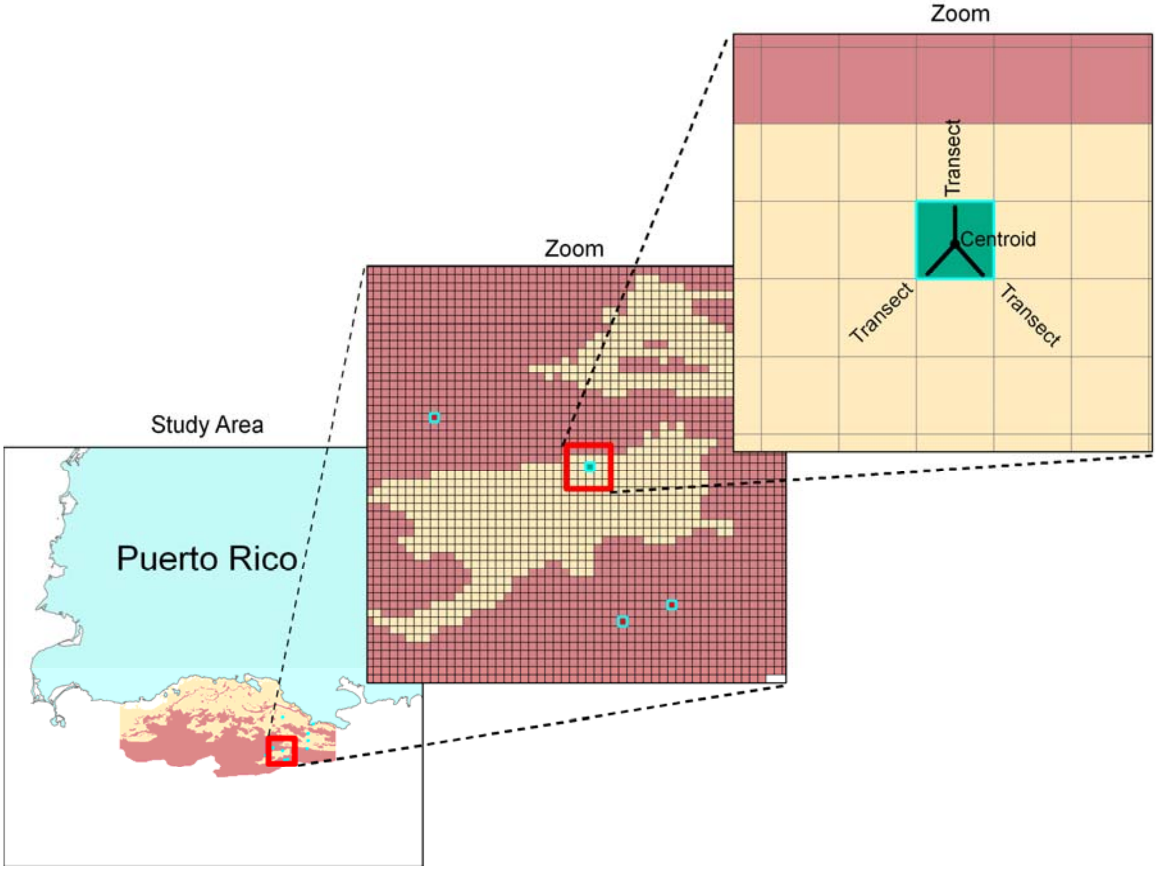
Analysis of Sample Frames and Subsampling Methods for Reef Fish Surveys



NOAA Technical Memorandum NOS NCCOS 72

Citation for this Report

Menza, C., C. Caldow, C. Jeffrey and M. Monaco. 2008. Analysis of Sample Frames and Subsampling Methods for Reef Fish Surveys. NOAA Technical Memorandum NOS NCCOS 72. 14 pp.

Analysis of Sample Frames and Subsampling Methods for Reef Fish Surveys

Charles Menza, Chris Caldow, Chris Jeffrey and Mark Monaco

Center for Coastal Monitoring and Assessment (CCMA)
NOAA/NOS/NCCOS
1305 East West Highway (SSMC-IV, N/SCI-1)
Silver Spring, MD 20910

NOAA Technical Memorandum NOS NCCOS 72

September 2007

United States Department of Commerce	National Oceanic and Atmospheric Administration	National Ocean Service
Donald L. Evans Secretary	Conrad C. Lautenbacher, Jr. Administrator	Jack Dunnigan Assistant Administrator

Table of Contents

Introduction

The National Oceanic and Atmospheric Administration's Biogeography Branch has conducted surveys of reef fish in the Caribbean since 1999. Surveys were initially undertaken to identify essential fish habitat, but later were used to characterize and monitor reef fish populations and benthic communities over time. The Branch's goals are to develop knowledge and products on the distribution and ecology of living marine resources and provide resource managers, scientists and the public with an improved ecosystem basis for making decisions.

