Hierarchical cluster analyses of the American Samoa and Guam boat-based creel data.

Robert Ahrens, Marc Nadon, Erin Bohaboy, Felipe Carvalho, Joseph O'Malley, T. Todd Jones

Background

Annual catch limits (ACLs) and accountability measures (AMs) are required for all federally managed stocks (MSRA 2006). Magnuson-Stevens Act requires a council to prepare a Fishery Management Plan (FMP) or Fishery Ecosystem Plan (FEP) for each fishery under its authority that requires conservation and management. Guidelines for determining conservation and management status for a stock are provided in the National Standards (NS) where 10 factors are presented for consideration. When fisheries target multiple species or when data are limited, developing ACLs or AMs for every species captured in a fishery may not be possible. In these situations, National Standard 1 (NS-1) recommends grouping stocks into 'complexes' where stocks have similar geographic distribution, life history characteristics, and vulnerabilities to fishing pressure. When possible, stock complexes should have 'indicator' stocks that are representative of the complex. Measurable and objective status determination criteria (SDC) can be established for the indicator stocks and used to help manage the stock complex.

Within a multispecies fishery, where complexes are required for management, the National Standard guidance implies a process where individuals subject to similar fishing pressure are identified first. These geographically similar assemblages can be further aggregated based on life history characteristics and vulnerability to define a complex and, in turn, indicator species can be identified. Fisheries in the U.S. Pacific territories are multi-gear and multispecies by nature with information extrapolated from limited intercept surveys where individuals may not be identified to species. These fisheries are candidates for establishing species complexes in FMPs or FEPs, and complexes have been developed in previous plans. The MSRA and NS guidelines recommend decisions regarding species complexes and indicator species be revisited to ensure that they are achieving the conservation and management goals. It is in this spirit, as well as to provide a transparent and repeatable process, that this analysis was conducted.

In this paper we present the results of hierarchical clustering of creel interviews for boat-based operations in American Samoa and Guam. The dendrograms are intended to delineate species aggregations that are experiencing similar fishing pressure to facilitate, when used in conjunction with life history information, the determination of species complexes for FMPs and FEPs.

Methods

To identify species assemblages that are likely subject to similar fishing pressure, hierarchical clustering was applied to the American Samoan boat-based intercept creel data (Pacific Islands Fisheries Science Center, 2021a) and the Guam boat-based intercept creel data (Pacific Islands Fisheries Science Center, 2021b). Working under the premise that species that are captured together, are likely subject to similar fishing effort, and can be viewed as occurring in a fishery, each data set was converted to presence and absence by interview. Clustering was performed

PIFSC Internal Report IR-22-01
Issued 7 January 2022

in R (R Core Team, 2020). The *dist* function was used to create a binary dissimilarity matrix. A tree structure was produced through agglomerative hierarchical clustering using the ward.D2 method (Ward Jr, 1963) and the *hclust* function. Data were analyzed in their entirety for each creel program as well as in specific time blocks. The ceiling of the 5% quantile for positive total occurrence of a species in the full records was used as a cutoff. Comparisons between dendrograms of different time blocks was conducted using the *tanglegram* function in the dendextend library.

Cluster significance was estimated using a permutation test where interview specific species occurrences were randomly permuted to generate novel random data sets of presence/absence of the same size as the original data sets. Clustering was then performed on these data sets. The number of nodes within a tree is equal to one less than the number of categories. A branch height was calculated for each node. A *p*-value was calculated as the proportion of branch heights—original and simulated data—that were larger or equal to the height at the same node in the original data. The smallest *p*-value that can be obtained is the reciprocal of the number of simulated data sets plus 1. This permutation test identifies a node height at which (and above) the grouped nodes could be considered not random clusters.

For American Samoa, boat-based creel data from 1986–2019 were included in the analysis. Three blocks of data were assessed: 1986–2019, 2010–2019, and 2016–2019. These blocks represent noticeable changes in the creel program. In particular, after 2016, improvements were made to species level identification. The boat-based data from Guam were assessed as: all years (1982–2020) and post-2000. For both American Samoa and Guam, analyses were also conducted for the same time blocks on a subset of the data where trips were declared as bottom fishing.

Findings

From a federal management perspective, the main objective of this analysis was to more fully understand the fishing pressure exerted on bottomfish species and the potential complexes that could be developed. Although the majority of more commonly caught species are included in the analysis, species currently in BMUS (**Table 1,** **Table 2**) will be the focus. Two plot types are presented, dendrograms of species association based on interview and density plots for each species indicating the distribution for positive records of the gears reported on each interview. When interpreting the dendrograms, it is important to remember that the relative location of individuals along the vertical axis does not necessarily indicate greater dissimilarity. Greater distances along the height axis do suggest greater dissimilarity. The density plots are useful for understanding the diversity of gears in which a species may be caught as well as understanding the label colors associated with each leaf on the dendrogram.

American Samoa Boat Based

Species associations are generally determined by the main gear of capture with some over distribution as a result of mixed gear categorization (**Figure 1,** **Figure 3,** **Figure 5**). Pelagic species cluster together consistently across time blocks; tunas cluster separately from billfishes as a result of the proportion of observations occurring in troll gear (**Figure 2, Figure 4, Figure 6**). Shallow reef species primarily captured using spears also group together with species clustering in a consistent manner over time. Bottom associated species tend to cluster based on depth preference and gear, bottomfish mainly occur in deeper depths and require heavier gear cluster apart from more moderate depth associated species that are captured on lighter tackle or spear. These associations are not necessarily consistent over time given the relative influence of mixed gear trips over time.

For species currently listed as bottomfish MUSs (**Table 1**), there are noticeable patterns of association that are time period dependent. The strongest association between these species is observed when all data are assessed (**Figure 5**). Moderate depth species (*Aprion virescens, Lethrinus rubrioperculatus,* and *Lutjanus kasmira*) cluster from the deeper (*Aphareus rutilans, Caranx lugubris, Etelis carbunculus, Etelis coruscans, Pristipomoides flavipinnis, Pristipomoides zonatus, Pristipomoides filamentosus, Variola louti*). This pattern changes as older data are excluded from analyses. The deep species cluster is somewhat consistent, but *Pristipomoides flavipinnis, Pristipomoides filamentosus,* and *Variola louti* become less associated (**Figure 3, Figure 5**) and become grouped with more moderate depth species due potentially to a higher proportion of interviews reporting the mixed bottomfish / trolling categories (**Figure 4, Figure 6**). *Lutjanus gibbus*, *Sphyraena forsteri*, and *Variola albimarginata* are frequently associated with some of the more moderate depth BMUS species. These patterns are also apparent when only bottomfish trips are analyzed (**Figure 7, Figure 8, Figure 9**).

Guam Boat Based

The Guam boat-based interviews cluster into patterns that are similar to those from American Samoa. For the pelagic species, a cluster can be seen between offshore and nearshore across time blocks (**Figure** 10**,** Figure **12**) as a result of gear type; trolling is dominant in offshore, and other methods influence nearshore associations (**Figure 11, Figure 13**). A cluster of shallow water reef species is evident when all time periods are analyzed, but the associations are less clear in the later time periods with smaller sample sizes. A deepwater bottomfish cluster and more moderate depth species clusters are apparent and consistent over time with greater consistency in the deepwater cluster.

For species currently listed as bottomfish MUSs (**Table 2**), there is a consistent cluster of deepwater species across time blocks (*Aphareus rutilans, Caranx lugubris, Etelis carbunculus, Etelis coruscans, Pristipomoides auricilla, Pristipomoides flavipinnis, Pristipomoides zonatus, Pristipomoides filamentosus)* though the association with *Pristipomoides seiboldii* is less apparent using more recent data. *Caranx ignobilis* appear unassociated with any of the BMUS falling into a mixed grouping for reef and midwaters species*. Lethrinus rubrioperculatus, Lutjanus kasmira*, and *Variola louti* tend to cluster together with more intermediate depth species. Notable associations with the deepwater cluster are *Seriola dumerili* and *Pristipomoides argyrogrammicus*. These patterns are also apparent when only bottomfish trips are analyzed, except *Pristipomoides seiboldii* occur in insufficient numbers to be included (Figure 14 and Figure 15).

Summary

One of the objectives of defining species complexes is to ensure that species in the complex are subject to similar fishing pressure. If clustering of species from interview data is indicative of the fishing pressure, then it is likely that current BMUS are experiencing different fishing intensities and a redefinition of complexes may be warranted. Unfortunately, no distinction has been made between federal vs. territorial waters because the spatial resolution of interview records is insufficient to allow for such a categorization of a trip. Presence in federal vs. territorial waters could be a criterion for determination of inclusion in BMUS list but would need to be considered outside of this analysis. In both the American Samoa and Guam data, a deepwater cluster comprised of *Etelis* and some *Pristipomoides* as well as *Caranx lugubris* and *Aphareus rutilans* is apparent. More moderate depth species tend to be associated with a broader suite of species and may warrant a separate complex. Finally, species that have been depleted historically may not appear at high levels in the creel data and would not be included in the dendrograms.

