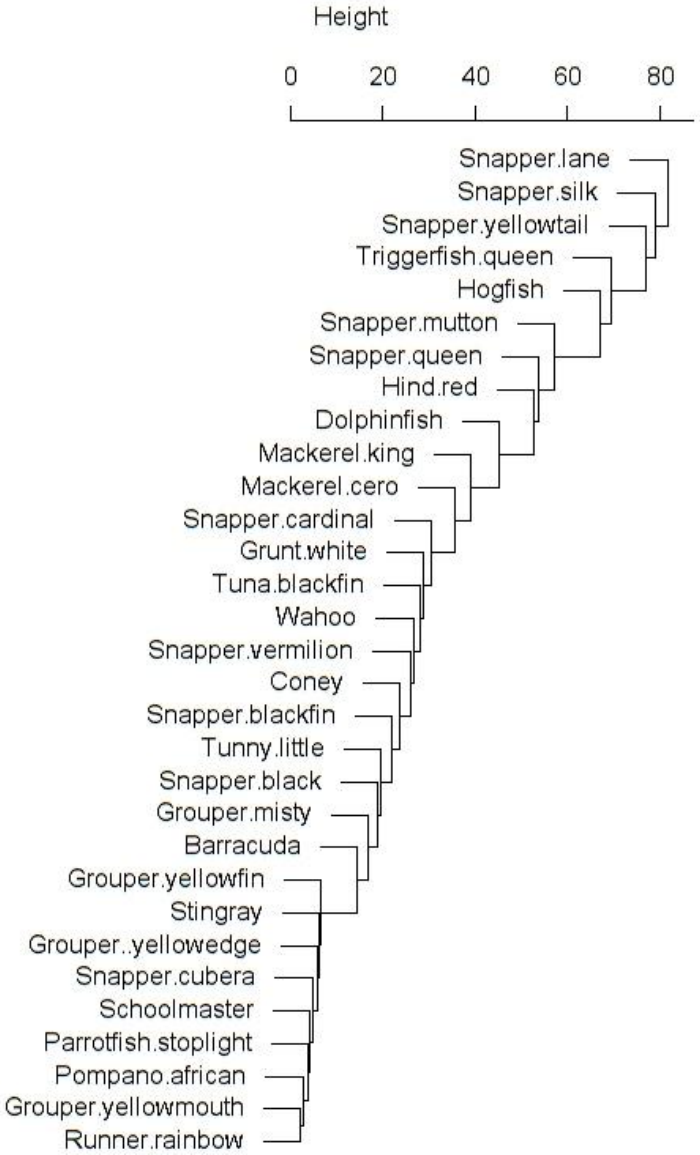

Figure 1. Dendrogram of fish species generated with a cluster analysis of the commercial data for the species chosen for management in Puerto Rico (n = 27,320 trips). The Caribbean council chose 64 species however 2010 through 2014 landings only had land for 31 of the 64 species.

The fisher's commercial catch report form only lists 23 of 64 species that the Council chose for management. The commercial data was filtered to only include the 23 species that were both

chosen for management and on the catch form. This resulted in 27,269 commercial trips, and a cluster analysis result of these data is shown in Figure 2.