The Biogeography Branch monitors reef fishes and benthic communities in three study areas: (1) St. John, USVI, (2) Buck Island, St. Croix, USVI, and (3) La Parguera, Puerto Rico. In addition, the Branch has characterized the reef fish and benthic communities in the Flower Garden Banks National Marine Sanctuary, Gray's Reef National Marine Sanctuary and around the island of Vieques, Puerto Rico.

Reef fish data are collected using a stratified random sampling design and stringent measurement protocols. Over time, the sampling design has changed in order to meet different management objectives (i.e. identification of essential fish habitat vs. monitoring), but the designs have always remained:

- Probabilistic – to allow inferences to a larger targeted population,
- Objective – to satisfy management objectives, and
- Stratified – to reduce sampling costs and obtain population estimates for strata.

There are two aspects of the sampling design which are now under consideration and are the focus of this report: first, the application of a sample frame, identified as a set of points or grid elements from which a sample is selected; and second, the application of subsampling in a two-stage sampling design. To evaluate these considerations, the pros and cons of implementing a sampling frame and subsampling are discussed. Particular attention is paid to the impacts of each design on accuracy (bias), feasibility and sampling cost (precision). Further, this report presents an analysis of data to determine the optimal number of subsamples to collect if subsampling were used.

Sampling Design

The sampling design used by the Biogeography Branch for monitoring was chosen to collect reef fish data in an unbiased manner with minimal logistic requirements and few decision rules. Over time, this sampling design has proven its value in acquiring unbiased data in moderate spatial scales (>10 km^2 and <150 km^2) of nearshore Caribbean reef ecosystems. Requirements include a team of 2-10 divers, 1-3 small boats, SCUBA equipment for each diver and 8-12 days in the field. Sampling costs are moderate ($1000 /day for boats and fuel, plus $40 /diver/day for tanks, plus $300 /diver/day for hotel and food), and generally fall within projection allowing between 90 and 140 dives during a 10 day mission. The sampling fraction is typically very low (<1%).

A random sample of reef fish survey locations is selected during each field mission by selecting random spatial coordinates within hardbottom (reef, pavement) and softbottom (sand, seagrass) benthic habitat types delineated on a benthic habitat map. A description of the benthic habitat maps used is provided by Kendall et al. (2001). At each selected geographic coordinate, a belt-transect is used to collect fish community and population measurements (Figure 1). The benthic habitat map defines areas to be sampled and the transects serve as sampling units. There is no sample frame as defined by Cochran (1977). Sample units are randomly placed in an area, not selected from a finite list of mutually-exclusive units.

The simplicity of the selection technique reduces biases imposed by decision rules, simplifies computations for population and community estimates and makes the design easily adaptable in the field to indeterminate events (e.g. strong currents, unfeasible sample location). The only requirements are a benthic habitat map and a computer program to select random coordinates from the map.

The use of points for sampling units simplifies the assignment of stratum designations in a stratified sampling design where strata are defined by a benthic habitat map. A point can only belong to one stratum; whereas a line or polygon requires a decision rule (e.g. sampling unit belongs to stratum A if majority of area in benthic habitat type A). Since the area of benthic habitats or strata can be calculated from the map it is simple to estimate probabilities of sampling unit selection. Sampling weights are estimated by comparing the area of the benthic habitat type sampled to the area surveyed (i.e. the transect). In addition, the designation of a sampling unit by a spatial coordinate or point allows workers in the field to easily adapt to often unknown and changing environmental conditions. Whenever possible, each belt-transect radiates from a coordinate at a random bearing in order to eliminate measurement bias imposed by a survey diver (i.e. transect towards reef), but in some cases this is not feasible. For instance, in cases of strong currents a transect can be pointed in the direction of flow. The ability to adapt to less than perfect field conditions allows data acquisition which may not be possible with more complicated systems.

There are several disadvantages of this sampling approach. First, the approach lacks a sample frame – defined as a finite list of mutually-exclusive sample units. The application of sampling theory, which is essential to data analysis, requires that every sampling unit has a known (prespecified) chance of selection. To get around this requirement, the benthic habitat map is used to estimate sampling weights, sampling probabilities and sampling fractions. Consequently, critical sampling parameters can be inaccurate and used by critics to devalue findings.

