

Catch Length Composition of Bottomfish Management Unit Species of Guam, 1982–2023

DRAFT for WPSAR CIE Review July 8–12, 2024

Do not distribute

Toby Matthews and Erin C. Bohaboy

Pacific Islands Fisheries Science Center National Marine Fisheries Service 1845 Wasp Boulevard Honolulu, HI 96818

EXECUTIVE SUMMARY

Two types of length composition were presented for the Guam Bottomfish Management Unit Species: 1) Recent length compositions meant to estimate the present-day size structure, generated using data from the Commercial Fisheries Biosampling Program; and 2) Time-series length composition meant to capture changes in size structure over time, generated using data from the Guam Department of Agriculture, Division of Aquatic and Wildlife Resources boat-based creel survey. Length data from each source were filtered to exclude lengths from the non-bottomfishing gears (with one exception) and lengths from incomplete catch records, when bias was detected in incomplete catches. When sufficient length data were available, lengths were weighed by area to account for unbalanced sampling of nearshore and banks regions around Guam. Larger fish of most species were caught at the banks than nearshore, causing the region-weighted and unweighted length compositions to deviate. Ultimately, usable recent length compositions were produced for 12 of the 13 species, while time-series could only be produced for 9 species due to data limitations.

INTRODUCTION

The Bottomfish Management Unit Species (BMUS) of Guam include 13 species of snappers, jacks, and a grouper that are managed in Federal waters by the Western Pacific Regional Fishery Management Council under the Fishery Ecosystem Plan (FEP) for the Mariana Archipelago (FEP; WPRFMC 2009). This report is one of four documents prepared ahead of an external review, to be conducted in July 2024 as part of the Western Pacific Stock Assessment Review (WPSAR), to present data that will be used in benchmark stock assessments of Guam BMUS. Previous stock assessments of the BMUS have been conducted on the entire multi-species complex, most recently in the 2019 benchmark stock assessment (Langseth et al. 2019), which was updated in 2024 (Bohaboy and Matthews, 2024). For the upcoming BMUS benchmark assessment, single-species assessments may be considered. As such, this report describes the length data available for each individual BMUS, and is accompanied by reports on species-specific catch, catch per unit effort (CPUE), and life history data.

The length composition data presented in this working paper can be classified as one of two overall types: recent and time-series. Recent length compositions provide a present-day snapshot of the catch, and include only the most recent years of available length data. Recent length compositions are derived from the Commercial Fisheries Biosampling Program (hereafter "biosampling program"), which provides 400 to 5543 length observations per year (aggregated for all BMUS) but is only available dating back to late 2009. Time-series length compositions provide information on how the catch

lengths of BMUS may have changed over the past several decades. Time-series length compositions are based on the Guam Department of Agriculture, Division of Aquatic and Wildlife Resources (DAWR) boat-based creel survey (BBS) interview data, which began in 1982. The biosampling program and BBS are each characterized by variability in sampling effort and information quality, and limited sample size. We account for this variability and limited sample size by presenting the appropriate recent and time-series data products for each BMUS, which depends primarily on whether incomplete (or sub-sampled) length observations are used and how fishing location is represented in length compositions.

METHODS

Data Sources

Commercial Fisheries Biosampling Program

The biosampling program was established in 2009 through cooperation between the Pacific Islands Fisheries Science Center (PIFSC) and local staff on Guam (Sundberg et al., 2015). The goal of the biosampling program is to support the collection of length data, weight data, and life history samples for a wide range of fishery species. Biosampling program staff establish cooperative relationships with local fish markets, fishermen, and vendors to gain access to their fish for data collection and to record trip-level information. Until March 2020, the Guam Fishermen's Co-op ("Co-op") was the primary participant. The biosampling program has two components: a "field" component that collects basic information from entire trip catches, and a "lab" component that collects more detailed information on and attains life history samples from specially selected fish.

For the field component, all fish provided from a fishing trip are typically identified to the species level. Length measurements are usually taken for all fish caught in a fishing trip and constitute a complete sample of the catch. However, in some instances, only a subsample of the fish are measured. Whether the biosampling record included the complete or incomplete catch has only been recorded since November, 2020. Basic trip-level information is also collected, such as the fishing methods (e.g., bottomfishing, hook-and-line, spearfishing, etc.) and region fished. The region fished may be detailed to specific named locations, e.g., 'Galvez bank', or may be ambiguous such as the general direction relative to the Island of Guam, e.g., 'South', 'Southwest', etc.

Additional information, including biological samples (e.g., otoliths, gonads, etc.), are collected for the lab component on a subset of fish encountered in the field component.

In addition to the catches recorded in the field component, opportunistically-collected fish, such as from markets, may also be included. Biosampling lab component fish are systematically selected based on the data needs for life history research, hence they often do not represent the lengths of fish caught in the fishery overall. Therefore, the analyses of biosampling program data presented in this working paper rely on only the field component.

Boat-based creel survey

The Guam Department of Agriculture. Division of Aguatic and Wildlife Resources (DAWR) has conducted its boat-based creel survey (BBS) since 1982 following a stratified design to estimate total catch from boat-based fishing across Guam (Jasper et al. 2016). Interviews are collected during the BBS on 8 survey days per month, which are randomly assigned on 4 weekdays and 4 weekend days. Each survey takes place at one of three main fishing ports, with the most active port, Agana Boat Basin, surveyed at twice the frequency of Agat Harbor and Merizo Pier. During each survey day there are two shifts: (1) a morning shift from approximately 0500-1200, and (2) an evening shift from approximately 1600-2400. All fishers returning during a shift are asked if they will participate in a voluntary interview. During the interview, surveyors collect trip-level information including the fishing method used (e.g., bottomfishing, trolling, spearfishing, etc.) and the location fished. Location is recorded using the DAWR BBS offshore location codes, which range in detail from specific banks (e.g., 11 Mile Bank, Area 14 offshore of Agana, etc.) to relatively undefined fishing locations such as quadrants (e.g., "Southwest") and cardinal directions from Guam (e.g., "North"). We grouped the offshore location codes into 5 larger regions: the east/northeast banks (45 Degree and Rota), the south/southwestern banks (11 Mile, Galvez, Baby, Santa Rosa, and White Tuna), the eastern side of Guam (offshore location codes 31, 32, 50-52), the northwestern side of Guam (offshore location codes 10–16), and the southwestern side of Guam (offshore location codes 69, 71–73) (Figure 1; Bohaboy and Matthews, in review, "Standardized Catch per Unit Effort Indices for Bottomfish Management Unit Species of Guam, 1982-2023").

