CREP REEF FISH DATA – OVERVIEW OF DATA GATHERING

Pacific Reef Assessment and Monitoring Program (RAMP)

NOAA Coral Reef Ecosystems Program (CREP) surveys coral reefs at a total of ~40 islands and atolls in the US and US-affiliated regions in the Pacific using consistent methods and survey design. The program, known as Pacific RAMP, is part of NOAA's national coral reef monitoring program (NCRMP) (NCRMP, n.d.).

RAMP surveys were initially conducted on a 2-year cycle (each region visited every two years), with a small team and relatively few surveys in each survey round. Between 2007 and 2009, the RAMP reef fish component was revamped, with the adoption of a stratified random sampling design within an explicit survey domain (<30 m hardbottom), increased survey replication, and implementation of a new survey methodology. Since 2012, RAMP has moved to a 3-year cycle.

Reef fish surveys - sampling design and monitoring domain

CRED maintains GIS bathymetric and bottom composition maps for each island. The monitoring target habitat (<30 m hardbottom) is stratified by reef zone (backreef, forereef, lagoon) and depth bin: shallow (0–6 m), mid (6–18 m), and deep (18–30 m). At some larger islands, the island is further stratified by section of coastline or management status. For example, Guam reef areas are classified as being "Marine Preserve"; "Guam Open East"; and "Guam West".

The locations of survey sites are selected using a formal randomization process, with a new set of sites selected prior to each survey mission. Typically, each island is visited for 3-5 days during a RAMP cruise, which allows for a total of ~30-50 surveys to be conducted. Exceptions include the small islands of Sarigan, Guguan, and Alamagan, each of which is visited for one day. For routine analyses those three adjacent islands are pooled together into a single unit ("SGA").

Survey methods

Two types of data are collected at each survey site: fish counts and benthic cover estimates.

Counting and sizing reef fishes using the stationary point count method (SPC)

Each site is surveyed by a pair of divers conducting simultaneous counts in adjacent, visually estimated 15-m-diameter cylindrical plots extending from the substrate to the limits of vertical visibility (Figure 1). Prior to beginning each SPC-pair, a 30-m line is laid across the substratum. Markings at 7.5 m, 15 m and 22.5 m enable survey divers to locate the midpoint (7.5 m or 22.5 m) and two edges (0 m and 15 m; or 15 m and 30 m) of their survey plots. Each count consists of two components: (i) a 5-min **species enumeration period** in which the diver records the taxa of all species observed within their cylinder; followed by (ii) the **tallying portion** of the count, in which divers systematically work through their species lists, recording the number and estimated size (total length, TL, to the nearest cm) of each individual fish. The tallying portion is conducted as a series of rapid visual sweeps of the plot, with one species/grouping counted per sweep. To the extent possible, divers remain at the center of their cylinders throughout most of the count; but small, generally site-attached and semi-cryptic species are left to the end of the tally period, at which time the observer swims through their plot counting those species. Surveys are not conducted if horizontal visibility is < 7.5 m.

Data are recorded as one of five different 'observation types'. The majority of records – those where a species is observed during the enumeration period and where individuals of that species are present in the cylinder at the time of the tallying portion for that taxa – are recorded as 'instantaneous' observations (OBS_TYPE="I"). When a species is observed during the enumeration period but is not present during the instantaneous sweep for that taxa, divers record size and number present in the cylinder when it was first observed during the enumeration period and mark the data record as 'non-instantaneous' (OBS_TYPE="N"). Since 2012, we also record three other types of observations: when a species is first observed in the cylinder between 5 and 10 minutes into the survey (i.e. in the first 5 minutes of the tallying portion), the diver conducts a rapid visual sweep of their cylinder for that species and records number and size as 'five-to-ten' (OBS_TYPE="F"). Data on species first observed inside the cylinder any time after that, up to 30 minutes into the survey, are recorded as 'ten-to-thirty' (OBS_TYPE="T"). The presence of other species of interest in the general vicinity of the survey, and seen at any time throughout the survey period are recorded as 'present' (OBS_TYPE="P"). 'Instantaneous' data therefore come from a 'closed count', i.e., represents density of fishes within a defined area at one point in time. Other data-types allows us to integrate data over longer time periods, i.e. to count fishes that are present in or move across the cylinder at some point through the course of the survey. That integrated data allows us to gather systematic data on relative abundance and size distribution of relatively rare or skittish and/or more mobile species. Depending on the question of interest, we filter the data by its observation type. By default, we pool 'I' and 'N' data for routine reporting, because that allows us to have the longest comparable data set – we have SPC data from 2007, but only recorded observation type from 2010 onwards.