The lack of a sample frame and the manner in which sample units are selected also means sample units can overlap or extend past the boundaries of strata or even the target population. The inaccuracy of estimates associated with transect overlap and containment can be considered negligible when the sampling fraction (f) is very small, but inaccuracy increases as the sampling fraction approaches 1. When transects are densely packed into an area (i.e. when f is close 1) they are more likely to overlap and are less likely to be completely contained within the boundaries of a given polygon. The

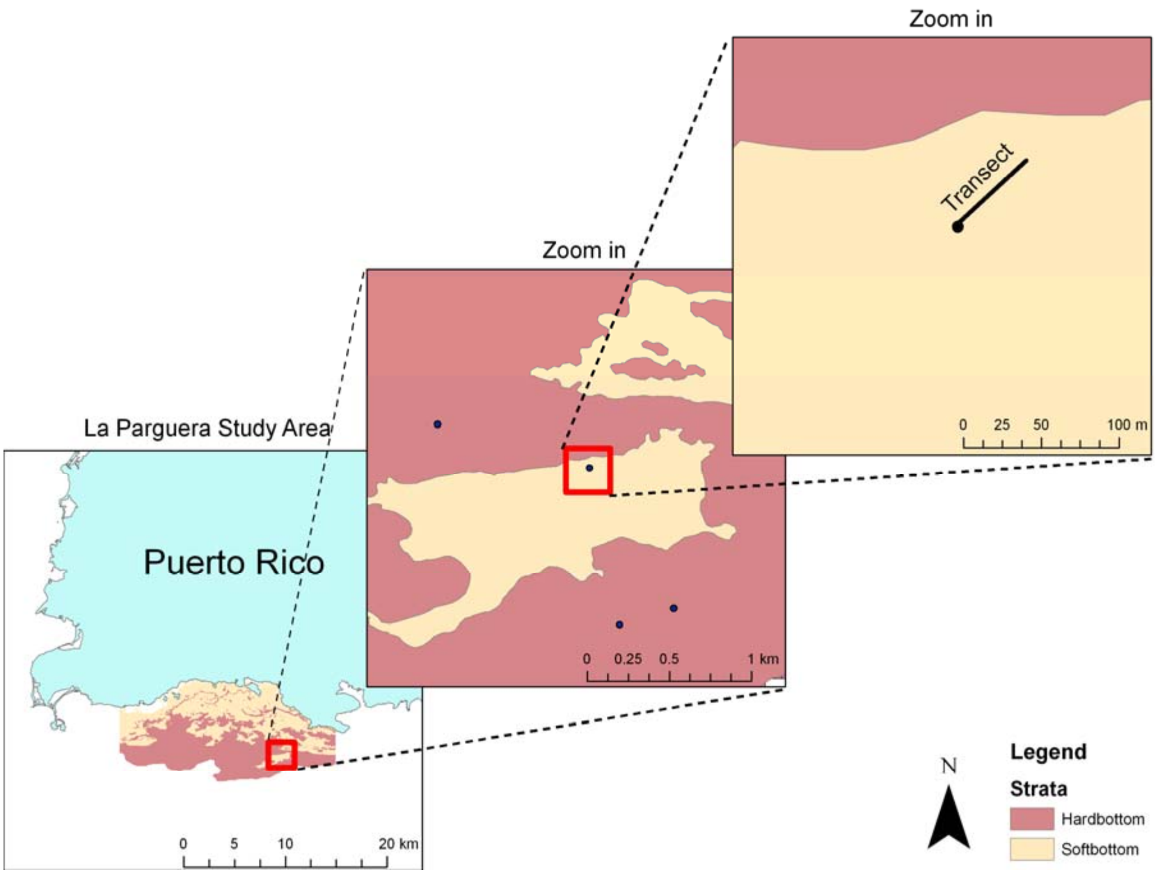


Figure 1: Sampling method used by the Biogeography Branch to sample reef fish communities in La Parguera, Puerto Rico

former decreases the variance of measured items, whereas the latter will have an unknown affect on the variance. These inaccuracies cannot be quantified easily.

The absence of a sample frame also means sample unit designation into particular strata cannot be changed if it is determined the unit belongs to a different stratum. If a transect which was supposed to be on reef habitat actually fell on sand habitat, there is no simple way to update the benthic habitat map. This revision is important because these errors increase the variance of estimates and decrease the accuracy of population estimates. Also, if the error in benthic habitat maps is not random among habitat types then this type of error will introduce a systematic bias into population estimates.

The application of a benthic habitat map presents a problem with the calculation of stratum weights. Commonly, there is non-random overlap of the fish community among habitat types. This is an issue, because a major goal of stratification is to divide the population into distinct fish communities (i.e. internally homogeneous groups). Take, for instance, an isolated reef in the middle of a sand patch. It is common to see species typically found on reef habitat moving some distance off the reef into sand habitat. The movement blurs any potential difference in fish communities among distinct habitats and thus decreases the utility of using benthic habitats to parse the fish community. The problem of community overlap is especially problematic when a subset of habitat types is sampled. By constraining sampling to a particular habitat type, some elements of the population may be selectively missed introducing bias into estimates. When habitat transitions are used to divide strata, the result can be an underestimate or overestimate of stratum weights and derived population parameters. The direction of the bias depends on the spatial relationship of measured reef fish to sampled habitats.