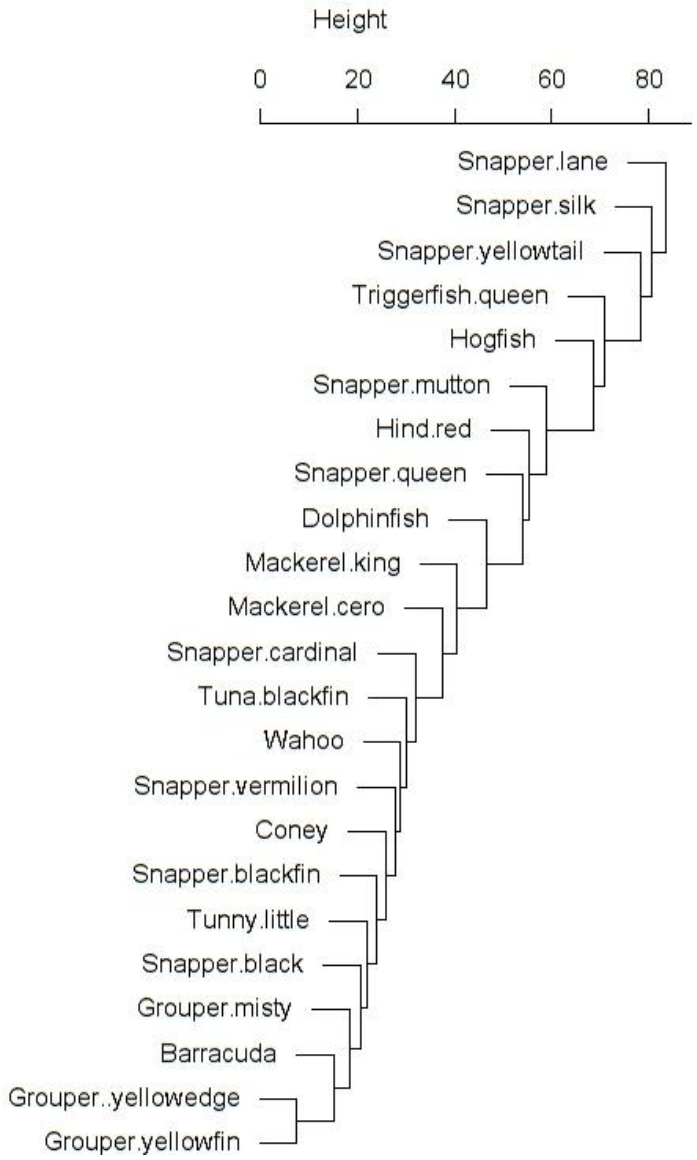

Figure 2. Dendrogram of fish species generated with a cluster analysis of commercial data for the 23 species that were both chosen for management and on the catch form in Puerto Rico (n = 27,269 trips).

The commercial trips were separated by gear and included all (if landings were available) of the 64 species chosen for management. There were 5,788 diving, 16,421 hook & line, 3,752 trap, and 1,359 net trips. A cluster analysis was done for all four of the gear types, and the results are shown in Figures 3, 4, 5, and 6.

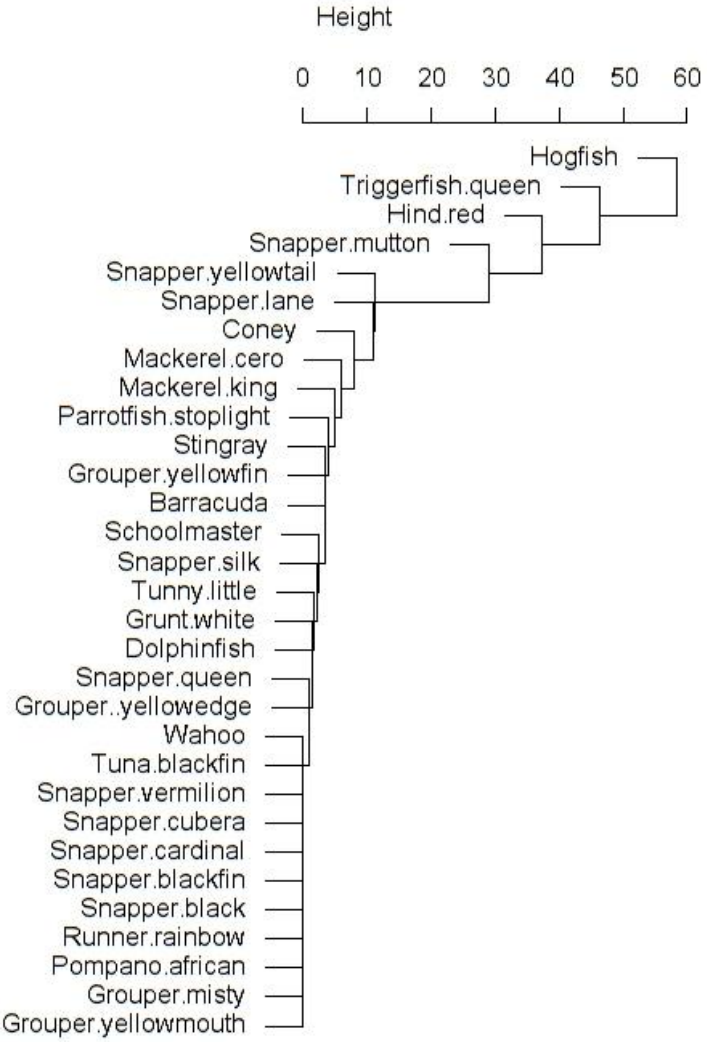

Figure 3. Dendrogram of fish species generated with a cluster analysis of commercial data for only the diving gear trips (n = 5,788 trips). The dendogram includes all of the 64 species chosen for management, however landings are not available for all of the species.