Table 1. Bottomfish management unit species (BMUS) that are identified in the relevant Fishery Ecosystem Plan and that are used for the bottomfish assessment for American Samoa.

Species name	Common name	Samoan name
Aphareus rutilans	Rusty jobfish	palu-gutusiliva
Aprion virescens	Green jobfish	asoama
Caranx lugubris	Black trevally	tafauli
Etelis carbunculus[1]	Ruby snapper	palu malau
Etelis coruscans	Flame snapper	palu-loa
Lethrinus rubrioperculatus	Redgill emperor	filoa-paomumu
Lutjanus kasmira	Bluestripe snapper	savane
Pristipomoides flavipinnis	Yelloweye snapper	palu-sina
Pristipomoides zonatus	Oblique-banded snapper	palu-ula, palu-sega
Pristipomoides filamentosus	Pink snapper	palu-ʻenaʻena
Variola louti	Lyretail grouper	papa, velo

Table 2. Bottomfish management unit species (BMUS) that are identified in the relevant Fishery Ecosystem Plan and that are used for the bottomfish assessments for Guam and the Commonwealth of the Northern Mariana Islands.

Species name	**Common name**	**Chamorro/Carolinian name**
Aphareus rutilans	Rusty jobfish	lehi/maroobw
Caranx ignobilis	Giant trevally	mamulan/etam
Caranx lugubris	Black trevally	tarakiton attelong/orong
Etelis carbunculus	Ruby snapper	buninas agaga'/falaghal moroobw
Etelis coruscans	Flame snapper	buninas/taighulupegh
Lethrinus rubrioperculatus	Redgill emperor	mafute'/atigh
Lutjanus kasmira	Bluestripe snapper	funai/saas
Pristipomoides flavipinnis	Yelloweye snapper	buninas/falaghal-maroobw
Pristipomoides sieboldii	Von Siebold's snapper	buninas/-
Pristipomoides zonatus	Oblique-banded snapper	buninas rayao amiriyu /falaghal-maroobw
Pristipomoides auricilla	Goldflag snapper	buninas/falaghal-maroobw
Pristipomoides filamentosus	Pink snapper	buninas/falaghal-maroobw
Variola louti	Lyretail grouper	gadau matingon/bwele

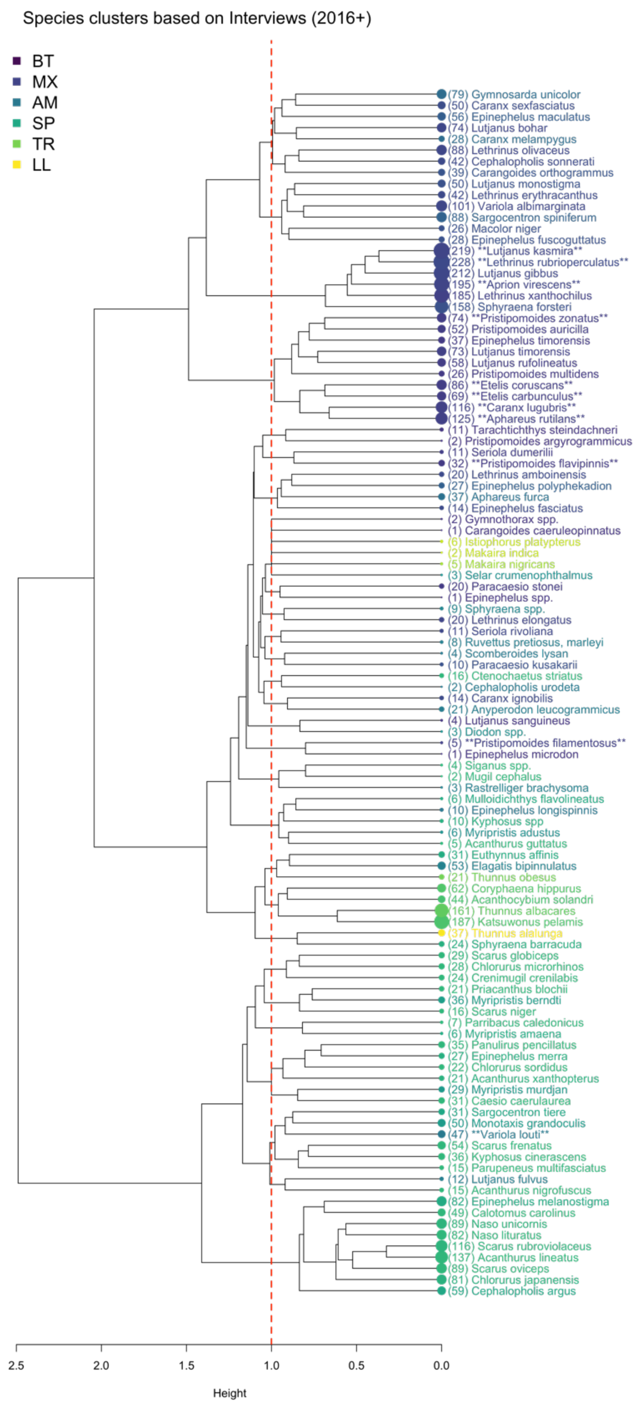

Figure 1. Dendrogram for American Samoa boat-based creel data from 2016–2019. Color of the species name indicates the main gear of capture: bottom gear (BT), mixed (MX), Atule mixed (AM), spear (SP), troll (TR), longline (LL). The leaf dot size indicates the number of positive records. Nodes with heights greater than the red vertical line can be used to determine clusters that have a low probability of being random.

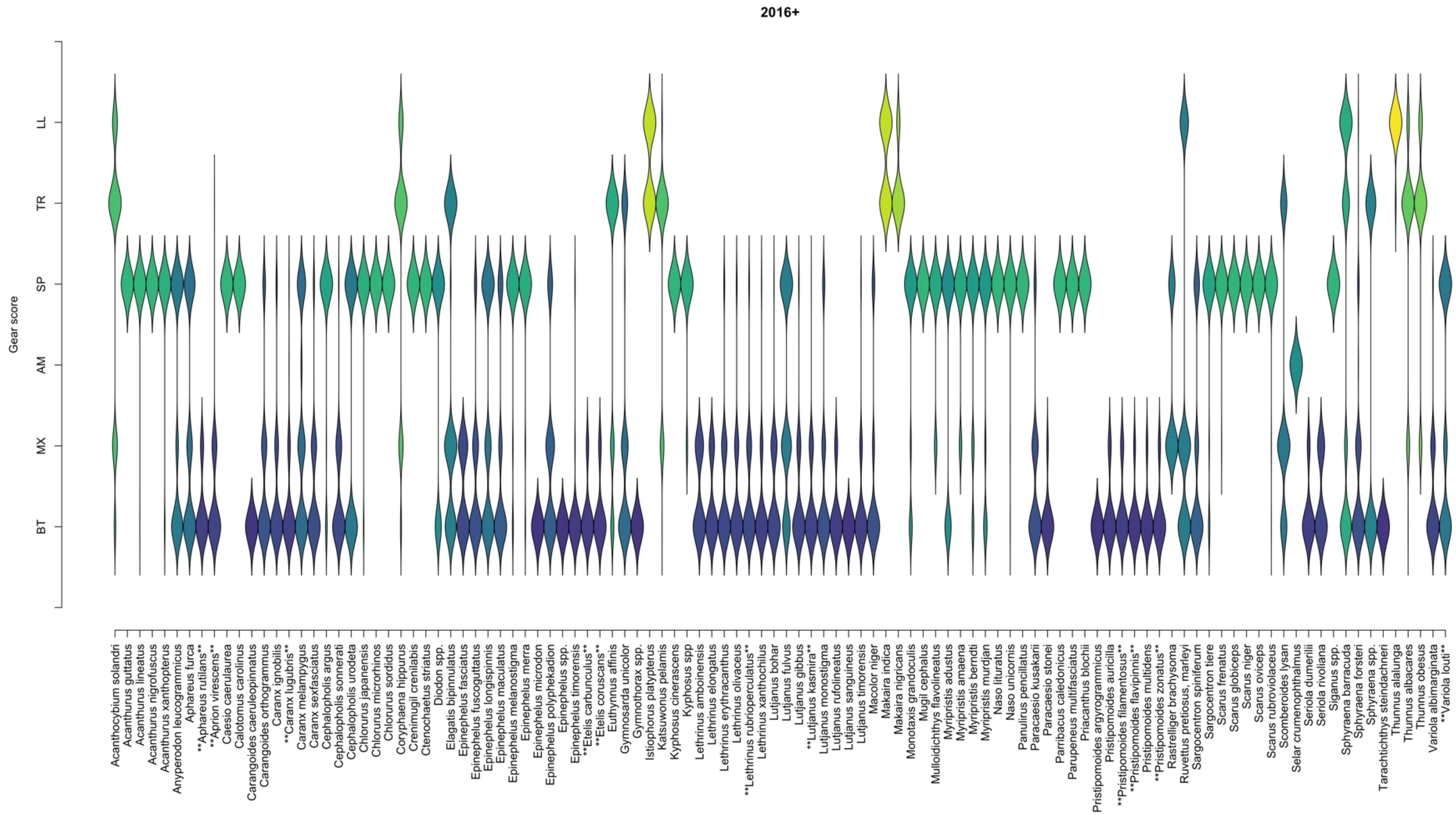

Figure 2. Violin plot of the density of species records by gear type for American Samoa boat-based creel data from 2016–2019. Color of the violin indicates the average gear score: bottom gear (BT), mixed (MX), Atule mixed (AM), spear (SP), troll (TR),longline (LL).