The application of a benthic habitat map also means sampling is associated with the spatial scale or minimum mapping unit used to generate the map. A spatial scale disparate from relevant ecosystem scales of fish-habitat linkages affects the ability of a sampling design to effectively sample a reef fish population. For instance, a minimum mapping unit (MMU) of 4000 m^2 (1 acre) misses all reef structures less than this area. These structures will be sampled if the entire map is available for sampling, but the ability to use the habitat map to effectively parse the fish community into strata is compromised. Further, it is possible habitat areas smaller than the MMU are not sampled and an important component of the habitat is under-represented or not represented at all in the sample.

Application of a Sample Frame

An alternative to the aforementioned sampling design is to place a uniform distribution of points or grids onto an area which corresponds to the sampled population. The population of points or grids serves as the sample frame from which a random sample can be selected. The sampling units can be parsed according to benthic habitat types in order to have a stratified random selection.

The ideal design will divide the target population into a set of $100m^2$ belt transects, but in many cases this is not possible or desired. First, the spatial accuracy of available benthic habitat maps rarely allow transects to be accurately allocated among discrete benthic habitats. Second, arriving at a precise location in the field is made difficult by inaccuracies in GPS systems, inexperienced boat drivers, and sea conditions. Third, the placement of rectangular transects onto an area such that they are mutually-exclusive and exhaustively over the target population requires a systematic placement procedure. This procedure may cause sampling errors derived from systematic patterns in the benthic habitats of reef systems. For instance, systematically distributed transects which run perpendicular to a linear reef may provide a different representation of a fish community than one with transects which run parallel to the reef due to measurement biases.

A more realistic approach is to divide the target population into a set of sample units which are larger than the spatial units used to make measurements (i.e. the belt transects). A common technique is to divide the target population into a grid composed of square elements, where each element serves as a sampling unit. Grid centroids can be used as starting points for belt transects. To ensure unbiased samples, transects can radiate from the centroid at random bearings. The corners of each grid cell would not be sampled, but since a random set grids and bearings are used for the sample, this should not incorporate any sampling bias.

The advantages to using a sampling frame are that “spheres of influence” can be incorporated into estimates, sampling weights are not estimated, and sampling unit stratum designations can be updated. As stated earlier, a drawback of not having a sampling frame is that sampling units designated as a particular strata cannot be updated if new information shows otherwise.

The sphere of influence of benthic habitats on fish species can be implicitly incorporated into a sampling design by organizing sampling units according to benthic habitats over a broader spatial scale. For instance, a small patch reef surrounded by sand will likely influence the fish community on adjacent sand causing the area of the patch reef to be a poor estimator of the fish population. A stratification method which incorporates this sphere of influence will likely be a better estimator of fish population and community statistics than one which does not. One method of incorporating the sphere of influence is to stratify points according to benthic habitats at a spatial scale relevant to species-habitat interactions. By altering this spatial scale a user can refine survey estimates to more accurately represent the reef fish community and populations. Alternatively, if the sphere of influence is improperly estimated the user can incorporate significant biases into estimates.

A problem with identifying the sphere of influence is that it is species specific and the Branch surveys multiple species simultaneously. The sphere of influence of a reef for a stationary damselfish (e.g., *Stegastes partitus*) will be much smaller than for a mobile snapper (*Ocyrus chrysurus*). Some decision rule is required to balance these differences. For the Biogeography Branch most decisions for sampling designs are based on obtaining accurate and precise data for rare species, since these species are generally more difficult to sample and important to natural resource managers. Other species or community metrics (e.g., species richness, overall abundance, species diversity) typically have lower coefficients of variation and are more likely to be adequately sampled independent of the chosen sampling design.

The determination of an appropriate spatial scale for sampling units is also important because it affects the spatial association between variables used to stratify sample units and the reef fish populations being sampled (i.e. sphere of influence), and affects counts of species and individuals. A sampling unit which is too small may count the same species and individuals found in an adjacent sampling unit due to fish mobility. A sampling unit which is too large may generate a sampling frame which does not reflect true spatial patterns among benthic habitats and negate the positive affects of stratification.

Additionally, the creation of a sample frame itself poses problems. The principle problems with generating a sample frame are related to the spatial distribution of sample units (origin, size, separation, boundaries) and stratum designations. The decisions needed for these parameters are not simple and poorly reasoned decisions can lead to biased reef fish population and community estimates. Take, for instance, the case where two habitat types are present in a study area. A decision to treat all sample units with a majority of habitat type A as belonging to Strata A will offer different samples than a decision to treat sample units with 20% of habitat type A as belonging to Strata A. The latter is desirable in circumstance where a hierarchical approach to habitat classification is warranted (i.e. coral-centric). The different samples and sampling weights from these alternative decisions will likely produce different reef fish population and community estimates and may lead to different decisions by managers.