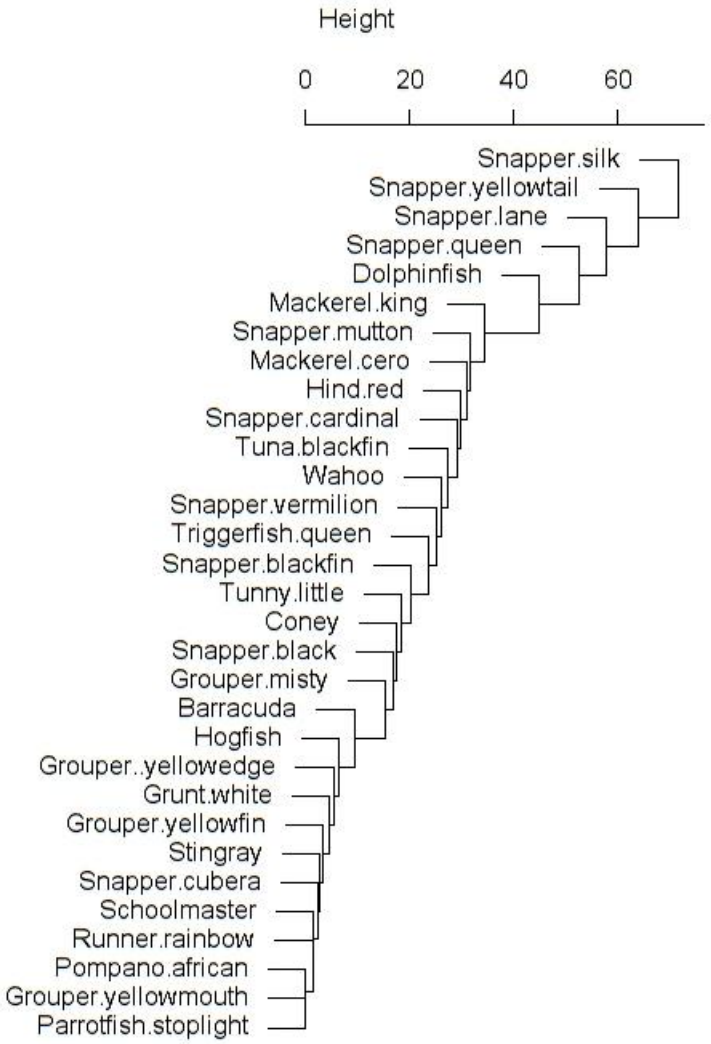

Figure 4. Dendrogram of fish species generated with a cluster analysis of commercial data for only the hook and line gear trips (n = 16,421 trips). The dendogram includes all of the 64 species chosen for management, however landings are not available for all of the species.