For each BBS interview, surveyors estimate the number of fish of each species in the catch and obtain individual fish length measurements. Generally, only up to 3 individuals of each species are measured (Jasper et al. 2016). As a result, for any given BMUS, a BBS interview may represent the complete catch if every individual was measured, or an incomplete catch if fewer length observations than the number of fish caught are recorded in the interview data.



Figure 1. Guam Department of Agriculture and Wildlife Resources (DAWR) Boat-based Creel Survey offshore location codes, grouped into 5 larger areas: the east/northeast banks (E_banks), the northwestern quadrant of Guam (NW), the southwestern nearshore areas (SW_nearshore), the south/southwestern banks (SW_banks), and the eastern nearshore areas (E_nearshore).

Length Composition Dataset Filtering

Recent Length Compositions from the Biosampling Program

We downloaded all biosampling program field data records for August 2009 – December 2023 from the Guam SQL-server Datawarehouse curated by the Western Pacific Fisheries Information Network (WPacFIN) on 12 June, 2024.

We followed a sequence of analyses and decisions to produce the filtered biosampling dataset for each BMUS (Figure 2a. Filter Biosampling Data flowchart). First, we filtered the dataset to retain only the lengths from the most common gear type for each BMUS. Specific fishing methods were grouped into three broad gear types: bottomfishing, line fishing, and spearfishing. Component fishing methods for line fishing (e.g. trolling, hook and line) and spearfishing (e.g. snorkel spear, scuba spear) were assigned logically, while bottomfishing was only represented by a single fishing method. For bottomfishing, length measurements were also included on a species-by-species basis from fishing trips that engaged in multiple fishing methods when the non-bottomfishing methods employed on the trip had not typically caught the species. The specific rules applied to identify bottomfishing lengths were:

- Length measurements from trips that only engaged in bottomfishing were included for all BMUS.
- Length measurements from trips that engaged in bottomfishing and line fishing were included for all species except *C. ignobilis*, *L. rubrioperculatus*, and *L. kasmira*.
- Length measurements from trips that engaged in bottomfishing and net fishing were included for all species except *C. ignobilis*.
- Length measurements from trips that engaged in bottomfishing and spearfishing were included for all species except *C. ignobilis*, *C. lugubris*, *L. rubrioperculatus*, *L. kasmira*, and *Variola louti*.
- Length measurements from trips with fully unspecified fishing methods were included for all species that were not excluded from one of the above groups.

We retained biosampling length measurements from bottomfishing for all BMUS except *C. ignobilis*, which had the most length measurements from line fishing.

Next, we compared the median lengths between complete and incomplete biosampling catches for each BMUS (Table 1). Note, for this analysis, we included 37 additional length measurements (across all BMUS) that were collected in November and December, 2020, because incomplete vs. complete catch information was first recorded in November 2020. We sought to determine whether lengths from incomplete catches were different from complete catches, which would indicate non-random subsampling and preclude the use of lengths from incomplete catches in this analysis. We found lengths from incomplete catches were larger for 3 BMUS (*Pristipomoides auricilla, P. filamentosus*, and *P. zonatus*) and smaller for 3 BMUS (*Aphareus rutilans, Etelis*)

carbunculus, and *L. kasmira*) with these differences being significant as indicated by the non-parametric Wilcoxon rank sum test statistic p-value at the alpha = 0.05 level (function wilcox.test() from the R stats package, R Core Team 2024). For these 6 BMUS, we filtered the 2021–2023 biosampling length dataset to include length measurements from complete catches only, but we included length measurements from both incomplete and complete catches for the other BMUS.



Figure 2. Flowcharts detailing the processes applied to each BMUS to (a) filter biosampling data to identify usable length data, and (b) identify the data product representing recent length composition that can be produced from the filtered lengths. Candidate length composition data products are described in (c).

Table 1. Biosampling program number of length measurements (*N*) and median length (L_{50}) of fish from fishing trips with complete and incomplete catches. The nonparametric Wilcoxon rank-sum test statistic p-value is provided as a means to detect significant bias when selecting fish from incomplete catches, i.e. a test of H_0 that the median length is equivalent from complete and incomplete catches.

	Complete	Incomplete	Complete	Incomplete	Wilcoxon
	Catch N	Catch N	Catch L_{50}	Catch L_{50}	Rank Sum p
A. rutilans	204	74	35.1	32.5	0.012
C. ignobilis	58	4	71.8	74.2	0.875
C. lugubris	79	21	37.9	38.5	0.823
E. carbunculus	524	74	29.1	27.1	0.011
E. coruscans	252	103	70.6	71.0	0.837
L. rubrioperculatus	236	5	26.2	24.2	0.140
L. kasmira	607	138	19.9	19.1	< 1e-4
P. auricilla	3346	1945	25.1	27.0	< 1e-4
P. filamentosus	62	10	26.0	42.0	< 1e-4
P. flavipinnis	303	61	27.2	28.5	0.201
P. sieboldii	117	7	30.9	27.7	0.533
P. zonatus	513	148	24.1	25.9	0.008
V. louti	14	0	32.3		

Time-series Length Compositions from the BBS

We downloaded 1982–2023 BBS interview records from the Guam SQL-server Datawarehouse curated by the Western Pacific Fisheries Information Network (WPacFIN) on 01May, 2024. We followed a sequence of analyses and decisions to produce the filtered BBS dataset that was used to generate the time-series length compositions for each BMUS (Figure 3a. Filter BBS Data flowchart). First, we counted the number of length measurements by fishing method for each BMUS over 1982–2023. For all BMUS, bottomfishing captured the greatest number of length measurements, so we filtered the BBS length data to include bottomfishing only.