A similar conundrum is presented at the boundaries of the target population or near areas which cannot be sampled. When half of a sample unit is too deep for diving, a decision must be made as to whether to include it in the sampled population or not. One option is to include the sample unit if the centroid of a sample unit falls within safe repetitive diving limits (i.e. 110ft). In shallow coral reef systems a more complicated problem arises due to emergent reefs or land. Sampling units may include areas on either side of the emergent reefs. A set of decisions are needed for sampling. Should the sampling unit be included in the sampled population? If it is, which side of the reef should be measured? These decisions should strive to eliminate sampling bias.

In designing a sample frame the sphere of influence and method of measurement should be evaluated to arrive at a suitable spatial scale for sample units. Based on available data, a spatial scale of 50 m (i.e. each point/centroid is 50 m apart from adjacent points/centroids in cardinal directions) would be ideal for the Biogeography Branch. This scale would minimize element area and thus be a more accurate representation of the real world and ensure 25 m long transects from adjacent sampling units do not overlap. A 25 m sampling unit was not feasible, because transects radiate from the centroid at a random bearing and may be directed towards each other. The 50 m spatial resolution of sampling units is also considered an adequate compromise for the spheres of influence of rare species.

The origin of the sample frame and spatial scale of sample units determines the position of all sample units. A consideration when deciding the origin of a sample frame should be the potential union of the frame with another administration's frame. The union can be assisted by associating sample frames with broad spatial coordinate systems, such as the Universal Transverse Mercator (UTM) grid.

The minimum mapping unit of the benthic habitat map used to characterize sampling units into strata was 4000 m^2 (1 acre). This spatial scale was larger than the spatial scale adopted for the sample frames – 2500 m^2. A drawback to having the minimum mapping unit for the benthic habitat map being larger than the spatial scale of the sample frame is that some sample units are incorrectly characterized. In reef fish communities, patch reefs are at risk of being missed or incorporated into other habitats if the MMU is large. Patch reefs can serve as important staging points for juvenile fish and if they are not classified appropriately, the associated reef fish community may be underrepresented or even missed in samples. The problem of missing or incorrectly classifying sampling units into strata will occur in all sampling designs considered, even the approach currently used by the Biogeography Branch; however, the application of a sampling frame allows sample units which were incorrectly identified to be modified if necessary.

It is expected that switching the current sampling design to one which uses a sample frame will not be conceptually or logistically challenging. Field methods, estimate computations, and logistical requirements will not change. Differences in fish population and community estimates are likely, and will be associated with spatial scales used.

Subsampling

Subsampling corresponds to the use of subsamples to garner information from larger sampling units. If subsamples or secondary sampling units (SSUs) are chosen among larger primary sampling units (PSUs), it is known as two-stage sampling. The advantages of such a design are that it may be more cost effective to sample a population if the cost of sampling PSUs is high and SSUs within each PSU are similar.

A sample frame defined by areal PSUs is required to subsample a reef community. Sampling units organized in a grid, as considered in the previous section, are appropriate. Multiple belt transects can be placed within each grid. A simple technique is to have several transects radiating from the centroid of each grid in a sample.

In order to determine if subsampling is more efficient than a simple random sample, the variance among PSUs and SSUs is required. This information was gathered using a pilot study in La Parguera, Puerto Rico. A two-stage sampling design was used to collect data (Figure 2). The study area was first divided into 2500 m^2 (50m X 50m) primary sampling units (PSUs) such that all PSUs formed an exhaustive grid. Then a benthic habitat map designed by the Biogeography Branch (Kendall et al. 2001) was used to organize each PSU into either a hardbottom or softbottom strata. PSUs were allocated to the hardbottom stratum if more than 10% of area within a unit was hardbottom habitat (e.g. reef, bedrock, scattered coral in sand); the remaining elements were allocated to the softbottom stratum.

Five PSUs were randomly selected within each stratum and three 100m^2 belt transects were used to collect fish community and population measurements in each PSU (see Menza et al. 2006 for information on fish surveys within a belt-transect). Each belt-transect was positioned to radiate from the grid cell centroid at 130° intervals to minimize overlap (see figure 2). This design provided the required data to answer questions associated with subsampling.

The intraclass correlation coefficient (ICC) was computed for each stratum as suggested by Lohr (1999) to determine the efficiency of two-stage sampling. The ICC provides a measure of homogeneity within primary sampling units (PSU) and can be computed from ANOVA table quantities as

$$ICC = \frac{M}{M-1} \times \frac{SSW}{SSTO} \quad \text{eq 1}$$

where SSW = the sum of squares (SS) within PSUs, SSTO = the total SS and M = the number of secondary sampling units in each PSU. In addition, the adjusted R_a^2 was computed with

$$R_a^2 = 1 - \frac{MSW}{S^2} \quad \text{eq 2}$$

where MSW = the mean square between PSUs and S^2=the population variance. Similar to the ICC, R_a^2 will be high when PSU means are highly variable relative to the variation of SSUs within PSUs.