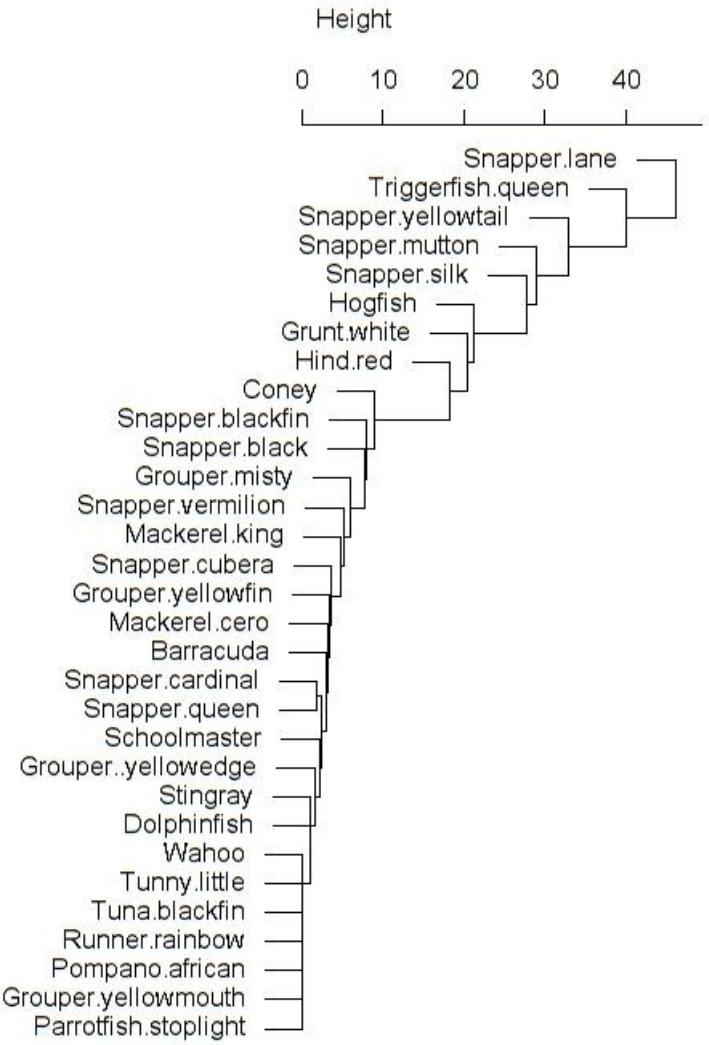

Figure 5. Dendrogram of fish species generated with a cluster analysis of commercial data for only the trap gear trips (n = 3,752 trips). The dendogram includes all of the 64 species chosen for management, however landings are not available for all of the species.

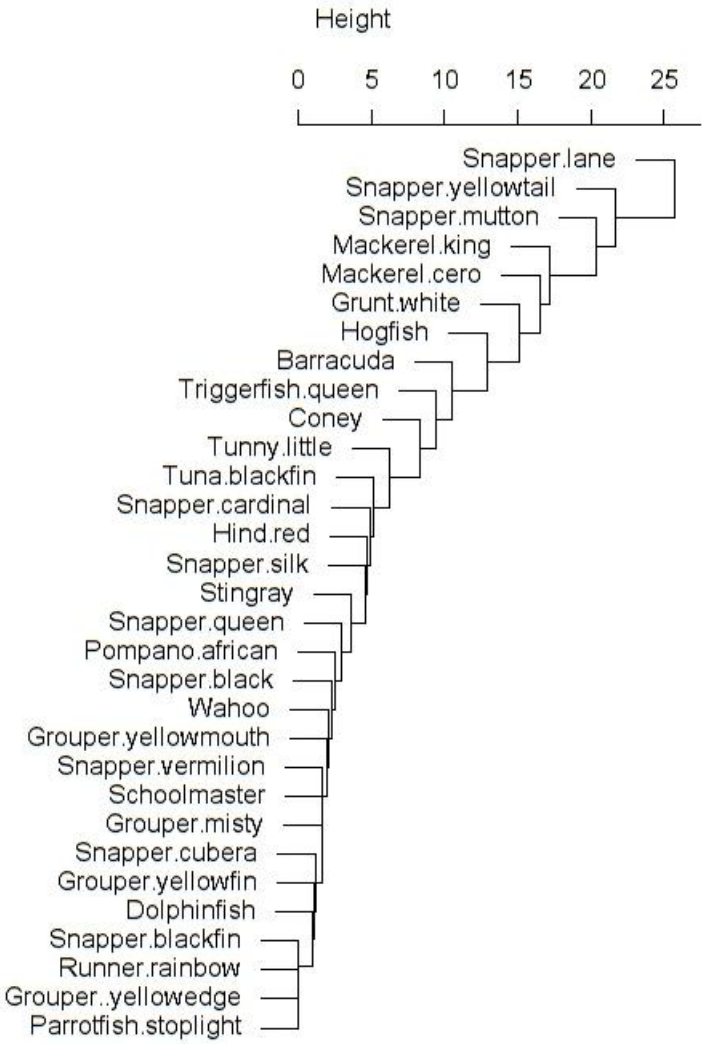

Figure 6. Dendrogram of fish species generated with a cluster analysis of commercial data for only the net gear trips (n = 1,359 trips). The dendogram includes all of the 64 species chosen for management, however landings are not available for all of the species.

The trips were separated into the depth bins of: 1) trips that recorded a depth of less than 100 feet, and 2) trips that recorded a depth of 100 feet or greater. This analysis included all (if landings were available) of the 64 species chosen for management. There were 14,825 trips that occurred less than 100 feet and 1,792 trips that recorded 100 feet or deeper. Some trips did not record any depth information (n = 10,703 trips, 39% of total trips) and were removed from the cluster analysis by depth. Cluster analysis results for the two different depth bins are shown in Figures 7 and 8.