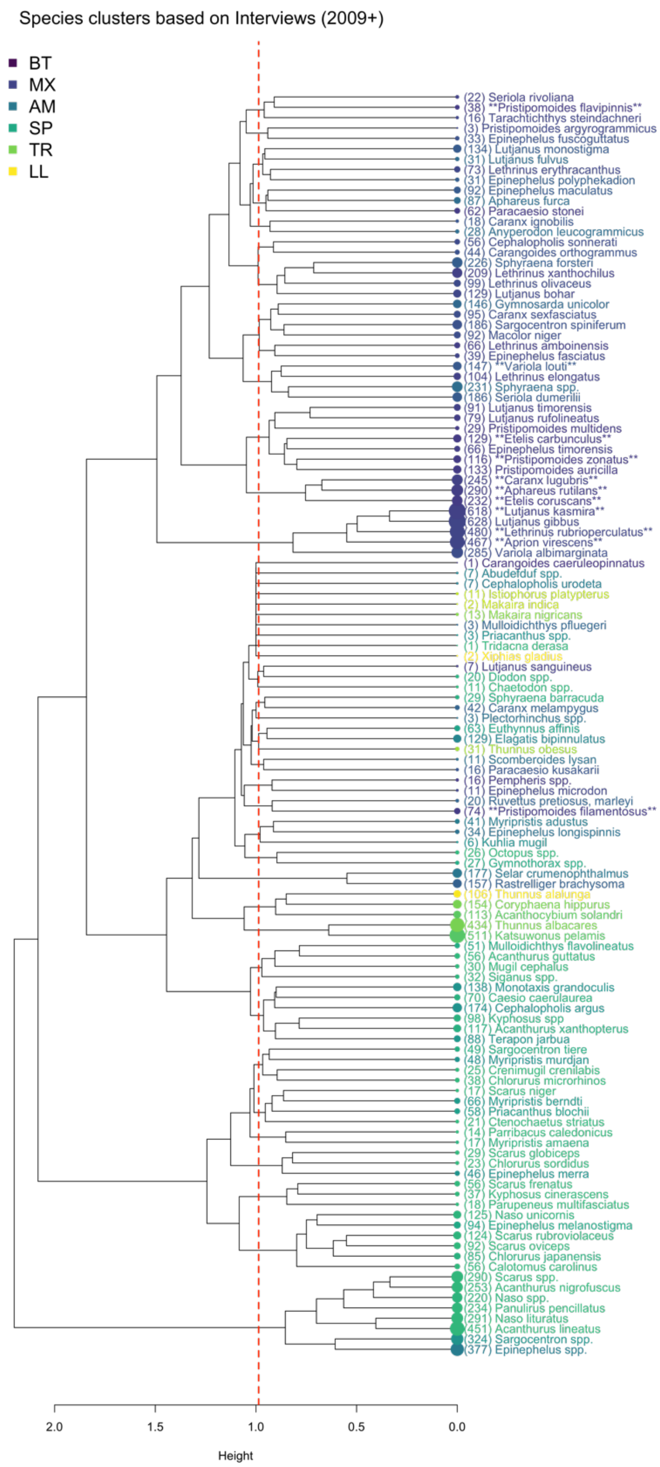

Figure 3. Dendrogram for American Samoa boat-based creel data from 2009–2019. Color of the species name indicates the main gear of capture: bottom gear (BT), mixed (MX), Atule mixed (AM), spear (SP), troll (TR), longline (LL). The leaf dot size indicates the number of positive records. Nodes with heights greater than the red vertical line can be used to determine clusters that have a low probability of being random.

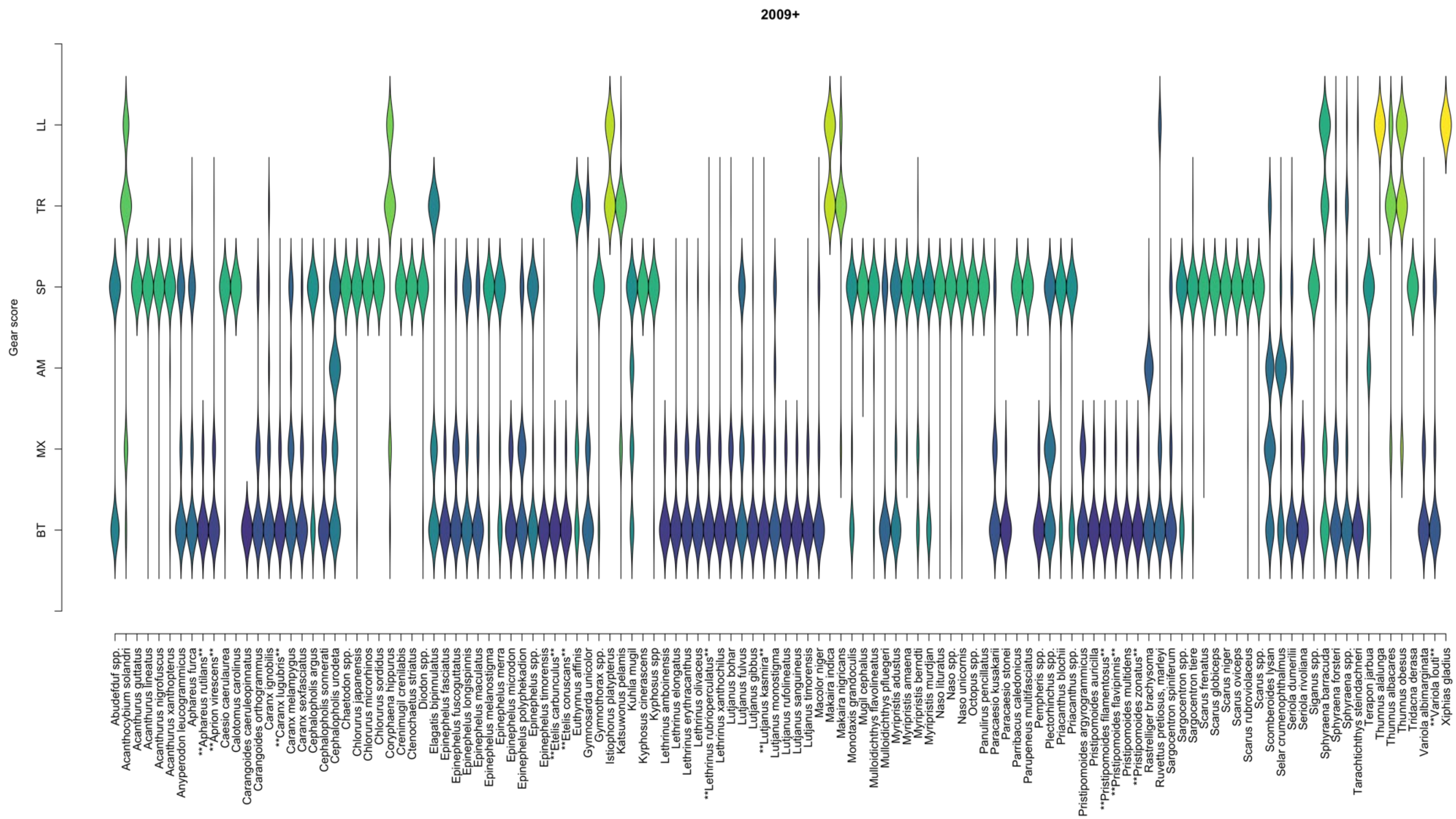

Figure 4. Violin plot of the density of species records by gear type for American Samoa boat-based creel data from 2009–2019. Color of the violin indicates the average gear score: bottom gear (BT), mixed (MX), Atule mixed (AM), spear (SP), troll (TR), longline (LL).