Next, we compared the median lengths between incomplete and complete BBS catch records for each BMUS (Table 2) to determine whether incomplete catch lengths were

different from complete catch lengths, which would indicate non-random subsampling and preclude the use of incomplete catch lengths in this analysis. Based on the non-parametric Wilcoxon rank sum test statistic p-value at the alpha = 0.05 level (R Core Team 2024), the analyses did not indicate the length distributions between incomplete and complete catch records were significantly different, hence we retained all lengths (from both incomplete and complete catch records) within the filtered BBS dataset for each BMUS.



Figure 3. Flowcharts detailing the processes applied to each BMUS to (a) filter boat-based creel survey (BBS) data to identify usable length data, and (b) identify the data product representing time-series length composition that can be produced from the filtered lengths. Candidate length composition data products are described in (c).

Table 2. Boat-based creel survey (BBS) program number of length measurements (*N*) and median length (L_{50}) of fish from fishing trips with complete and incomplete catches. The nonparametric Wilcoxon rank-sum test statistic p-value is provided as a means to detect significant bias when selecting fish from incomplete catches, i.e. a test of H_0 that the median length is equivalent from complete and incomplete catches.

	Complete	Incomplete	Complete	Incomplete	Wilcoxon
	Catch N	Catch N	Catch L ₅₀	Catch L ₅₀	Rank Sum p
A. rutilans	80	26	37.1	36.3	0.075
C. ignobilis	25	3	50.0	71.5	0.235
C. lugubris	34	12	36.5	50.4	0.078
E. carbunculus	96	68	34.0	34.1	0.640
E. coruscans	102	102	64.8	63.5	0.907
L. rubrioperculatus	126	302	23.7	24.6	0.181
L. kasmira	94	147	18.5	19.0	0.727
P. auricilla	102	256	25.4	24.6	0.185
P. filamentosus	28	15	31.0	31.3	0.558
P. flavipinnis	47	42	28.1	29.1	0.934
P. sieboldii	0	0			
P. zonatus	73	111	26.8	25.4	0.081
V. louti	76	36	28.8	26.8	0.067

Length Composition Data Products

A primary concern when using fishery-dependent data in stock assessment models is to account for spatial heterogeneity in fishing (or sampling) activity over time and space. This heterogeneity is accounted for in abundance indices through CPUE standardization, as is described for the Guam bottomfishery in preparation for the upcoming benchmark stock assessment (Bohaboy and Matthews, *in review*, "Standardized Catch per Unit Effort Indices for Bottomfish Management Unit Species of Guam, 1982-2023"). It is also important to standardize length composition data. The biosampling and BBS data include information on catch location ranging in detail from specific banks to unspecified locations, and the relatively small number of length observations across all regions for each BMUS and presence of ambiguous location identifiers (such as 'North') make it impossible to incorporate location information at the level provided in the datasets. Based on information provided by Guam fishermen

(Iwane et al. 2023), a reasonable minimum level of location information describing bottomfishing grounds around Guam is the distinction of the offshore banks from the slopes surrounding the main island of Guam. For the BBS data, this can be done by grouping the previously described 5 larger regions (Figure 1) into 'banks' and 'nearshore', and including locations not identifiable to either banks or nearshore within a third category, 'unspecified'. For the biosampling data, we assigned the 75 unique location names to 'banks', 'nearshore', or 'unspecified' following Appendix Table A1.

We made several assumptions regarding the bottomfish stocks and bottomfishery of Guam in our preparation of the length composition data. First, we assumed fish in the banks and nearshore regions are part of the same overall population, such that the population dynamics in all regions are common over time. This assumption is supported by hydrodynamic models that predict reef fish larval dispersion occurs between islands within the Mariana Archipelago (Kendall et al. 2018) and general east-west movement of subsurface ocean drifter tracks across the Mariana Archipelago in relatively short timeframes (1–10 days; Kendall and Poti 2014). Second, we assumed fishing gear specifications, such as hook size, are not different between regions. As a result, the gear (or contact) selectivity among regions is identical, while any differences in catch length composition among regions are driven solely by differential availability of fish populations to exploitation.

We performed analyses of the biosampling and BBS datasets, described below, to determine whether it was necessary to standardize the length composition data for each BMUS to account for the influence of uneven sampling among the banks and nearshore regions. When it was warranted, we used relative habitat area (e.g., area of the seafloor) to weight the contribution of length observations by region to the overall length composition. The weight of a nearshore length observation was calculated as:

$$\left(\frac{\text{area nearshore}}{\text{total area}}\right) / \left(\frac{N_N}{N_N + N_B}\right)$$

Where N_N is the total number of nearshore length measurements and N_B is the total number of banks length measurements for a given BMUS. The weight of a banks length observation was similarly calculated as:

$$\left(\frac{area \ banks}{total \ area}\right) / \left(\frac{N_B}{N_N + N_B}\right)$$

The overall ("weighted") length composition was then generated as the sum of the area-weighted length observations from the two regions.

Maunder et al. (2020) recommend spatial weighting of length composition data should be based on region-specific stock abundance when the objective is to represent the *population* size structure (as may be done for an 'index' fleet; Hoyle et al. 2024), but based on region-specific catch when the objective is to represent the *catch* size structure (as may be done for an 'extraction' fleet; Hoyle et al. 2024, particularly when there are multiple fleets that may be modeled separately within an integrated assessment model). There are no fisheries-independent survey data for Guam BMUS and there is only a single fleet for the Guam bottomfish fishery, hence there would be no data streams available to inform a stock assessment model on the population length compositions (e.g., fleet catchability) if the biosampling or BBS length data were assumed to describe only the catch. We chose instead to use the Guam bottomfish fishery length composition data to represent the *population* size structure.

We use area to weight the length compositions, instead of abundance, for three primary reasons: 1) the effects of region on the standardized CPUE indices are uncertain and commonly overlap among areas (Bohaboy and Matthews, in review, "Standardized Catch per Unit Effort Indices for Bottomfish Management Unit Species of Guam, 1982-2023"), hence incorporating region-specific CPUE estimates to weight regions by abundance risks the introduction of additional uncertainty into the length compositions. 2) Given the absence of reliable region-specific CPUE estimates, it may be appropriate to assume catch rates are approximately equal among regions, which would be expected in minimally fished populations, hence abundance is proportional to area. 3) In an extreme case where a population has been heavily depleted in one region but is essentially unfished in another, strongly weighting the length compositions based on the relatively high abundance in the unfished region would greatly under-represent the contribution of the length compositions in the depleted regions (Dichmont 2015). Instead, the resulting stock-wide length compositions would be similar to those expected in an unfished population, hence data-limited stock assessments based on fish length would likely underestimate stock-wide depletion and fishing mortality.