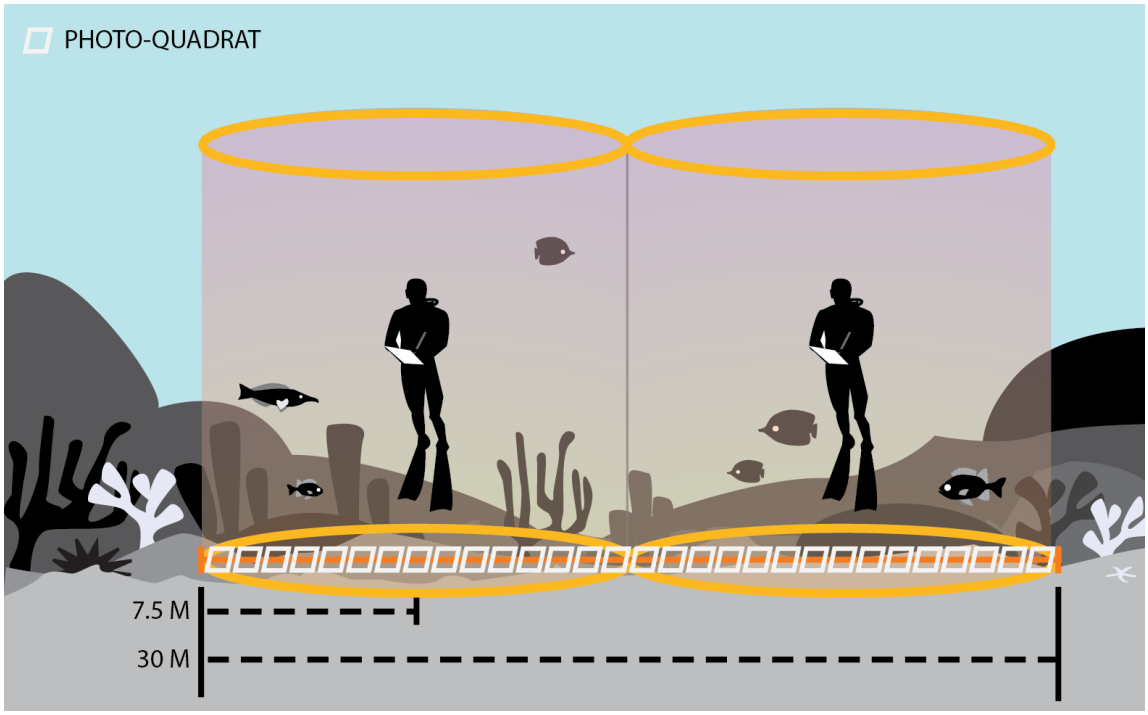


Figure 1. Side view of the stationary point count method. Surveys typically consist of one SPC-pair but some sites are surveyed by means of 2 SPC-pairs. During counts, dive partners count and size fishes within their cylinders measuring 7.5 m in radius. Once the fish survey is complete, divers estimate benthic habitat composition and conduct a benthic photo-transect spanning the two cylinders.

Assessing benthic habitat characteristics

Two complementary methods are used to assess benthic composition within survey cylinders: (i) divers conduct **rapid visual assessments** of percentage cover of major functional categories (e.g.: 'coral', 'macroalgae', sand, other); and (ii) divers conduct a **photo transect** through the middle of their SPC-pair. Analysis of photo-quadrat provides taxonomically finer-scale information, but requires post-survey processing and hence there is a lag before those data are available. Because the visually-estimated benthic data from each survey are immediately available and generally reliable, especially for hard coral cover (McCoy, Williams, & Heenan, 2015), we generally use those data to produce timely reports, but supplement our database with the photo-transect data once it has been processed.

Divers also record reef habitat complexity by visually estimating the percentage of their cylinder that falls into the following levels of vertical relief: <0.20 m from substrate, 0.20–0.50 m, 0.50–1.00 m, 1.00–1.50 m, and >1.50 m. Divers also record the abundance of free (e.g., *Tripneustes, Heterocentrotus, Diadema* and *Echinothrix* spp.*)* and boring (e.g., *Echinometra* and *Echinostrephus* spp.) urchins within their cylinder. Finally, divers also classify habitat type using the geomorphological categories identified by the NOAA National Ocean Service Biogeography Branch: aggregate reef, individual patch reef, aggregated patch reefs, spur and groove, pavement, pavement with sand channels, pavement with patch reefs, reef rubble, sand with scattered coral/rock and rock / boulder (Kendall & Poti, 2011).