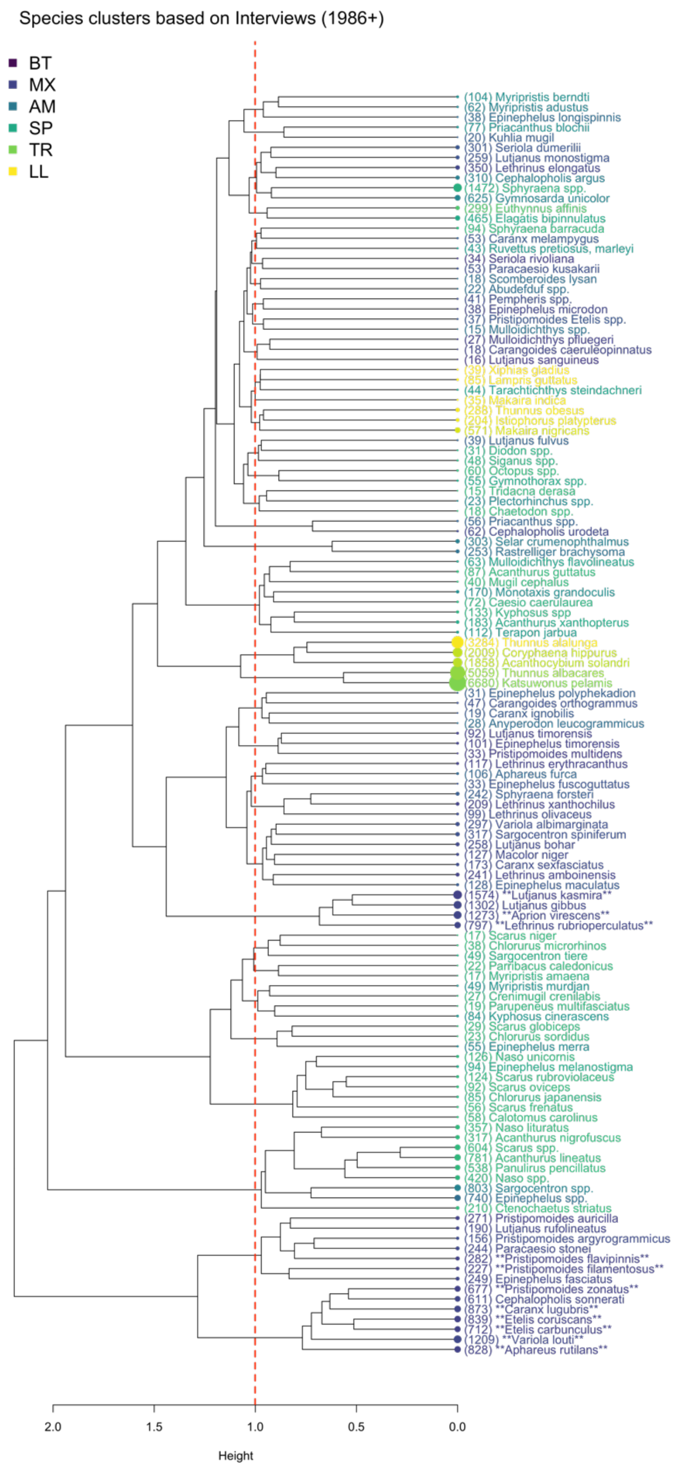

Figure 5. Dendrogram for American Samoa boat-based creel data from 1986–2019. Color of the species name indicates the main gear of capture: spear (SP), bottom gear (BT), mixed (MX), Atule mixed (AM), troll (TR), longline (LL). The leaf dot size indicates the number of positive records. Nodes with heights greater than the red vertical line can be used to determine clusters that have a low probability of being random.

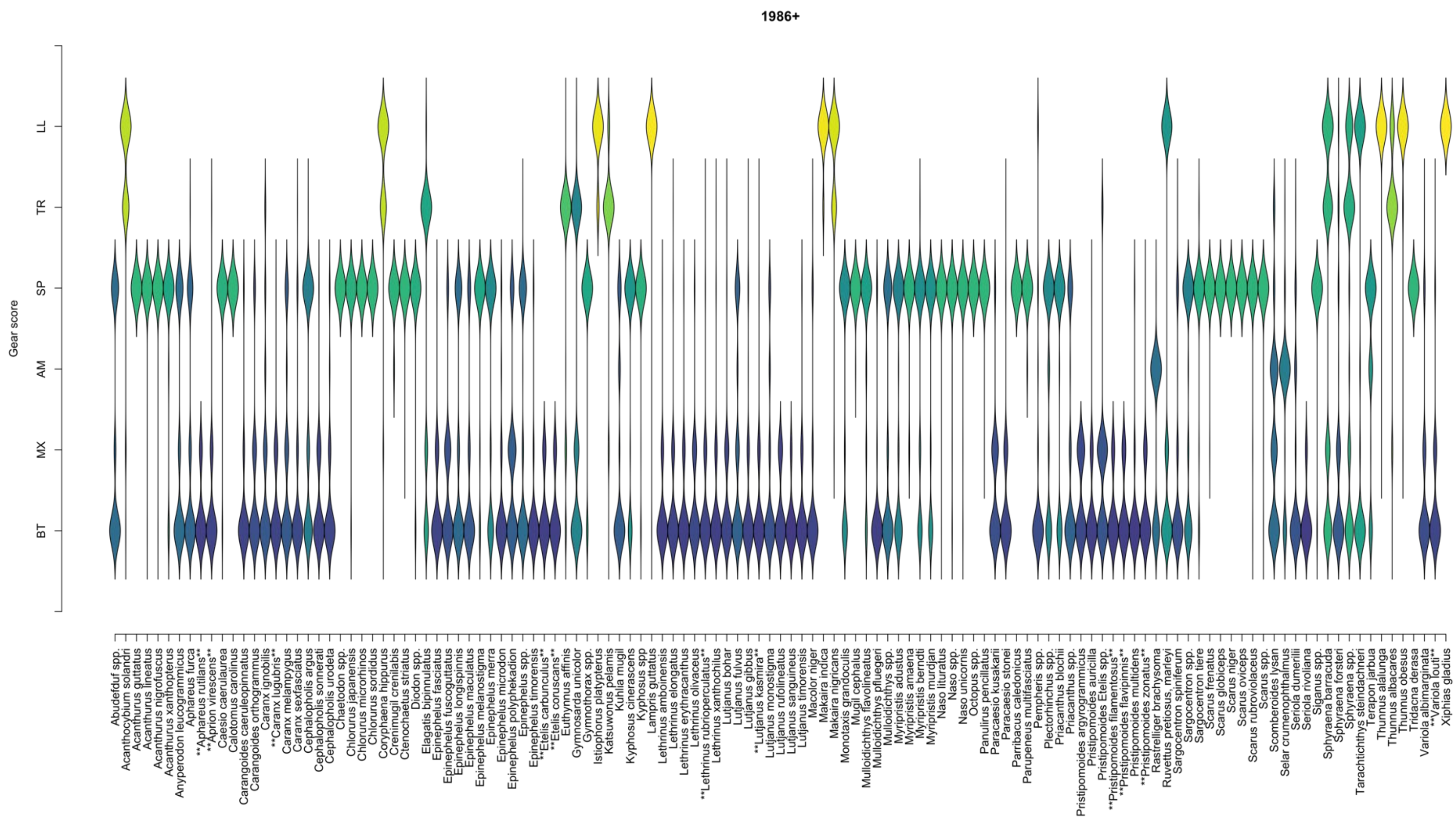

Figure 6. Violin plot of the density of species records by gear type for American Samoa boat-based creel data from 1986–2019. Color of the violin indicates the average gear score: bottom gear (BT), mixed (MX), Atule mixed (AM), spear (SP), troll (TR),longline (LL).