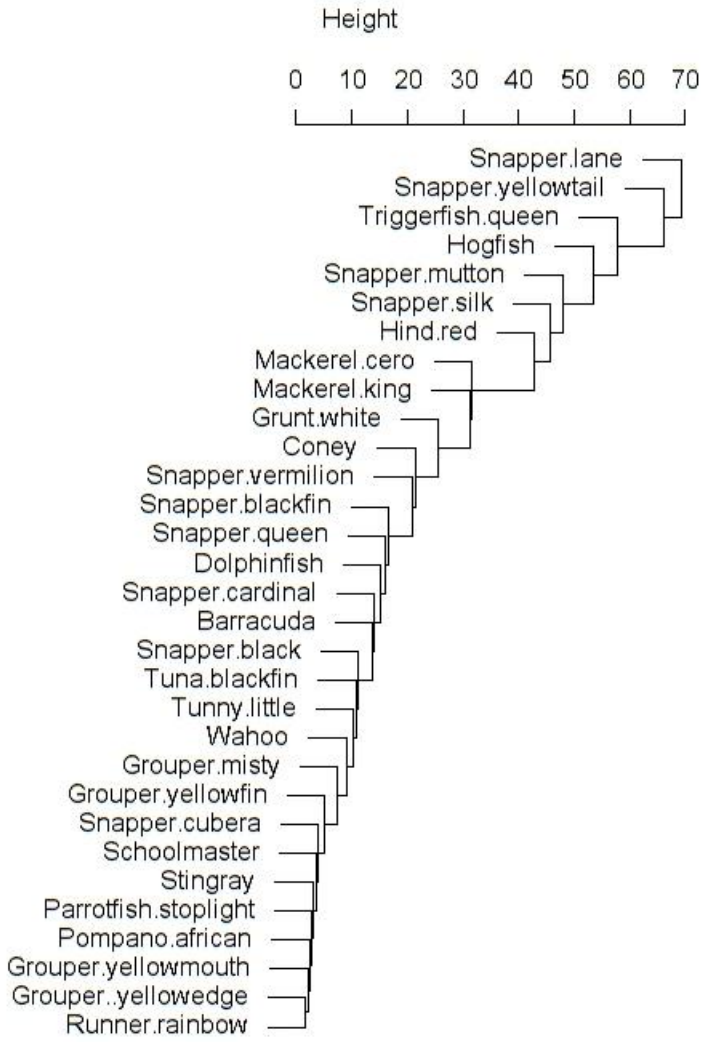

Figure 7. Dendrogram of fish species generated with a cluster analysis of commercial data for only trips that occurred in depths of 100 feet or less (n = 14,825 trips). The dendogram includes all of the 64 species chosen for management, however landings are not available for all of the species.