A more complete description of the survey approach is given in the CREP fish survey standard operating procedure document (Ayotte, McCoy, Heenan, Williams, & Zamzow, 2015).

Synthesizing data

Generating reef fish biomass estimates

Survey data are generally synthesized to generate estimates of biomass, i.e. weight of fish per unit area surveyed. Fish weights are derived from lengths using taxon-specific length-weight conversion parameters. To generate biomass (as g/m^2), total fish weight is summed for a divers' count and then divided by the area surveyed (176.7 m^2 per SPC cylinder). Data from the two adjacent SPC cylinders are averaged to create a mean estimate for the SPC-pair. In some cases, a site is surveyed by means of 2 SPC-pairs. When that happens, data are averaged within the SPC-pairs, and then between SPC-pairs to generate site-level estimates. Fish data can be pooled in a variety of ways, e.g. per species, family or trophic group. The NCRMP default is to pool data by consumer group, i.e. as 'PRIMARY' (herbivores and detritivores); 'SECONDARY' (invertivores and omnivores); PLANKTIVORES; and PISCIVORES. Consumer group classifications are listed in CREP data reports (Heenan et al., 2014, 2015).

Generating benthic habitat estimates

The replicate level benthic data is similarly averaged to create site level benthic estimates, including benthic cover, substrate complexity and maximum height. Two types of substrate complexity information are gathered by divers: between 2010 and 2011, divers classified complexity on a 5-point scale from 1 (very low) to 5 (very high). From 2012 onwards, divers record the complexity of the substrate by estimating the proportion of their survey area in

different height bins (0-20cm from reef plane; 20-50cm; 50-100cm; 100-150cm; >150cm). Divers also estimate the maximum height from the plane of hard substrate inside their cylinders. Together that information is used to derive estimates of mean substrate height inside the cylinder and mean and standard deviation of the difference in height from that cylinder average.

Pooling site level data to per-strata and per-island levels

Summary statistics (e.g., mean and variances) of survey quantities, e.g., biomass, are calculated from the surveys within each stratum. To pool those up into larger units (e.g. 'island'), CREP typically uses the approach described by (Smith et al., 2011) which weights each stratum by its relative size (i.e., if a stratum is 50% of the total area in an island then is weighting factor will be 0.5, and total of all weighting factors sums to 1). Per strata mean and variance values are aggregated to a higher level (e.g., to island scale) using the formulas below:

(1) pooled mean biomass (*X)* across S strata: $X = \sum_1^S (X_i * w_i)$ and;
(2) pooled variance of mean biomass (*VAR)* across S strata: $VAR = \sum_1^S (VAR_i * w_i^2)$

where X_i is the estimate of mean biomass within stratum i, VAR_i is the estimated variance of X_i and w_i is the stratum-weighting factor.

References

Ayotte, P., McCoy, K., Heenan, A., Williams, I., & Zamzow, J. (2015). *Coral Reef Ecosystem Division standard operating procedures: data collection for Rapid Ecological Assessment fish surveys.* Ayotte P., McCoy K., Heenan A., Williams I., Zamzow J. 2015. Coral Reef Ecosystem Division standard operating procedures: data collection for Rapid Ecological Assessment fish surveys. Retrieved from http://www.pifsc.noaa.gov/library/pubs/admin/PIFSC_Admin_Rep_15-07.pdf

Heenan, A., Ayotte, P., Gray, A., Lino, K., McCoy, K., Zamzow, J., & Williams, I. D. (2014). *Pacific Reef Assessment and Monitoring Program - Data Report - Ecological Monitoring 2012-2013 - reef fishes and benthic habitats of the main Hawaiian Islands , American Samoa , and Pacific Remote Island Areas*.

Heenan, A., McCoy, K. S., Asher, J., Ayotte, P., Gorospe, K., Gray, A., … Williams, I. D. (2015). *Pacific Reef Assessment and Monitoring Program - Data Report - Ecological Monitoring 2014 - stationary point count surveys of the Northwestern Hawaiian Islands, Mariana Islands, and Wake Atoll*.

Kendall, M., & Poti, M. (2011). *A biogeographic assessment of the Samoan Archipelago*. *NOAA Tech Memo NOS NCCOS*. Silver Spring, MD.

McCoy, K., Williams, I., & Heenan, A. (2015). *A Comparison of Rapid Visual Assessments and Photo- Quadrat Analyses to Monitor Coral Reef Habitats Data Report*.