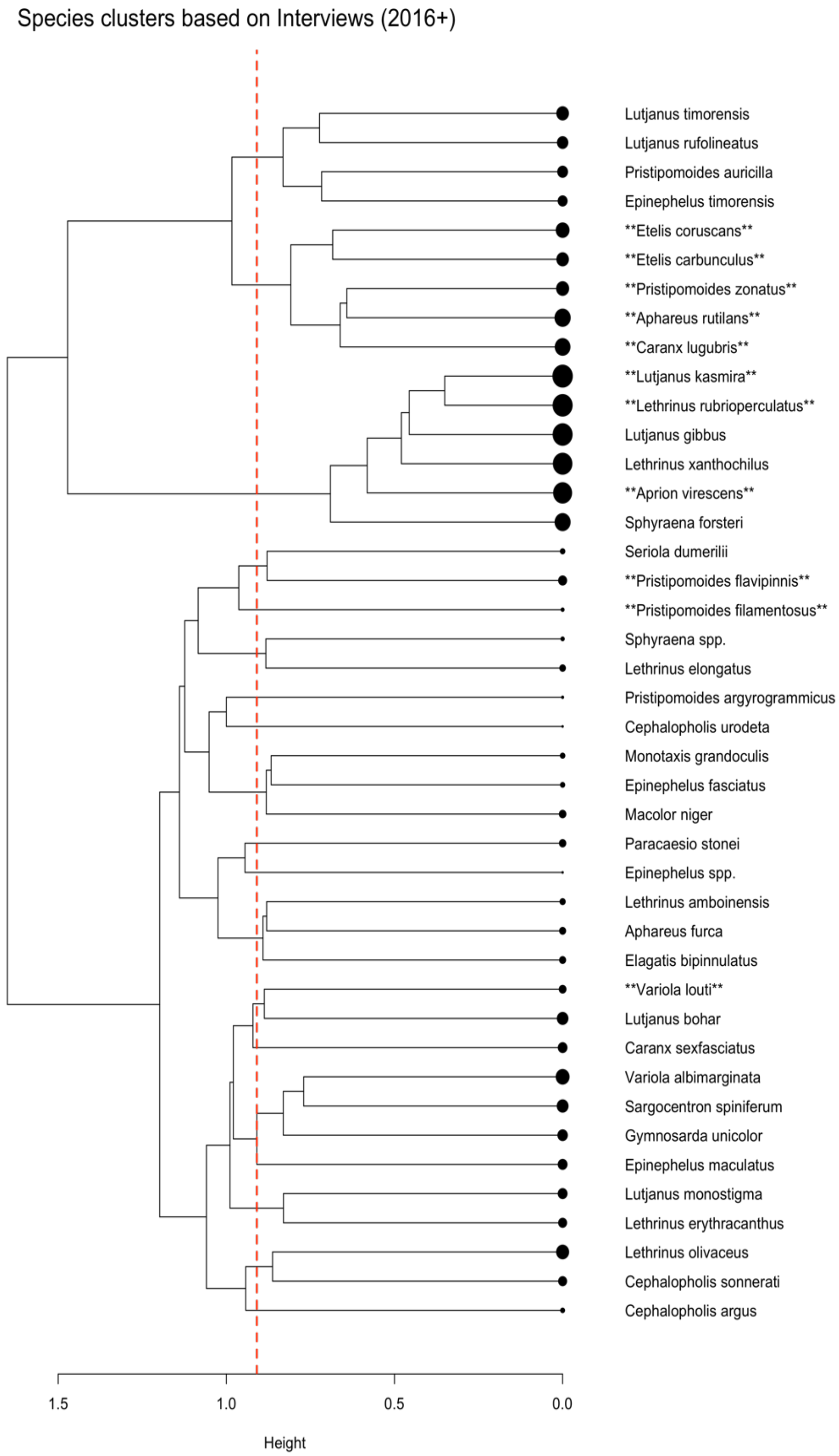

Figure 7. Dendrogram for American Samoa boat-based creel data from 2016–2019 for bottom fishing declared trips only. The leaf dot size indicates the number of positive records. Nodes with heights greater than the red vertical line can be used to determine clusters that have a low probability of being random.

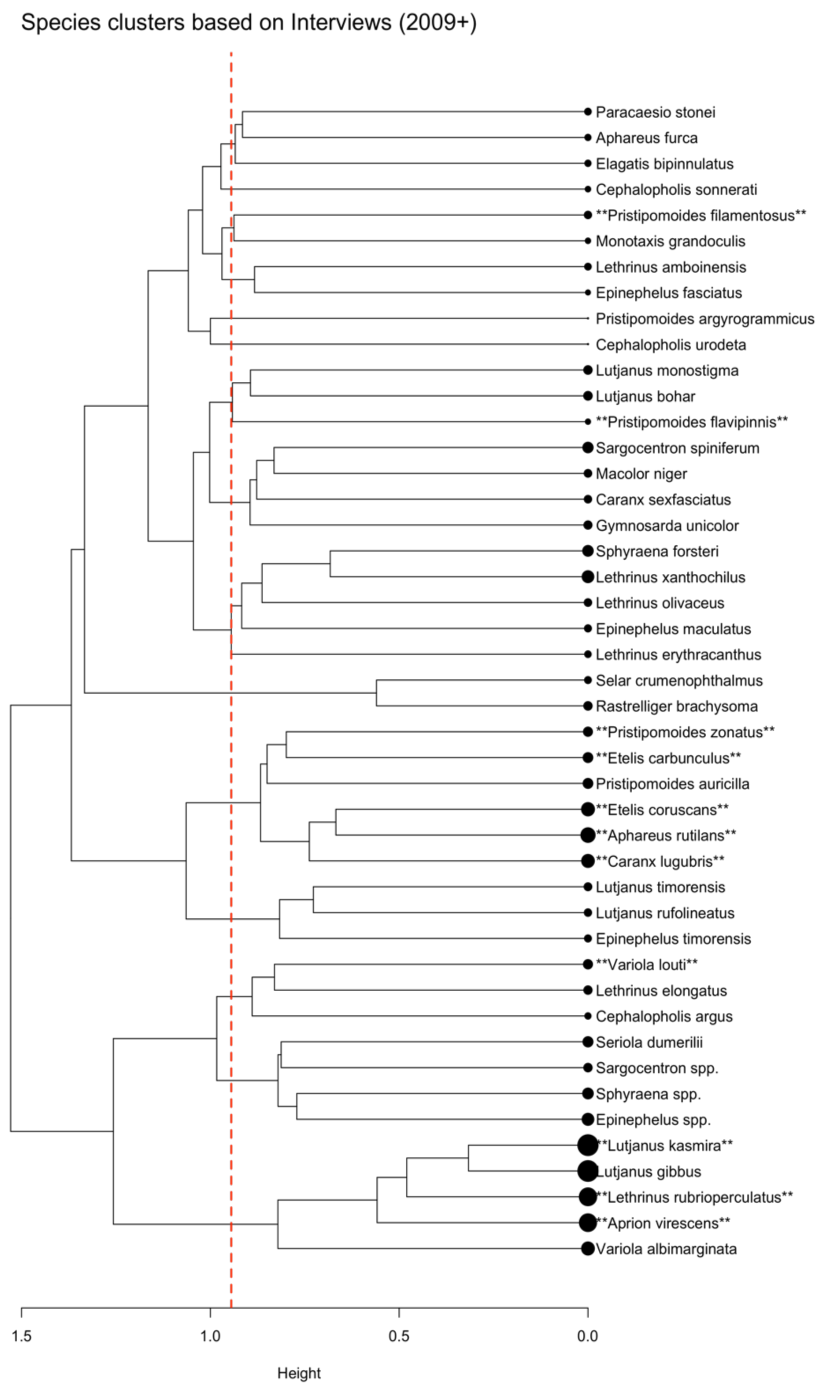

Figure 8. Dendrogram for American Samoa boat-based creel data from 2009–2019 for bottom fishing declared trips only. The leaf dot size indicates the number of positive records. Nodes with heights greater than the red vertical line can be used to determine clusters that have a low probability of being random.