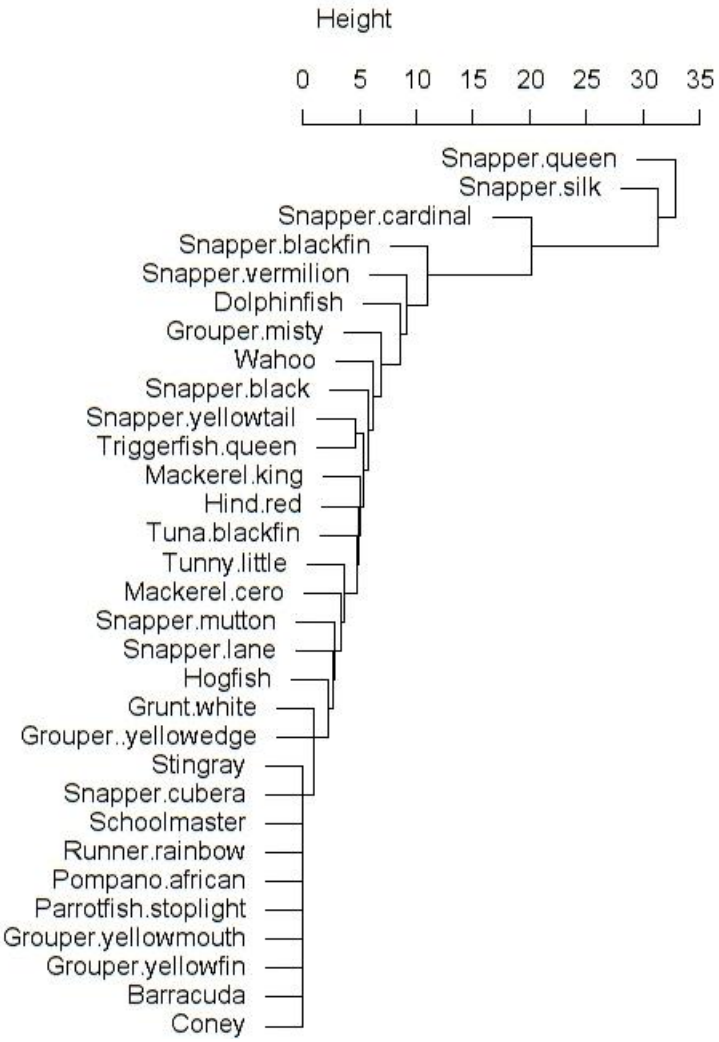

Figure 8. Dendrogram of fish species generated with a cluster analysis of commercial data for trips that occurred in depths greater than 100 feet (n = 1,792 trips). The dendogram includes all of the 64 species chosen for management, however landings are not available for all of the species.

The recreational data had 22 species that were chosen by the Council that did not have any landings from 2010-2014. These species were black snapper, goliath grouper, red grouper, tiger grouper, yellowfin grouper, yellowedge grouper, blue parrotfish, schoolmaster, coney, graysby, yellowmouth grouper, princess parrotfish, redband parrotfish, blue tang, ocean surgeonfish, doctorfish, queen angelfish, manta ray, cardinal snapper, African pompano, pompano dolphin, and cero mackerel. The recreational data for the 42 relevant species that did have landings from 2010-2014 produced a total of 995 recreational trips. Cluster analysis results for these trips are shown in Figure 9.

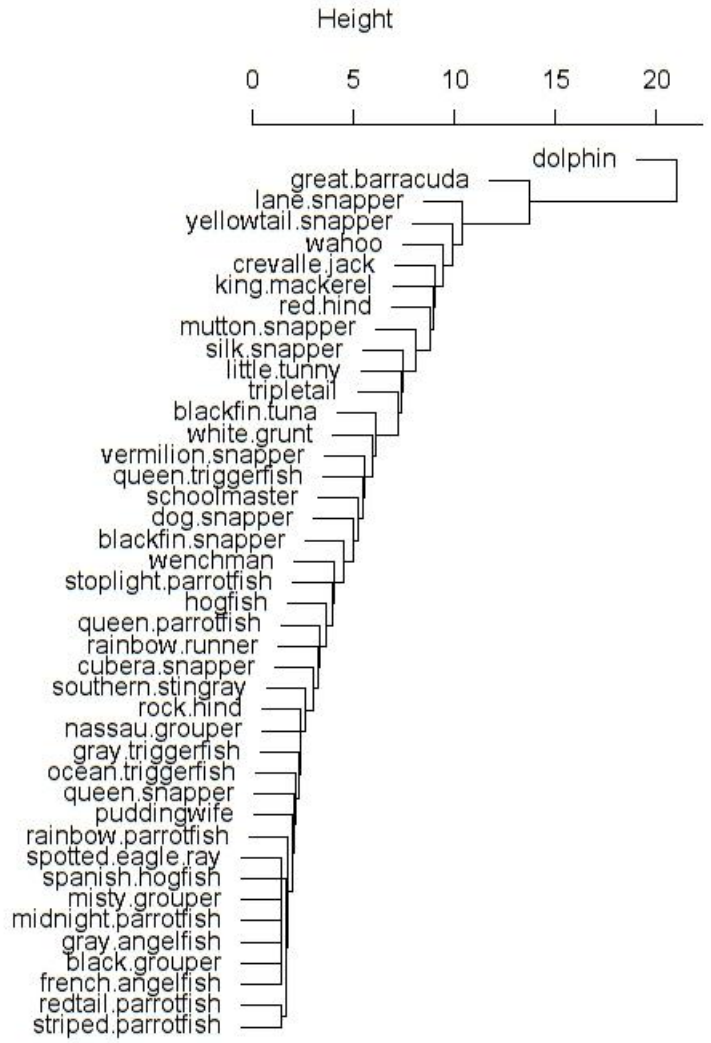

Figure 9. Dendrogram of fish species generated with a cluster analysis of the recreational data for the species chosen for management in Puerto Rico (n = 995 trips). The Caribbean council chose 64 species however 2010 through 2014 landings only had landings for 42 of the 64 species.