We defined the depth range used to calculate banks and nearshore area according to the general occurrences of each BMUS, based on feedback from Guam bottomfishermen (Bohaboy and Matthews, *in review*, "Standardized Catch per Unit Effort Indices for Bottomfish Management Unit Species of Guam, 1982-2023"; Iwane 2023). The shallower group consisted of *C. ignobilis*, *C. lugubris*, *L. rubrioperculatus*, *L. kasmira*, and *V. louti* and was assigned a habitat depth range of 0–100 m. The deeper group consisted of *A. rutilans*, *E. carbunculus*, *E. coruscans*, *P. auricilla*, *P. filamentosus*, *P. flavipinnis*, *P. sieboldii*, and *P. zonatus* and was assigned a habitat depth range of 100–400 m. The bottom habitat area within each depth range for the nearshore and banks regions (Figure 4) was calculated from the General Bathymetric Chart of the Oceans (GEBCO) 2023 global bathymetry 15 arc-second spatial resolution model (GEBCO Compilation Group, 2023).



Figure 4. Map of Guam's nearshore and banks regions, separated into shallower and deeper BMUS habitat.

Recent Length Compositions from the Biosampling Program

We followed a sequence of analyses and decisions to produce the biosampling data product for each BMUS (Figures 2b and 2c). Beginning with the filtered biosampling dataset, we qualitatively compared length frequency distributions of the nearshore and banks regions. If the modes of the two length distributions appeared to be notably different, or if low sample sizes from either region made it impossible to assess, we determined it would not be reasonable to combine length distributions from the two regions because it was possible that either of the assumptions regarding identical gear selectivity or similar population dynamics between regions may have been violated. In such instances, if the number of length measurements was large enough, we preserved the distinct length compositions for each region in the data product. If at least 20 length measurements were available for each of the nearshore and banks regions, we presented the area-specific biosampling data product, which contained separate length compositions for each of the nearshore and banks regions. If the majority of length measurements (40 or more) were from one region, and the other region was insufficiently represented (with fewer than 20 length observations), then only a single-region data product was presented, consisting of the length composition from the better-represented region. For BMUS with apparent unique length distributions by region that did not have a sufficient number of length measurements (e.g., one region had less than 20 samples and both regions had less than 40 samples), we did not present a data product.

If qualitative examination of the nearshore and banks length frequency distributions did not indicate a difference between the two regions, we considered the length measurements from the two regions were not sufficiently different to warrant weighting by area, and we used the aggregate length compositions across all regions (nearshore, banks, and unspecified) in the biosampling data product if there were at least 40 length measurements.

If qualitative examination of the nearshore and banks length frequency distributions indicated a difference between the two regions, and there were at least 20 length measurements for each region, we calculated the area-weighted length composition. If either region had less than 20 length measurements, we considered the sample size too small to calculate the area-weighted length composition, and instead used a single-area length composition if there were at least 40 length measurements from a region. For BMUS with notably different length compositions between regions that did not have a sufficient number of length measurements (e.g., one region had less than 20 samples and both regions had less than 40 samples), we evaluated appropriate data products on a case-by-case basis.

Time-series Length Compositions from the BBS

We followed a sequence of analyses and decisions to produce the BBS time-series data product for each BMUS (Figures 3b and 3c). The number of BBS length measurements for each BMUS was too small for yearly length compositions, so we had to bin years into time blocks. We selected time intervals for each BMUS that would allow for the greatest number of time blocks over the time-series and at least 40 length measurements per time block. Beginning with the filtered BBS dataset, we qualitatively compared the length distributions between the banks and nearshore regions (over all years combined). If there was no apparent difference between the two regions, we concluded the length measurements from the two regions were not sufficiently different to warrant weighting by area, and we used the aggregate length compositions across all regions (nearshore, banks, and unspecified) in the BBS data product if there were sufficient length measurements to allow the formation of time blocks as described previously (i.e., if there were at least 40 length measurements in each of 2 time blocks).

If qualitative examination of the nearshore and banks length frequency distributions indicated a difference between the two regions, and there were at least 20 length measurements for each region in each of 2 time blocks, we calculated the area-weighted length composition (by time block). If either region had less than 20 length measurements per time block, or time blocks could not be assigned due to very low sample size, we did not present a time-series data product.

RESULTS

The large quantity of length data from the biosampling program allowed usable recent length compositions to be generated for 12 of the 13 BMUS. The exception was *A. rutilans*, for which peculiar patterns in the length composition could not be explained. With length data pooled across all available years (2009–2023), 10 of the 13 BMUS had sufficient region-specific data to calculate an area-weighted length composition. For most of these species, larger fish were caught at the banks than nearshore. The relative proportion of lengths by region rarely matched the estimated species habitat by region, causing the weighted length composition to deviate from the aggregate length composition. Forcusing on recent years (2021–2023), only 4 BMUS had sufficient region-specific data. For the other species, the aggregate length composition from recent years was used if there was no difference detected between regions when pooling across years. Otherwise, region-specific length compositions were produced.

The BBS provides a smaller quantity of length data per year, but time-series length composition could still be generated for 9 of the 13 BMUS. Of the 4 remaining species, 3 had insufficient data and one (*E. carbunculus*) could not be used prior to 2020 due to species identification issues. Data limitations still impacted many of the species with generated time-series. Weighted length compositions were required for 8 of the 9 species due to differences in length between the nearshore and banks regions. The small quantity of banks data consequently limited the temporal resolution of the time-series, with a maximum of 6 time blocks possible across the species. Still, time-series for the 9 species are a valuable resource to capture changes over the 42-year time span of the BBS.

Aphareus rutilans

Recent Length Composition

There are sufficient length data by region when pooled across all years for A. rutilans, and larger fish are caught at the banks than nearshore. Only 16.6% of lengths from identifiable regions are from the banks, while an estimated 45.6% of its habitat is at the banks. This causes the weighted length composition to deviate from the aggregate length composition, with a greater density of large individuals when lengths from the banks are given appropriate weight. Unfortunately, insufficient banks data are available in recent years, so the nearshore length composition must be used.