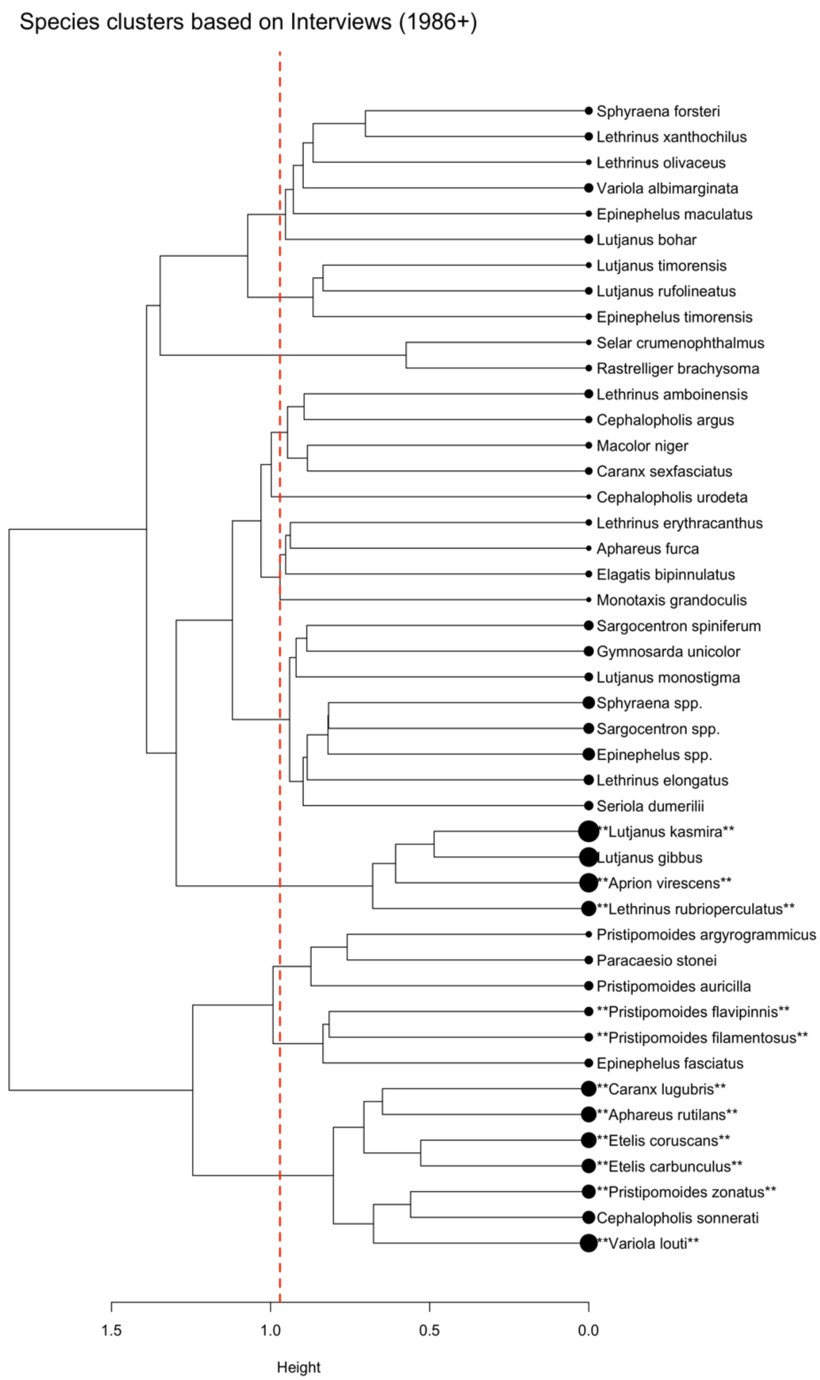

Figure 9. Dendrogram for American Samoa boat-based creel data from 1986–2019 for bottom fishing declared trips only. The leaf dot size indicates the number of positive records. Nodes with heights greater than the red vertical line can be used to determine clusters that have a low probability of being random.

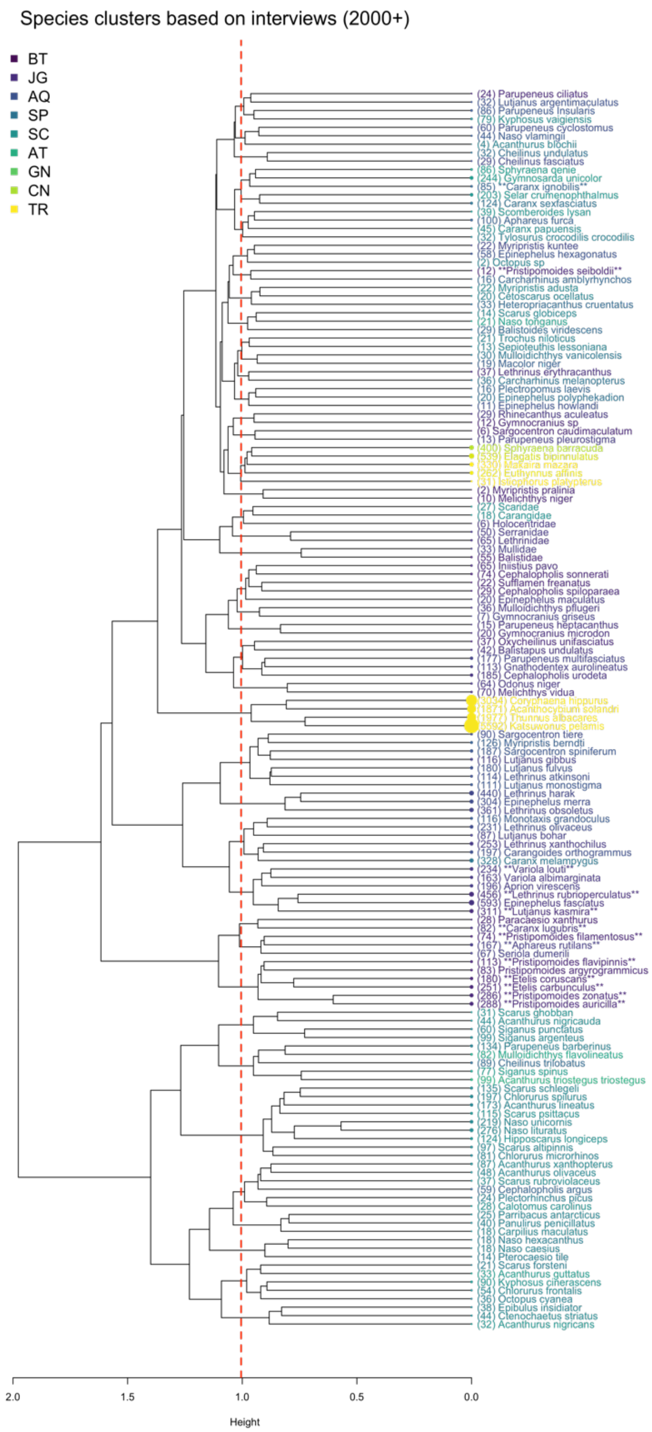

Figure 10. Dendrogram for Guam boat-based creel data from 2000–2020. Color of the species name indicates the main gear of capture: bottom gear (BT), jigging (JG), aquarium trade collection (AQ), spear (SP), spin casting (SC), Atulai (AT), gillnet (GN), cast net (CN), troll (TR). The leaf dot size indicates the number of positive records. Nodes with heights greater than the red vertical line can be used to determine clusters that have a probability of being random.