There were 23 species that were landed in both the commercial and recreational trips from 2010-2014. This data came from 27,639 trips (26,697 commercial and 942 recreational trips). The commercial and recreational cluster analyses similarity matrixes were combined with two different methods. The cluster analysis results for method 1 are shown in Figure 10 and the results for method 2 are shown in Figure 11.

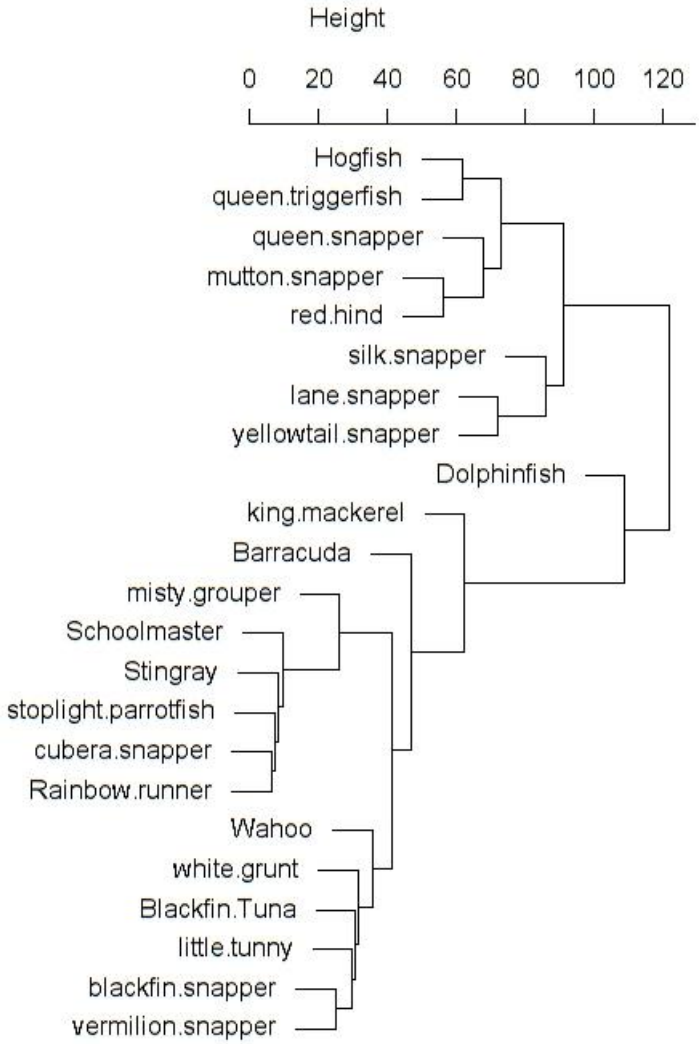

Figure 10. Dendrogram of fish species generated with a cluster analysis with both commercial and recreational data. This dendogram was generated from the first method which gave equal weight to each sector. The dendogram only includes species that had landings in both the commercial and recreational sectors.