There is also a peculiar bimodal structure to the length compositions, with a second peak around 80 cm. This could arise from the use of unique fishing gear configurations to target large *A. rutilans*, or frequent misidentifications with smaller *Aphareus spp*. Both issues could preclude the use of length data for this species.

Outcome: Unless further investigations can identify the cause of the bimodal size structure, it is unlikely that length composition data for *A. rutilans* can be used.



Figure 2. *A. rutilans* length frequency for (a) 2021–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

There are limited data to construct a time-series length composition for *A. rutilans*. As with biosampling data, the BBS data indicate that larger fish are caught at the banks than nearshore. Unfortunately, limited lengths are available from the banks over the last two decades. This restricts the number of time blocks to two, while retaining sufficient banks data within each to generate weighted length compositions.

Outcome: The time-series length composition for *A. rutilans* can be used.



Figure 3. *A. rutilans* length frequency (a) time-series, (b) data availability by year, time block, and region (N = nearshore, B = banks, U = unspecified, T = total), and (c) aggregated across 1982–2023.

Caranx ignobilis

Recent Length Composition

There are insufficient length data from bottomfishing for *C. ignobilis* at all regional and temporal scales. It is rarely targeted by bottomfishers (Iwane et al. 2023), as corroborated by its low occurrence in bottomfishing interviews (1.5%) during the boat-based creel survey. Instead, data from line fishing were used. Still no length data are available from the banks, so it is unclear if length composition differs between regions. For this reason, the nearshore length composition from recent years must be used.

Outcome: The nearshore length composition from recent years for *C. ignobilis* is available, however, there are insufficient data from the banks to understand how length compositions may differ between regions.



Figure 4. *C. ignobilis* length frequency for (a) 2021–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

There are extremely limited data from the BBS to construct a time-series length composition for *C. ignobilis*. There are also insufficient lengths from the banks to identify potential regional differences, which was also an issue with biosampling data. A nearshore-only time-series with two time blocks could be generated, but with extremely limited length data in each block (N = 49 and N = 40).

Outcome: There is no usable time-series length composition for *C. ignobilis.*



Figure 5. *C. ignobilis* (a) data availability for length frequency time-series by year and region (N = nearshore, B = banks, U = unspecified, T = total), and (b) length frequency aggregated across 1982–2023.

Caranx lugubris

Recent Length Composition

There are sufficient length data by region when pooled across years for C. lugubris, and length data from the banks and nearshore are broadly similar. Subsequently, the weighted and aggregate length compositions pooled across years are very similar. Insufficient banks data are available from recent years, so only an aggregate length composition is possible. Given the similar length composition by region across years and moderate sample size from recent years, the aggregate length composition from recent years may be usable. Notably, lengths data from nearshore has been over-represented in recent years, with 88.5% of lengths from identifiable regions in recent years, compared to 56.9% from all years. Its estimated nearshore habitat contribution is intermediate these values at 75.8%. If a difference in length compositions between the regions has emerged in recent years, the aggregate length composition for recent years may be biased. However, given the similarity between regions across all years, the length composition for recent years is usable.

Outcome: The aggregate length composition from recent years for *C. lugubris* can be used.



Figure 6. *C. lugubris* length frequency for (a) 2021–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

There are limited data to construct a time-series length composition for *C. lugubris*. Unlike with biosampling data, the BBS data indicate that slightly larger fish are caught nearshore than at the banks. This restricts the number of time blocks to three, while retaining sufficient region-specific data within each to generate weighted length compositions.

Outcome: The time-series length composition for *C. lugubris* can be used.



Figure 7. *C. lugubris* length frequency (a) time-series, (b) data availability by year, time block, and region (N = nearshore, B = banks, U = unspecified, T = total), and (c) aggregated across 1982–2023.

Etelis carbunculus

Recent Length Composition

There are sufficient length data by region from recent years for *E. carbunculus*. Length compositions are similar between regions, though slightly larger fish may be caught at the banks. Nearshore lengths are over-represented relative to the habitat they provide, but the weighted and aggregate length compositions are still similar.

There is evidence that *E. boweni* are present in the data. In 2022 five *E. carbunculus* over 60 cm were caught, including four in one trip. This is larger than *E. carbunculus* maximum size (Dahl et al. 2024), meaning they were likely *E. boweni*. The lengths are not included in Figure 5 only because the entire trip's catch was not measured. Given the different growth trajectories of the two species (Wakefield et al. 2020), the potential presence of *E. boweni* may compromise the use of these data.

Dahl and others (2024) used otolith morphometrics and FT-NIR spectroscopy trained with a set of *E. carbunculus* and *E. boweni* voucher otoliths to identify probable *E. boweni* otoliths collected during the biosampling program through 2019. Both methods estimated that only 8% of identified *E. carbunculus* were actually *E. boweni*. Large individuals (> 60 cm) were disproportionately represented in the purported *E. boweni*, which could indicate that smaller size classes are minimally contaminated with *E. boweni*. This biased size structure could arise from Guam existing beyond the initially published geographic range of *E. boweni* (Andrews et al. 2021; K. Dahl, pers. comm.). These results indicate that *E. boweni* contributes minimally within the size range of *E. carbunculus*.

Outcome: The weighted length composition from recent years for *E. carbunculus* can be used.



Figure 8. (a) *E. carbunculus* length frequency for 2021–2023, and length measurements by region for (c) each year and (d) cumulatively going back in time.

Given misidentification issues with *E. boweni* prior to the year 2020, we did not attempt to construct a time-series length composition for *E. carbunculus*.

Outcome: There are no usable time-series length compositions for *E. carbunculus*.



Figure 9. *E. carbunculus* (a) data availability for length frequency time-series by year and region (N = nearshore, B = banks, U = unspecified, T = total), and (b) length frequency aggregated across 2020–2023.

Etelis coruscans

Recent Length Composition

A large amount of region-specific length data are available for *E. coruscans* when pooled across years, and a moderate but usable amount is available for recent years. Samples pooled over all years suggest larger fish are caught at the banks, but little difference is detectable for recent years.