NCRMP. (n.d.). *NOAA Coral Reef Conservation Program. National Coral Reef Monitoring Program*. Retrieved from http://docs.lib.noaa.gov/noaa_documents/NOS/CRCP/noaa_crcp_national_coral_reef_monitoring_plan_2014.pdf

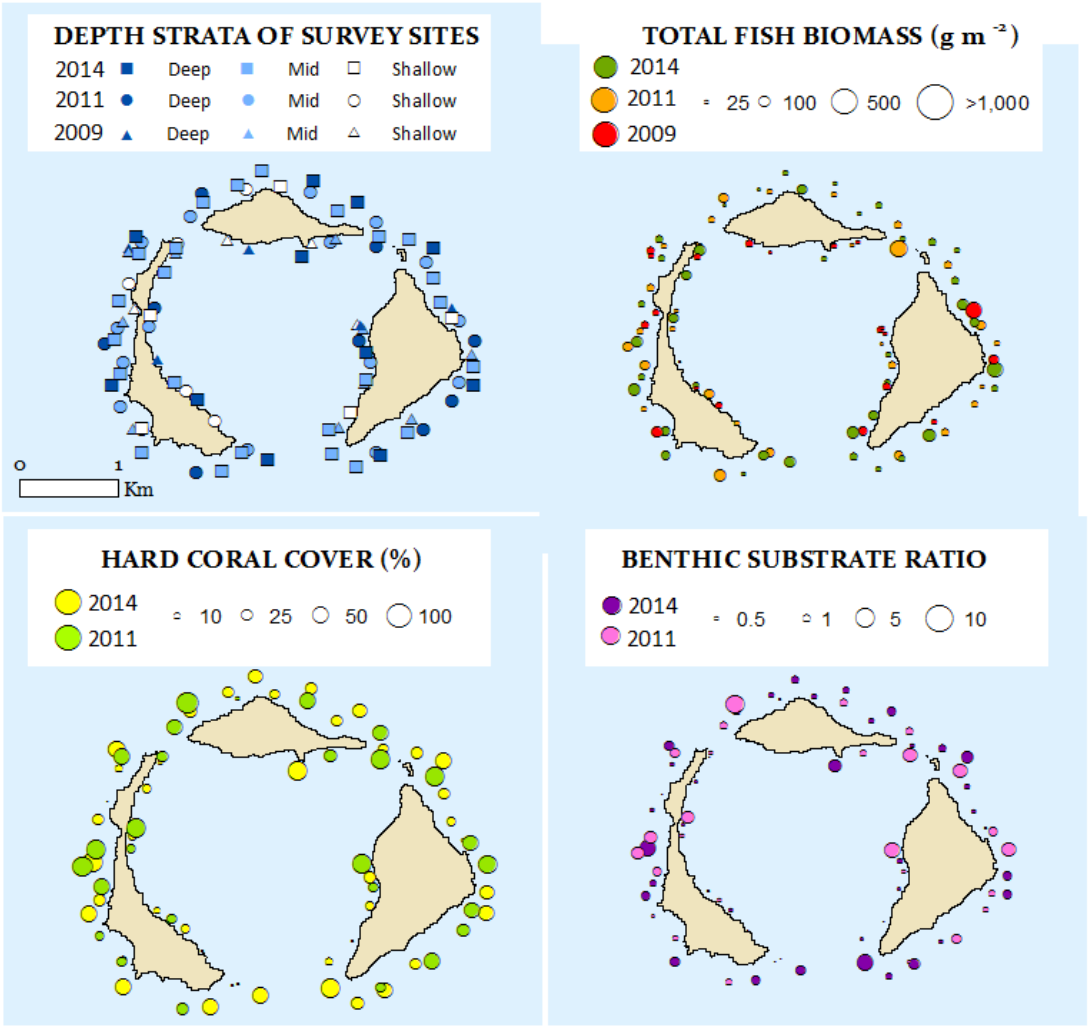
Smith, S. G., Ault, J. S., Bohnsack, J. A., Harper, D. E., Luo, J. G., & McClellan, D. B. (2011). Multispecies survey design for assessing reef-fish stocks, spatially explicit management performance, and ecosystem condition. *Fisheries Research*, *109*(1), 25–41. http://doi.org/10.1016/j.fishres.2011.01.012

Reef fish survey data – standard site level output field descriptions

DATA FIELD	DESCRIPTION
SITEVISITID	A unique identifier for a site-survey
METHOD	Fish survey method used. Is always "nSPC" for data gathered using the current CRED stationary-point-count survey method