The ICC is positive and R_a^2 is close to 1 if measurements are similar within PSUs and little new information would be obtained by sampling more than one element in a PSU. The ICC is negative and R_a^2 is close -1 if measurements within PSUs are more variable than random and indicate a two-stage sampling design is best. Tables 1A and 1B show ICC and R_a^2 results for hardbottom and softbottom strata, respectively. Insufficient data in the softbottom stratum precluded calculation for 7 of the 13 investigated metrics.

The ICC and R_a^2 values for most investigated community and population metrics indicated little new information would be obtained by sampling more than one element in a PSU. The lowest ICC and R_a^2 values were found for abundance of all fish species, followed by *T. bifasciatum, A. bahianus, S. partitus,* and *X. martinicensis.* None of the grouper and snapper species investigated had ICCs less than 0 or low R_a^2; thus for these species simple random sampling is likely better than two-stage sampling.

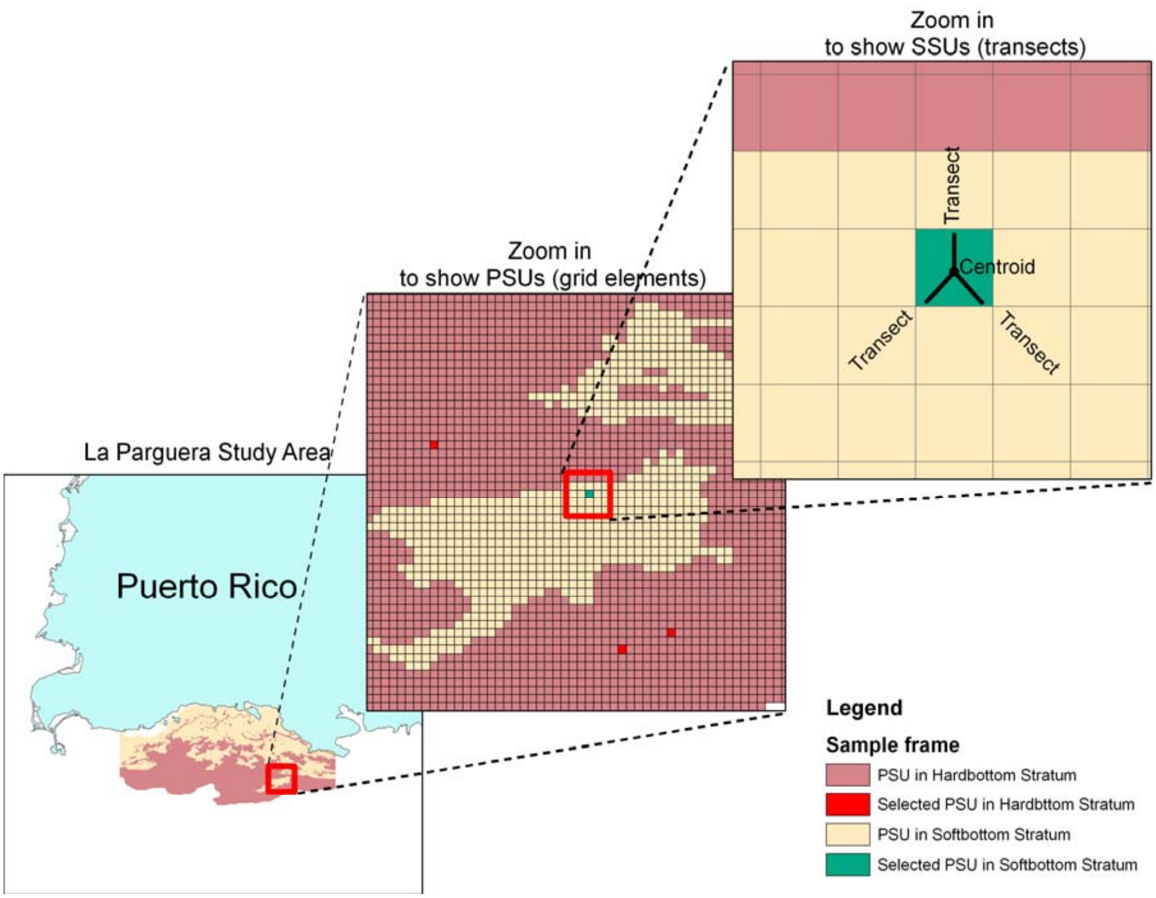


Figure 2: Schematic of pilot study used to acquire subsampling data in La Parguera, Puerto Rico.

Table 1A: Results of ICC calculations for reef fish community and population metrics surveyed within the hardbottom stratum.