Banks length data are overrepresented at both temporal scales, with 67.5% and 61.6% of lengths from the banks for recent and all years, respectively, compared to an estimated 45.6% of its habitat. Still, the weighted and aggregate length compositions are quite similar to each other in both cases. Because sufficient region-specific length data are available for recent years, the weighted length composition is preferred.

Outcome: The weighted length composition from recent years for *E. coruscans* can be used.



Figure 10. *E. coruscans* length frequency for (a) 2021–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

There are limited data from early years of the BBS to construct time-series length compositions for *E. coruscans*. Whereas biosampling data indicate mixed results for differences between regions, the BBS length data from the banks and nearshore are broadly similar. Even though an aggregate length composition can be used, the limited data prior to the year 2000 allows only four time blocks to be assigned.

Outcome: The time-series length composition for *E. coruscans* can be used.



Figure 11. *E. coruscans* length frequency (a) time-series, (b) data availability by year, time block, and region (N = nearshore, B = banks, U = unspecified, T = total), and (c) aggregated across 1982–2023.

Lethrinus rubrioperculatus

Recent Length Composition

There are sufficient length data by region when pooled across years for L. rubrioperculatus, which indicate slightly larger fish are caught at the banks. The weighted and aggregate length compositions are quite similar, even though banks length measurements (57.0% of lengths from identified regions) are overrepresented relative to the habitat they are estimated to provide (24.2%). Notably, the majority of fish were caught at the banks from 2010–2014, while the majority of lengths from identified regions since then are from nearshore. It is not clear why this shift occurred, though the vast majority of lengths collected since 2015 have been from unspecified regions and may include abundant data from the banks. Still, insufficient data known to be from the banks are available from recent years, so only an aggregate length composition is possible. Given the similarity between the weighted and aggregate length compositions when pooled across years and good data quantity from recent years, the aggregate length composition from recent years is usable.

Outcome: The aggregate length composition from recent years for *L. rubrioperculatus* can be used.



Figure 12. *L. rubrioperculatus* length frequency for (a) 2021–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

There are abundant length data for *L. rubrioperculatus* from the BBS, though most are from the early years of the survey and banks data are much less numerous than nearshore data. As with biosampling data, the BBS data indicate slightly larger fish are caught at the banks than nearshore. Six time blocks can be assigned, while retaining sufficient region-specific data within each to generate weighted length compositions.

Outcome: The time-series length composition for *L. rubrioperculatus* can be used.



Figure 13. *L. rubrioperculatus* length frequency (a) time-series, (b) data availability by year, time block, and region (N = nearshore, B = banks, U = unspecified, T = total), and (c) aggregated across 1982–2023.

Lutjanus kasmira

Recent Length Composition

There are sufficient length data by region when pooled across years for L. kasmira, which indicate larger fish are caught at the banks. Only 10.3% of lengths from identifiable regions are from the banks, while an estimated 24.2% of its habitat is at the banks. This causes the weighted length composition to deviate from the aggregate length composition, with a greater density of large individuals when the banks lengths are given appropriate weight. Despite the large amount of length data from recent years, only 2.7% of lengths from identifiable regions are from the banks. The aggregate length composition from recent years is likely biased, so the nearshore length composition from recent years must be used.

Outcome: The nearshore length composition from recent years for *L. kasmira* is available, however, there are insufficient data from the banks to understand how length compositions may differ between regions.



Figure 14. *L. kasmira* length frequency for (a) 2021–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

There are abundant length data for *L. kasmira* from the BBS, though banks data are much less common than nearshore data. As with biosampling data, the BBS data indicate that larger fish are caught at the banks than nearshore, with a more pronounced distinction in earlier years. Four time blocks can be assigned, while retaining sufficient region-specific data within each to generate weighted length compositions.

Outcome: The time-series length composition for *L. kasmira* can be used.



Figure 15. *L. kasmira* length frequency (a) time-series, (b) data availability by year, time block, and region (N = nearshore, B = banks, U = unspecified, T = total), and (c) aggregated across 1982–2023.

Pristipomoides auricilla

Recent Length Composition

A very large amount of region-specific length data are available for *P. auricilla*, which indicate larger fish are caught at the banks. Only 15.9% of lengths across all years and 9.7% of lengths from recent years are from banks, while an estimated 45.6% of its habitat is at the banks. This causes the weighted length composition to have a greater density of large individuals than the aggregate length composition, particularly for the recent years. This is an excellent example of a length composition with sufficient data.

Outcome: The weighted length composition from recent years for *P. auricilla* can be used.



Figure 16. *P. auricilla* length frequency for (a) 2021–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

There are abundant length data for *P. auricilla* from the BBS, though banks data are much less common than nearshore data. As with biosampling data, the BBS data indicate that larger fish are caught at the banks than nearshore. Five time blocks can be assigned, while retaining sufficient region-specific data within each to generate weighted length compositions.

Outcome: The time-series length composition for *P. auricilla* can be used.



Figure 17. *P. auricilla* length frequency (a) time-series, (b) data availability by year, time block, and region (N = nearshore, B = banks, U = unspecified, T = total), and (c) aggregated across 1982–2023.

Pristipomoides filamentosus

Recent Length Composition

There are insufficient banks length data for P. filamentosus, even pooled across all years. Only 11.9% of lengths from identifiable regions are from the banks, while an estimated 45.6% of its habitat is at the banks. It is unclear if length composition differs between these regions, but if it does, the aggregate length compositions will be biased toward the nearshore composition. For this reason, the nearshore length composition from recent years must be used. With only 28 length measurements from 2021-2023, a total of seven years need to be aggregated to meet the minimum required sample size of 40.

P. filamentosus can be difficult to distinguish from *P. sieboldii* (Iwane et al. 2023) and there it is possible these species have been mis-identified in the biosampling data. Given the different growth trajectories of the two species, presence of *P. sieboldii* in the *P. filamentosus* length data may compromise the use of these data in an assessment.

Outcome: The nearshore length composition from recent years for *P. filamentosus* is available, however, there are insufficient data from the banks to understand how length compositions may differ between regions.