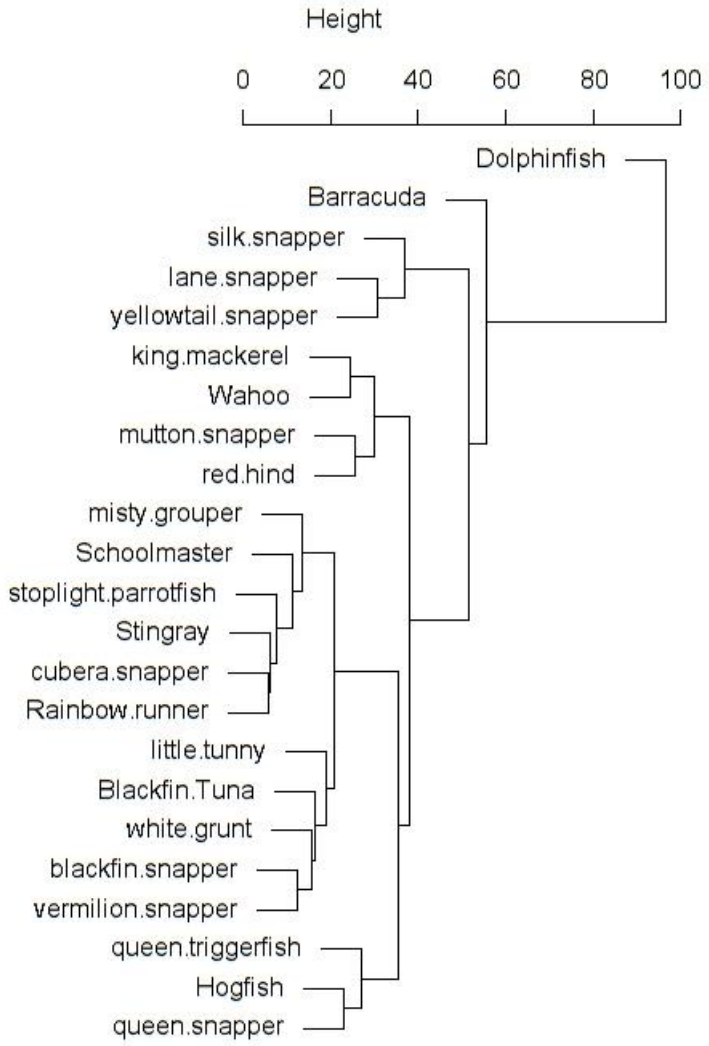

Figure 11. Dendrogram of fish species generated with a cluster analysis with both commercial and recreational data. This dendogram was generated from the second method which gave more weight to the commercial sector since it had higher landings. The dendogram only includes species that had landings in both the commercial and recreational sectors.

Discussion

Obviously, each trip does not catch all of the species the council chose for management. Some species were not caught on any of the trips in recent years, and other species were caught on a small number of trips. The species that were landed on a small number of trips are not able to be robustly separated in the cluster analysis. There simply is not enough data to form a robust distinction in the clusters. For example, in the commercial data the species of yellowmouth grouper and rainbow runner were on a combined total of 5 trips, and on Figure 1 they are shown close to the bottom of the dendogram without clear separation of a cluster. The recreational data had a similar situation where redtail parrotfish and striped parrotfish were on a combined total of 4 trips, and on Figure 9 they are shown at the bottom of the dendogram with low separation from the other clusters. In contrast, more clear clusters were formed for

species that were frequently landed on trips. For example, in the commercial data lane snapper and silk snapper were each landed on over 5,000 trips, and in Figure 1 there is a clear cluster of lane snapper and silk snapper. The recreational data had dolphin landed on 421 trips and there is a clear separation of a cluster for dolphin in Figure 9.

Figure 2 had more defined clusters than Figure 1. The data for the cluster analysis shown in Figure 2 had the write-in species removed. The fishermen may not consistently report their landings for the write-in species on the catch form which explains why these species were on a relatively small number of trips.

The isolation of the trips by gear (Figures 3, 4, 5, and 6) and their corresponding cluster analysis made different clusters of species. Interestingly, some species that were frequently landed on trips for one gear were infrequently caught with another gear. Thus, leading to different clusters of species. This same result was found when the data was filtered into the two depth bins.

It's unlikely that the Council will choose species groups for certain gears or depths. However, the cluster analysis by gear and depth allows evaluation of how any species groups that the Council chooses will impact species observed by the different gears and depth.

The dendrogram was pruned using the cluster analysis for all available data (both commercial and recreational data, Figures 10 and 11) to generate a list of clusters. Following the second method, which gave more weight to the commercial sector, Table 1 provides a list of potential species groupings. These groupings are based on the height of the dendrogram branches.

Table 1. Species clusters generated from a cluster analysis of both the commercial and recreational data. The results were generated from the second method which gave more weight to the commercial sector since it had more landings. The table only includes species that had landings in both the commercial and recreational sectors.

Cluster	Species
1	Dolphin
2	Barracuda
3	Silk Snapper Lane Snapper Yellowtail Snapper
4	King Mackerel Wahoo
5	Mutton Snapper Red Hind
6	Misty Grouper Schoolmaster Stoplight Parrotfish Stingray

	Cubera Snapper Rainbow Runner
7	Little Tunny Blackfin Tuna White Grunt Blackfin Snapper Vermilion Snapper
8	Queen Triggerfish Hogfish Queen Snapper

References

Southeast Data Assessment and Review (SEDAR) 46. 2016. U.S. Caribbean Data-Limited Species data workshop report. 278 pp.