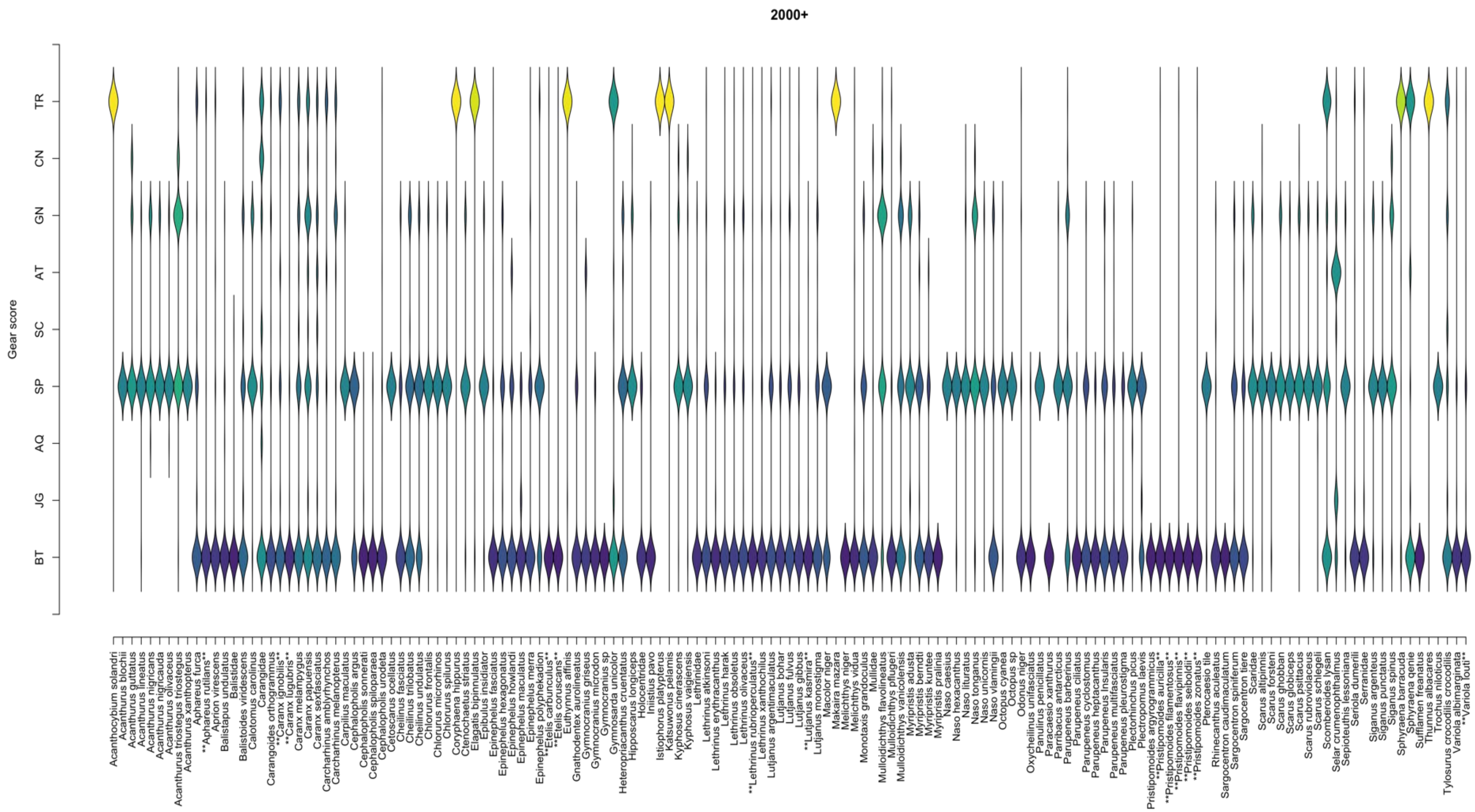

Figure 11. Violin plot of the density of species records by gear type for Guam boat-based creel data from 2000–2020. Color of the violin indicates the average gear score: bottom gear (BT), jigging (JG), aquarium trade collection (AQ), spear (SP), spin casting (SC), Atulai (AT), cast net (CN), troll (TR).

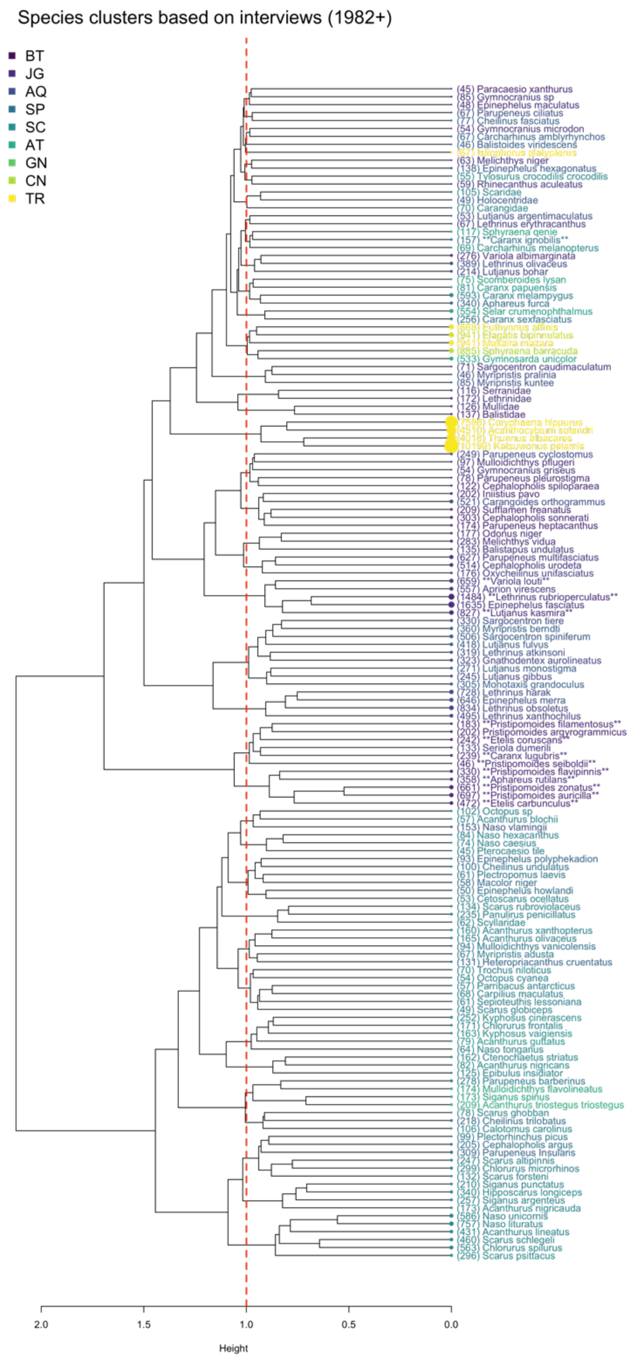

Figure 12. Dendrogram for Guam boat-based creel data from 1982–2020. Color of the species name indicates the main gear of capture: bottom gear (BT), jigging (JG), aquarium trade collection (AQ), spear (SP), spin casting (SC), Atulai (AT), gillnet (GN), cast net (CN), troll (TR). The leaf dot size indicates the number of positive records. Nodes with heights greater than the red vertical line can be used to determine clusters that have a low probability of being random.

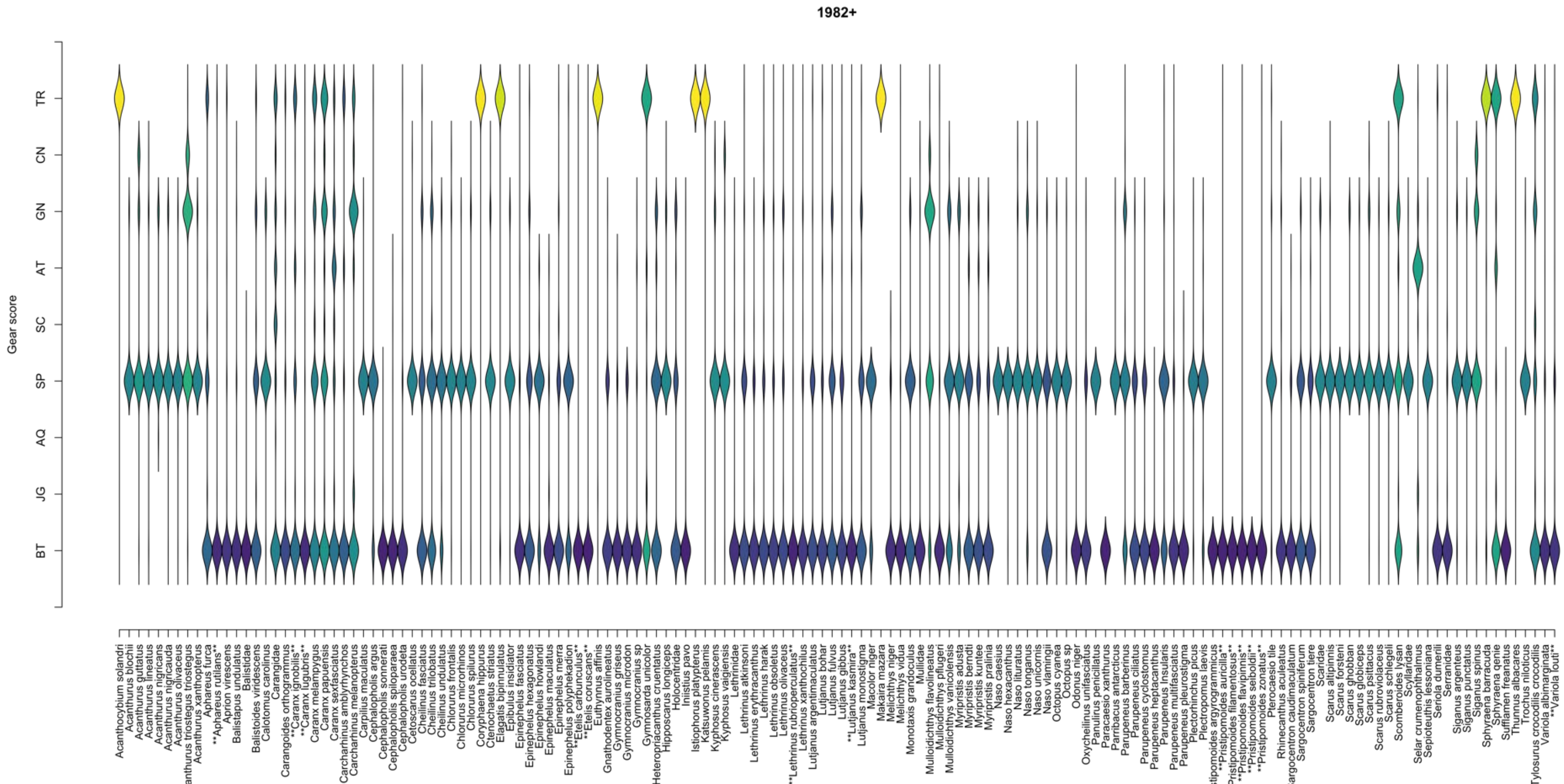

Figure 13. Violin plot of the density of species records by gear type for Guam boat-based creel data from 1982–2020. Color of the violin indicates the average gear score: bottom gear (BT), jigging (JG), aquarium trade collection (AQ), spear (SP), spin casting (SC), Atulai (AT), gillnet (GN), troll (TR).