Figure 18. *P. filamentosus* length frequency for (a) 2017–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

There are limited data to construct a time-series length composition for *P. filamentosus*. Whereas there were too few banks lengths from the biosampling program to identify regional differences, the BBS data indicate that larger fish are caught at the banks than nearshore. This requires weighted length compositions to be used, but unfortunately, there are insufficient data from the banks to even generate a time-series with two time blocks.

Outcome: There is no usable time-series length composition for *P. filamentosus*.



Figure 19. *P. filamentosus* (a) data availability for length frequency time-series by year and region (N = nearshore, B = banks, U = unspecified, T = total), and (b) length frequency aggregated across 1982–2023.

Pristipomoides flavipinnis

Recent Length Composition

There are sufficient length data by region when pooled across years for P. flavipinnis, and it indicates that larger fish are caught at the banks. Only 9.8% of lengths from identifiable regions are from the banks, while an estimated 45.6% of its habitat is at the banks. This causes the weighted length composition to differ from the aggregate length composition, with a greater density of large individuals when the banks lengths are given appropriate weight. Despite the large amount of length data from recent years, only 3.1% of lengths from identifiable regions are from the banks. The aggregate length composition from recent years is likely biased, so the nearshore length composition from recent years must be used.

Outcome: The nearshore length composition from recent years for *P. flavipinnis* is available, however, there are insufficient data from the banks to understand how length compositions may differ between regions.



Figure 20. *P. flavipinnis* length frequency for (a) 2021–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

There are limited data to construct a time-series length composition for *P. flavipinnis*. As with biosampling data, the BBS data indicate that larger fish are caught at the banks than nearshore, with a more pronounced distinction in earlier years. This restricts the number of time blocks to three, while retaining sufficient region-specific data within each to generate weighted length compositions.

Outcome: The time-series length composition for *P. flavipinnis* can be used.



Figure 21. *P. flavipinnis* length frequency (a) time-series, (b) data availability by year, time block, and region (N = nearshore, B = banks, U = unspecified, T = total), and (c) aggregated across 1982–2023.

Pristipomoides sieboldii

Recent Length Composition

There are sufficient length data by region when pooled across years for *P. sieboldii*, and length data from the banks and nearshore are broadly similar. Consequently, the weighted and aggregate length compositions pooled across years are very similar. While a small but usable amount of length data are available from recent years to generate a weighted length composition, the aggregate length composition is preferred given its greater sample size and similarity to the weighted length composition.

There is evidence that large P. filamentosus are misidentified as P. sieboldii. Kami (1973) reports that P. sieboldii reaches 40 cm fork length in Guam, while several very large (55+ cm) individuals have been logged in the biosampling program. P. filamentosus is reported to reach 75 cm (Kami 1973) and given reported difficulty distinguishing the two species (Iwane et al. 2023), these large fish were likely misidentified. Given the different growth trajectories of the two species, the likely presence of *P. filamentosus* in the data may compromise the use of these data in an assessment.

Outcome: The aggregate length composition from recent years for *P. sieboldii* can be used if *P. filamentosus* is known to contribute negligibly.



Figure 21. *P. sieboldii* length frequency for (a) 2021–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

Time-series Length Composition

Outcome: No length data have been collected for *P. sieboldii* in the BBS, so no time-series can be generated.

Pristipomoides zonatus

Recent Length Composition

A large amount of region-specific length data are available for P. zonatus when pooled across years, which indicate larger fish are caught at the banks. Only 24.2% of lengths across all years and 6.5% of lengths from recent years are from banks, while an estimated 45.6% of its habitat is at the banks. This causes the weighted length composition to have a greater density of large individuals than the aggregate length composition, particularly for the recent years. Though banks data are limited for P. zonatus in recent years, it still indicates the same trend of greater fish lengths and is sufficient to compute the weighted length composition.

Outcome: The weighted length composition from recent years for *P. zonatus* can be used.



Figure 22. *P. zonatus* length frequency for (a) 2021–2023 and (b) all years, and length measurements by region for (c) each year and (d) cumulatively going back in time.

There are a moderate amount of length data for *P. zonatus* from the BBS, though banks data are much less common than nearshore data. As with biosampling data, the BBS data indicate that larger fish are caught at the banks than nearshore. Five time blocks can be assigned, while retaining sufficient region-specific data within each to generate weighted length compositions.

Outcome: The time-series length composition for *P. zonatus* can be used.



Figure 23. *P. zonatus* length frequency (a) time-series, (b) data availability by year, time block, and region (N = nearshore, B = banks, U = unspecified, T = total), and (c) aggregated across 1982–2023.

Variola louti

Recent Length Composition

There are insufficient nearshore length data from the biosampling program for V. louti even pooled across all years. Interestingly, only 7.6% of lengths from identifiable regions are from nearshore, while an estimated 75.8% of its habitat is nearshore. Schemmel and Dahl (2023) report that V. louti is caught in Guam's shallow bottomfishery, so the nearshore length composition will be underrepresented in the aggregate length composition unless it is similar to the banks length composition. While the biosampling program has collected limited data from recent years, the BBS provides sufficient nearshore lengths to generate a length composition for recent years. A total of six years need to be aggregated to meet the minimum required sample size of 40.

Outcome: The nearshore length composition from recent years for *V. louti* is available, however, there are insufficient data from the banks to understand how length compositions may differ between regions.



Figure 24. *V. louti* length frequency for (a) 2018–2023 from the boat-based creel survey and (b) all years from the biosampling program, and length measurements by region for (c) each year and (d) cumulatively going back in time from the biosampling program.

There are limited data to construct a time-series length composition for *V. louti.* Whereas there were too few nearshore lengths from the biosampling program to identify regional differences, the BBS data indicate that slightly larger fish are caught at the banks than nearshore, with a more pronounced distinction in earlier years. This restricts the number of time blocks to three, while retaining sufficient region-specific data within each to generate weighted length compositions.

Outcome: The time-series length composition for *V. louti* can be used.



Figure 25. *V. louti* length frequency (a) time-series, (b) data availability by year, time block, and region (N = nearshore, B = banks, U = unspecified, T = total), and (c) aggregated across 1982–2023.