Metric	SSB	SSW	SSTO	R_a^2	ICC
C. cruentatus	1.64	0.67	2.31	0.97	0.56
C. fulvus	1.90	3.33	5.23	0.92	0.04
O. chrysurus	76.93	13.83	90.77	0.98	0.77
S. iseri	439.76	237.17	676.92	0.96	0.47
X. martinicensis	2.10	4.67	6.77	0.91	-0.03
H. flavolineatum	17.17	19.83	38.00	0.93	0.22
S. aurofrenatum	261.67	444.33	706.00	0.92	0.06
A. bahianus	103.69	294.00	397.69	0.91	-0.11
T. bifasciatum	1667.86	9155.83	10823.69	0.99	-0.27
S. partitus	1578.36	4039.33	5617.69	0.91	-0.08
Richness	128.74	50.33	179.08	0.96	0.58
Abundance	3449.85	41623.83	45073.69	0.88	-0.39
Diversity	1.55	0.15	2.71	0.95	0.92

Table 1B: Results of ICC calculations for reef fish community and population metrics surveyed within the softbottom stratum.

Metric	SSB	SSW	SSTO	R_a^2	ICC
C. cruentatus			Insufficient data		
C. fulvus			Insufficient data		
O. chrysurus			Insufficient data		
S. iseri			Insufficient data		
X. martinicensis	3497.67	4800.67	8297.73	0.94	0.13
H. flavolineatum			Insufficient data		
S. aurofrenatum			Insufficient data		
A. bahianus	5.07	16.67	21.73	0.92	-0.15
T. bifasciatum			Insufficient data		
S. partitus	4.27	10.67	14.93	0.93	-0.07
Richness	139.07	142.67	281.73	0.95	0.24
Abundance	10550.93	6910.00	17460.93	0.96	0.41
Diversity	2.81	4.16	6.98	0.94	0.11

A second analysis of subsampling was carried out by examining the optimal allocation of time devoted to sampling PSUs and SSUs. For this analysis different sampling scenarios with different subsampling amounts were evaluated according to their affect on projected variance. Sampling scenarios differed by relative costs for PSUs and SSUs, and subsample size. The optimal number of subsamples was determined using a cost function given by

$$C = c_1 n + c_2 nm \qquad \text{eq 3}$$

where C = total cost, c_1=cost per PSU and c_2= cost per SSU. The optimal number of subsamples, m*, was computed using

$$\hat{m}_{opt} = \sqrt{\frac{s_2^2}{s_1^2 - s_2^2/M}} \times \sqrt{\frac{c_1}{c_2}} \qquad \text{eq 4}$$

where *M*=the number of secondary sampling units in primary sampling units of equal size, s_2^2 = variance among SSUs within PSUs, and s_2^2 = variance among PSUs, as given by Cochran (1977).

Time was taken as the principle determinant of sampling cost. Time limits the number of samples which can be taken during a field mission and affects monetary resources required to collect all necessary information in a survey (e.g., dive boat rental, SCUBA tank rental, salary). The cost per PSU was given by the average time required to travel among PSUs. The cost per SSU was the sum of time required to enter the water, take measurements and return to the vessel. Table 2 presents sampling costs associated with monitoring in Puerto Rico and were determined from 5 years of experience.

Table 3A and 3B indicate the optimal number of samples for different cost scenarios (i.e. C1/C2). Table 3A shows that for most metrics investigated on hardbottom habitat only a single SSU per PSU is sufficient for sampling when sampling costs are approximately average (c1/c2=0.29 from Table 2). The exceptions to this rule include abundance and *A. bahianus* which are sampled most efficiently when the number of SSUs sampled in each PSU is more than one; in the case of abundance it is close to 5. The paucity of data collected over softbottom habitat precludes analysis for many metrics; however, for those that were examined subsampling with more than 1 SSU per PSU was beneficial.

It is important to note that the results provide in Tables 3A and 3B are derived from scenarios applicable to current study areas monitored by the Biogeography Branch. Sampling costs are likely to be different in other study areas. If the relative cost of sampling PSUs to SSUs is much higher than 1.00, the optimal subsample size increases for most metrics to two or three. Ault et al. (2002) sample reef fish over a much larger area than typically examined by the Biogeography Branch and thus the relative cost of sampling PSUs to SSUs will likely be greater than 1.00. They have found that subsampling with two SSUs per PSU is optimal for their sampling design.

Table 2: Sampling times (costs in minutes) associated with sampling reef fish communities.