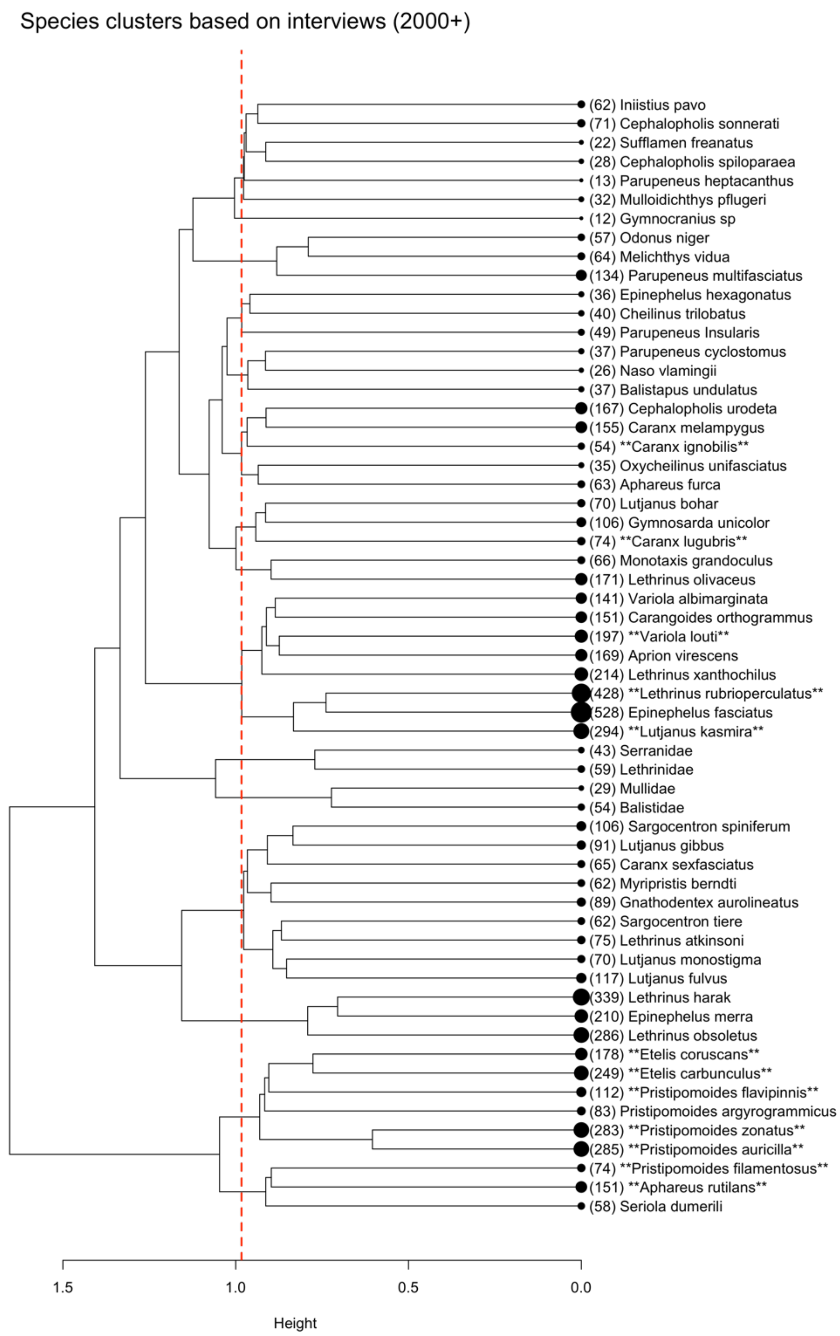

Figure 14. Dendrogram for Guam boat-based creel data from 2000–2020 for bottom fishing declared trips only. The leaf dot size indicates the number of positive records. Nodes with heights greater than the red vertical line can be used to determine clusters that have a low probability of being random.

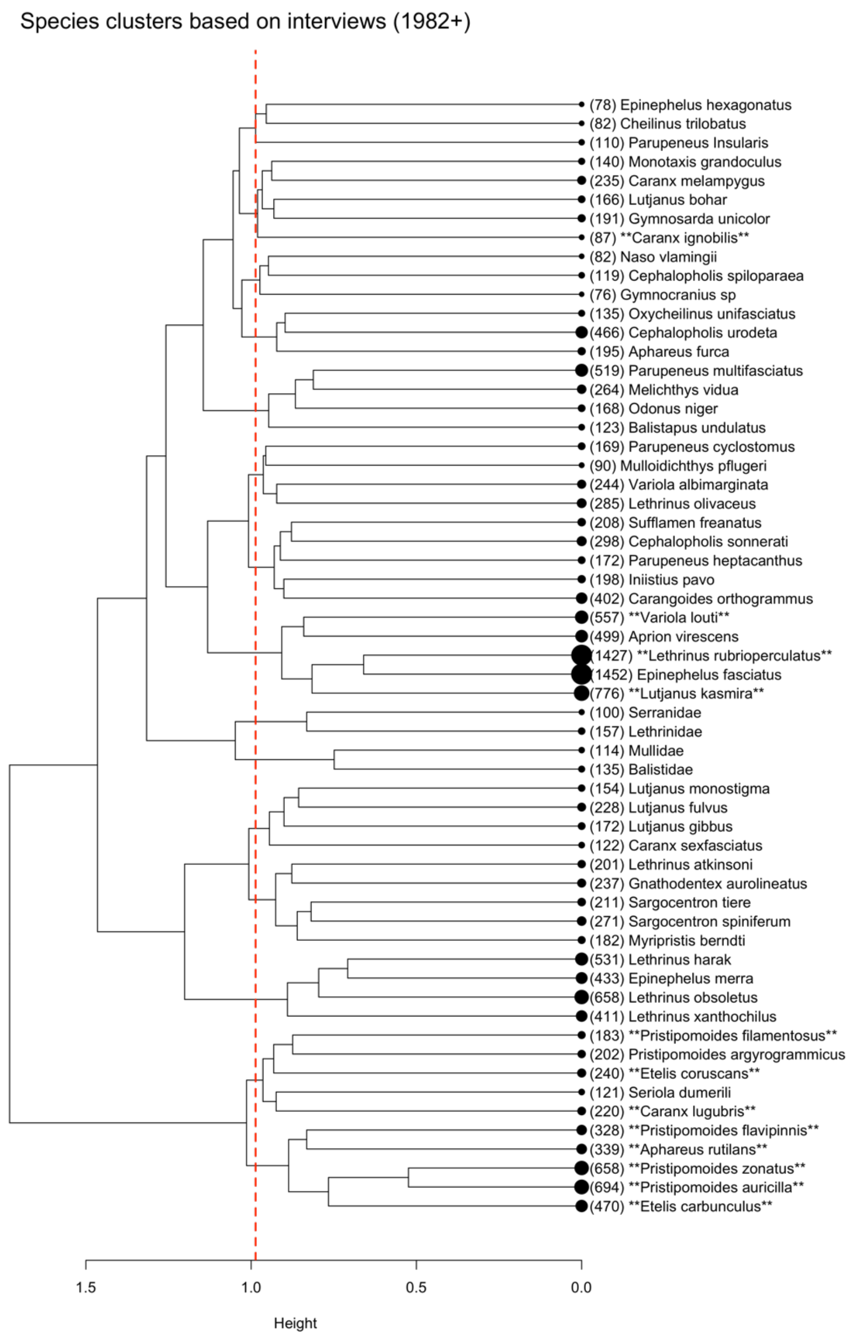

Figure 15. Dendrogram for Guam boat-based creel data from 1982–2020 for bottom fishing declared trips only. The leaf dot size indicates the number of positive records. Nodes with heights greater than the red vertical line can be used to determine clusters that have a low probability of being random.

References

Pacific Islands Fisheries Science Center. 2021a. American Samoa Boat-based Creel Survey. NOAA National Centers for Environmental Information, https://www.fisheries.noaa.gov/inport/item/5612.

Pacific Islands Fisheries Science Center. 2021b. Guam Boat-based Creel Survey. NOAA National Centers for Environmental Information, https://www.fisheries.noaa.gov/inport/item/5620.

R Core Team. 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.R-project.org/

Ward Jr JH. 1963. Hierarchical grouping to optimize an objective function. J Am Stat Assoc.58(301): 236–244.