OBS_YEAR	The year in which the survey was conducted
REGION	Survey region – generally a spatially meaningful grouping, e.g. MHI are the populated Main Hawaiian Islands, NWHI are the unpopulated Northwestern Hawaiian Islands. For Mariana Archipelago islands, islands are grouped into “S.MARIAN” (the populated islands from Guam to Saipan), and “N.MARIAN” (the unpopulated or very lightly populated islands from Sarigan to Farallon de Pajaros)
ISLAND	The island where survey was conducted
SITE	A unique survey site code. First three letters are an island identifier, last 4 digits are a numeric value to make each SITE code unique to a particular survey location.
DATE_	Date of the survey
REEF_ZONE	“Forereef", "Backreef' or "Lagoon"
HABITAT CODE	AGR (aggregate reef), APR (aggregate patch reef), APS (aggregate patch reefs), MIX (mixed habitat), PAV (pavement), PPR (pavement with patch reefs), PSC (pavement with sand channels), ROB (rock boulder), RRB (reef rubble), SAG (spur and groove), SCR (scattered coral/rock), UNK (unknown), WAL (Wall). Habitat code is intended to represent the general area in which the survey is conducted, rather than the exact area of the survey cylinders. Nominally, this is based on thinking of the survey itself being at the center of a 50m*50m cell and habitat code is for the entire cell.
DEPTH_BIN	Depth strata of the survey site: either Shallow (<6m); Mid (6-18m); or Deep (18-30m).
LATITUDE & LONGITUDE	Survey site location, taken from the diving platform (small boat) above the divers.
DEPTH (m)	Mean site depth (in meters), recorded by divers during the survey.
HARD_CORAL MA, CCA, SAND, and OTHER	Benthic cover, visually estimated by divers during surveys. Categories depend on survey location and period, but include hard coral, upright macroalgae, crustose coralline algae, sand, and other.
ComplexityValue	Complexity estimated on the 5 point scale (2010-2011 surveys only)
MEAN_SH	Complexity metric (mean height of substrate from the reef plane)
SD_SH_DIFF	A measure of variation in complexity (standard deviation of the difference in substrate height inside the cylinders from the substrate mean height)
MAX_HEIGHT	Height from reef plane of highest hard substrate inside each diver’s survey cylinder

PISCIVORE PLANKTIVORE PRIMARY SECONDARY	Fish biomass (g/m^2) by 'consumer group'. PRIMARY are herbivores and detritivores; SECONDARY includes omnivores, and invertivores; PISCIVORES are fishes that feed primarily on fishes.
TotFish	Total biomass (g/m^2) of all fishes combined.



Example figure: Maug Island site survey data 2009, 2011, and 2014 identified by depth strata (top left). Total fish biomass recorded at each site per year (top right). Hard coral cover (%) assessed by rapid visual assessment (bottom left). Benthic substrate ratio (hard coral plus encrusting algae / turf and macroalgae) (bottom right). This ratio indicates the balance between the benthic components that contribute to reef accretion (coral and crustose coralline algae) compared to fleshy macroalgae and turf algae that compete for space on the reef

Timeline of methodological changes

2000-2006	Survey conducted by means of belt transects at haphazardly located semi-fixed sites (lat/long used to approximately relocate sites), mostly in mid-depth forereef.
2007-2009	Partner agency, Papahānaumokuākea Marine National Monument (PMNM) develop and trial a point count fish survey method (METHOD="nSPC") and true stratified-random survey design approach for surveys in the NWHI. CRED begins calibration period, in which both belt and nSPC surveys are conducted at existing sites in all locations other than MHI, where belt surveys remain the single method used.
2010	(i) CRED fully switches to nSPC and stratified random design. (ii) Divers begin to record OBS_TYPE – as either 'I' (instantaneous) or 'N' (non-instantaneous, but observed during the 5 minute species enumeration period). Fishes that enter the survey area the completion of the species-enumeration period are not recorded. Data from earlier surveys is recorded as OBS_TYPE 'U' (unknown, but it was either 'I' or 'N'). (iii) Divers begin to take visual estimates of benthos and structural complexity on a 5-point scale.
Start of 2012	(i) Divers begin to record urchin (free and boring) abundance. (ii) Substrate complexity measures changed to current methods based on MAX_HEIGHT and proportion of substrate in various height bins, based on height from reef plane.
mid-2012 to present	Divers add additional observation types to cover fishes that first enter the survey cylinder after the completion of the species enumeration period. Those are either OBS_TYPE='F' (species that first enter the cylinder in the first 5 minutes after completion of the species enumeration period, i.e., 5-10 minutes into the count), and OBS_TYPE 'T' (fishes of species that first enter the survey are 10-30 minutes into the survey). Additionally OBS_TYPE 'P' is added for fishes of interest that are observed in the vicinity of the survey area in the course of the survey dive, but are not otherwise recorded.