ACKNOWLEDGEMENTS

We greatly appreciate the biosampling program and DAWR survey staff, fishermen, and fish vendors of Guam who contributed to the collection of the BMUS length data presented in this working paper. We also appreciate Guam fishermen and community members who provided us further knowledge informally and through the BMUS data workshops held on Guam in order to improve our analyses of the data. We thank the NOAA Pacific Islands Fisheries Science Center Western Pacific Fisheries Information Network for managing the biosampling program and BBS databases and working with us to access and understand the data. We would also like to recognize several staff at the Pacific Islands Fisheries Science Center for their contributions to the writing of this working paper and the information and analyses presented herein: Robert Ahrens, Felipe Carvalho, Eric Cruz, Mia Iwane, Marc Nadon, and Marlowe Sabater.

LITERATURE CITED

- Andrews, K. R., Fernandez-Silva, I., Randall, J. E., & Ho, H. 2021. *Etelis boweni* sp. Nov., a new cryptic deepwater eteline snapper from the Indo-Pacific (Perciformes: Lutjanidae). Journal of Fish Biology, 99(2), 335–344.
- Bohaboy, E. C. & Matthews, T. 2024. Stock assessment update of the bottomfish management species of Guam, 2024. U.S. Dept. of Commerce, NOAA Technical Memorandum NMFS-PIFSC-##, ## p.
- Dahl, K., O'Malley, J., Barnett, B., Kline, B., & Widdrington, J. 2024. Otolith morphometry and Fourier transform near-infrared (FT-NIR) spectroscopy as tools to discriminate archived otoliths of newly detected cryptic species, *Etelis carbunculus* and *Etelis boweni*. Fisheries Research, 272.
- Dichmont, C. M. 2015. Center for Independent Experts (CIE) independent peer review of length-based assessment methods of coral reef fish stocks in Hawaii and other U.S. Pacific territories.
- GEBCO Compilation Group. (2023). General Bathymetric Chart of the Oceans (GEBCO) 2023 Grid.

https://doi.org/10.5285/f98b053b-0cbc-6c23-e053-6c86abc0af7b.

- Hoyle, S. D., R. A. Campbell, N. D. Ducharme-Barth, A. Grüss, B. R. Moore, J. T. Thorson, L. Tremblay-Boyer, H. Winker, S. Zhou, and M. N. Maunder. 2024. Catch per unit effort modelling for stock assessment: A summary of good practices. Fisheries Research 269:106860.
- Iwane, M., Cruz, E., & Sabater, M. 2023. 2023 Guam bottomfish management unit species data workshops. U.S. Dept. of Commerce, NOAA Administrative Report H-23-07, 69 p.

- Jasper, W., Matthews, T., Gutierrez, J., Flores, T., Tibbatts, B., Martin, N., Bass, J.,
 Franquez, R., Manibusan, F., Ducusin, J., Regis, A., Lowe, M. K., & Quach, M.
 2016. DAWR creel survey methodology. Division of Aquatic & Wildlife Resources,
 Guam Department of Agriculture, Technical Report 1.
- Kami, H. T. 1973. The *Pristipomoides* (Pisces: Lutjanidae) of Guam with notes on their biology. Micronesia, 9(1), 97-118.
- Kendall, M. S., and M. Poti. 2014. Potential larval sources, destinations, and self-seeding in the mariana archipelago documented using ocean drifters. Journal of Oceanography 70:549–557.
- Kendall, M. S., M. Poti, and A. Winship. 2018. Is Guam a regional source, destination, or stepping-stone for larvae of three fisheries species? Fisheries Oceanography 28:159–170.
- Langseth, B., Syslo, J., Yau, A., & Carvalho, F. (2019). Stock Assessments of the Bottomfish Management Unit Species of Guam, the Commonwealth of the Northern Mariana Islands, and American Samoa, 2019. https://doi.org/10.25923/bz8b-ng72.
- Maunder, M. N., J. T. Thorson, H. Xu, R. Oliveros-Ramos, S. D. Hoyle, L. Tremblay-Boyer, H. H. Lee, M. Kai, S. K. Chang, T. Kitakado, C. M. Albertsen, C. V. Minte-Vera, C. E. Lennert-Cody, A. M. Aires-da-Silva, and K. R. Piner. 2020. The need for spatio-temporal modeling to determine catch-per-unit effort based indices of abundance and associated composition data for inclusion in stock assessment models. Fisheries Research 229:105594. https://doi.org/10.1016/j.fishres.2020.105594.
- R Core Team. 2024. The R Stats Package. https://www.r-project.org/.
- Schemmel, E., & Dahl, K. 2023. Age, growth, and reproduction of the yellow-edged lyretail *Variola louti* (Forssakal, 1775). Environmental Biology of Fishes, 106, 1247–1263.
- Sundberg, M., Humphreys, R., Lowe, M. K., Cruz, E., Gourley, J., & Ochavillo, D. 2015. Status of life history sampling conducted through the commercial fisheries bio-sampling programs in the Western Pacific Territories of American Samoa and Guam and in the Commonwealth of the Northern Mariana Islands. U.S. Dept. of Commerce, NOAA Administrative Report H-15-08.

APPENDIX

Banks	Nearshore	Unspecified	
11 mile bank 45 or Pati banks Galvez bank Ice Box Rota Banks Santa Rosa Southern Bank White Tuna Bank	Adelup Agana Agat Ague Cove Alupang Anai Apra Harbor Asan Asiga/Malojloj Behind Cocos Island Boat Basin Cabras Castro Beach Cetti Bay Cocos Lagoon Cocos/Umatac Double Reef East Agana East/Southeast Facpi Facpi to Fort Apugan Guam Port Authority Hanom Haputo Hawaiian Rock (E/NE) Hospital Point Inarajan Inarajan to Merizo backside Cocos	Ipan Magandas Malojloj Marble Cave Merizo NCS NW Quadrant Oka Point Orote Point Pagat Pago Bay Paseo Pati Point Piti Port Ritidian Point Sarigan Southeast Talofofo Talofofo to Marble Cave Tamuning Tarage Togcha Tumon Bay Two Lover's Point Umatac Urunao to Ritidian Ylig Yona	East NE Quadrant North Rota South Southwest SW and SE Quadrants West

Table A1. Biosampling program location names categorized by region (banks, nearshore, and unspecified).