Cost Type	Minimum	Maximum	Average
PSU	5	20	10
SSU	25	40	35
C1/C2	0.125	0.8	0.29

Table 3A: Optimal number of subsamples for different sampling cost scenarios on hardbottom habitat. Shaded boxes indicate more than one subsample is optimal.

	c1/c2				
Metric	0.125	0.25	0.5	0.75	1.00
C. cruentatus	0.23	0.33	0.46	0.57	0.66
C. fulvus	0.55	0.78	1.10	1.35	1.56
O. chrysurus	0.18	0.25	0.36	0.44	0.50
S. iseri	0.31	0.43	0.61	0.75	0.87
X. martinicensis	0.60	0.85	1.21	1.48	1.71
H. flavolineatum	0.39	0.56	0.79	0.96	1.11
S. aurofrenatum	0.53	0.76	1.07	1.31	1.51
A. bahianus	0.73	1.04	1.46	1.79	2.07
T. bifasciatum	0.28	0.39	0.55	0.68	0.78
S. partitus	0.67	0.95	1.35	1.65	1.90
Richness	0.23	0.32	0.45	0.56	0.64
Abundance	2.78	3.94	5.57	6.82	7.87
Diversity	0.31	0.44	0.63	0.77	0.88

Table 3B: Optimal number of subsamples for different sampling cost scenarios on softbottom habitat. Shaded boxes indicate more than one subsample is optimal.

	c1/c2				
Metric	0.125	0.25	0.5	0.75	1.00
C. cruentatus					
C. fulvus					
O. chrysurus					
S. iseri					
X. martinicensis	0.47	0.66	0.94	1.15	1.33
H. flavolineatum					
S. aurofrenatum					
A. bahianus	0.77	1.08	1.53	1.88	2.17
T. bifasciatum					
S. partitus	0.65	0.92	1.31	1.60	1.85
Richness	1.74	2.46	3.48	4.27	4.92
Abundance	1.20	1.69	2.39	2.93	3.39
Diversity	2.76	3.90	5.51	6.75	7.79

Conclusions

The application of a sample frame would be beneficial in places where a validated or accurate benthic habitat map is not available or when the sampling fraction is likely to be high. There are considerable costs to implementing a sampling frame, namely the requirement of several complex decision rules. Since the Biogeography Branch samples in regions where a suitable benthic habitat map exists and the sampling fraction is very low, a sample frame is not considered essential.

In the event a sample frame is used in either a uniform distribution of points or an exhaustive grid, the optimal spatial scale would be 50 m. At this scale, 25 m long transects radiating from adjacent sampling units would not overlap and the loss of spatial resolution needed to characterize sample units using a benthic habitat map would be minimized.

In general, it is optimal to sample only one SSU per PSU. This decision is based on the fact that most investigated rare species are sampled best with one SSU. These species, which are considered representative of other rare species and include many fishery species important to managers, generally have high coefficients of variations and high sample size requirements. Most decisions regarding reef fish sampling methods are based on how to minimize the sample size requirements associated with these species. Other species or community metrics (e.g., species richness, overall abundance, species diversity) typically have lower coefficients of variation and are more likely to be adequately sampled independent of the chosen sampling design. By this rationale we expect species with optimal subsample sizes greater than one to be adequately sampled even if one SSU per PSU is sampled.

Acknowledgements

Thanks to Jerry Ault and Steve Smith at the University of Miami Rosentiel School of Marine and Atmospheric Science for providing statistical advice and helping to collect data in the field.

References

Ault, J.S., S.G. Smith, J. Luo, G.A. Meester, J.A Bohnsack., and S.L. Miller. 2002. Baseline multispecies coral reef fish stock assessments for the Dry Tortugas. NOAA Technical Memorandum NMFS-SEFSC-487. 117 pp.

Cochran, W.G. 1977. Sampling Techniques, third edition. John Wiley & Sons, Inc., New York, 428 pp.

Kendall, M.S., M.E. Monaco, K.R. Buja, J.D. Christensen, C.R. Kruer, and M. Finkbeiner and R.A. Warner. 2001. (On-line). Methods Used to Map the Benthic Habitats of Puerto Rico and the U.S. Virgin Islands URL: http://biogeo.nos.noaa.gov/projects/mapping/caribbean/ startup.htm. Also available on U.S. National Oceanic and Atmospheric Administration. National Ocean Service, National Centers for Coastal Ocean Science Biogeography Program. 2001. (CD-ROM). Benthic Habitats of Puerto Rico and the U.S. Virgin Islands. Silver Spring, MD: National Oceanic and Atmospheric Administration.

Lohr, S.L. 1999. Sampling: Design and Analysis. Duxbury Press, Pacific Grove, CA, 494 pp.

Menza, C., J. Ault, J. Beets, J. Bohnsack, C. Caldow, J. Christensen, A. Friedlander, C. Jeffrey, M. Kendall, J. Luo, M. Monaco, S. Smith and K. Woody. 2006. A Guide to Monitoring Reef Fish in the National Park Service's South Florida / Caribbean Network. NOAA Technical Memorandum NOS NCCOS 39. 166 pp.

United States Department of Commerce

Donald L. Evans
Secretary

National Oceanic and Atmospheric Administration

Vice Admiral Conrad C. Lautenbacher, Jr. USN (Ret.)
Under Secretary

National Ocean Service

Jack Dunnigan
Assistant Administrator