Please provide the following information, and submit to the NOAA DM Plan Repository.

Reference to Master DM Plan (if applicable)

As stated in Section IV, Requirement 1.3, DM Plans may be hierarchical. If this DM Plan inherits provisions from a higher-level DM Plan already submitted to the Repository, then this more-specific Plan only needs to provide information that differs from what was provided in the Master DM Plan.

URL of higher-level DM Plan (if any) as submitted to DM Plan Repository:

1. General Description of Data to be Managed

1.1. Name of the Data, data collection Project, or data-producing Program:
Benthic Cover from Automated Annotation of Benthic Images Collected at Coral Reef Sites in the Pacific Remote Island Areas and American Samoa in 2018

1.2. Summary description of the data:
The coral reef benthic community data described here result from the automated annotation (classification) of benthic images collected during photoquadrat surveys conducted by the NOAA Pacific Islands Fisheries Science Center (PIFSC), Ecosystem Sciences Division (ESD, formerly the Coral Reef Ecosystem Division) as part of NOAA's ongoing National Coral Reef Monitoring Program (NCRMP). SCUBA divers conducted benthic photoquadrat surveys in coral reef habitats according to protocols established by ESD and NCRMP during the ESD-led NCRMP mission to the islands and atolls of the Pacific Remote Island Areas (PRIA) and American Samoa from June 8 to August 11, 2018. Still photographs were collected with a high-resolution digital camera mounted on a pole to document the benthic community composition at predetermined points along transects at stratified random sites surveyed only once as part of Rapid Ecological Assessment (REA) surveys for corals and fish and permanent sites established by ESD and resurveyed every ~3 years for climate change monitoring. Overall, 30 photoquadrat images were collected at each survey site.

The benthic habitat images were quantitatively analyzed using the web-based, machine-learning, image annotation tool, CoralNet (https://coralnet.ucsd.edu; Beijbom et al. 2015). Ten points were randomly overlaid on each image and the machine-learning algorithm "robot" identified the organism or type of substrate beneath, with 300 annotations (points) generated per site. Benthic elements falling under each point were identified to functional group (Tier 1: hard coral, soft coral, sessile invertebrate, macroalgae, crustose coralline algae, and turf algae) for coral, algae, invertebrates, and other taxa following Lozada-Misa et al. (2017). These benthic data can ultimately be used to produce estimates of community composition, relative abundance (percentage of benthic cover), and frequency of occurrence.

1.3. Is this a one-time data collection, or an ongoing series of measurements?
One-time data collection

1.4. Actual or planned temporal coverage of the data:
2018-06-08 to 2018-06-15, 2018-06-19 to 2018-07-18, 2018-07-28 to 2018-08-11

1.5. Actual or planned geographic coverage of the data:
W: -176.626077, E: -159.971695, N: 6.451465, S: -0.382607
Phoenix (Baker and Howland) and Line Islands (Jarvis, Kingman, and Palmyra) of the Pacific Remote Island Areas.

W: -171.092425, E: -168.141129, N: -11.045273, S: -14.558022
American Samoa, including Tutuila, Ta`u, Ofu and Ologega, and Swains islands, and Rose Atoll.

1.6. Type(s) of data:
(e.g., digital numeric data, imagery, photographs, video, audio, database, tabular data, etc.)
Table (digital)

1.7. Data collection method(s):
(e.g., satellite, airplane, unmanned aerial system, radar, weather station, moored buoy, research vessel, autonomous underwater vehicle, animal tagging, manual surveys, enforcement activities, numerical model, etc.)

1.8. If data are from a NOAA Observing System of Record, indicate name of system:

1.8.1. If data are from another observing system, please specify:

2. Point of Contact for this Data Management Plan (author or maintainer)

2.1. Name:
Annette M DesRochers

2.2. Title:
Metadata Contact

2.3. Affiliation or facility:

2.4. E-mail address:
annette.desrochers@noaa.gov

2.5. Phone number:
(808)725-5461

3. Responsible Party for Data Management
Program Managers, or their designee, shall be responsible for assuring the proper management of the data produced by their Program. Please indicate the responsible party below.

3.1. Name:
Courtney S Couch

3.2. Title:
Data Steward

4. Resources
Programs must identify resources within their own budget for managing the data they produce.

4.1. Have resources for management of these data been identified?
Yes

4.2. Approximate percentage of the budget for these data devoted to data management (specify percentage or "unknown"):
Unknown

5. Data Lineage and Quality
NOAA has issued Information Quality Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information which it disseminates.

5.1. Processing workflow of the data from collection or acquisition to making it publicly accessible
(describe or provide URL of description):

Lineage Statement:

Protocol describing the automated annotation of benthic images to generate benthic community data. Still photographs (30 per survey site) collected during photoquadrat surveys at random and permanent sites in coral reef habitats by the PIFSC Ecosystem Sciences Division were later annotated by a human-trained robot in CoralNet.

Process Steps:

- RANDOM SURVEY SITES: A stratified random sampling (StRS) design was employed to select the random sites surveyed for reef fish and corals. The survey domain encompassed the majority of the mapped area of reef and hard bottom habitats in the 0-30 m depth range. The stratification scheme included island, reef zone, and depth. Sampling effort was allocated based on strata area and sites were randomly located within strata. The StRS design effectively reduces estimate variance through stratification using environmental covariates and by sampling more sites rather than sampling more transects at a site. Therefore, site-level estimates and site-to-site comparisons should be used with caution.   Photoquadrat surveys conducted during surveys for reef fish (Ayotte et al. 2015) occur along one 30-m transect that spans the length of two stationary point count (SPC) cylinders used to assess fish abundance (30 images total, Rep A). Benthic photos are taken at 1-m intervals along the right hand side of the 30-m transect line.   Photoquadrat surveys conducted during coral demographic surveys (Swanson et al. 2018) occur along two, independent 18-m transects. Photos of the benthic substrate are taken at 1-m intervals starting at meter one, along the left hand side of each 18-m transect (15 images/transect, 30 images total, Rep A and B).

- PERMANENT SURVEY SITES: Permanent sites were chosen in hard-bottom habitat at ~15-m depths, and a subset of the permanent sites called climate stations were established at north, south, east, and west points around each of the islands and atolls. A minimal suite of climate monitoring activities are conducted at permanent sites, whereas a full complement of activities are conducted at climate stations.   Divers estimated and delineated each site's rectangular perimeter by temporarily placing measuring tapes with 1-m markers starting from a permanently installed reference stake on the reef. Along the nearshore 10-m side of the survey site and the downslope 5-m side, the measuring tapes marked every meter of the L-shaped 15-m transect used for photoquadrat documentation. The divers photographed the reef at 1-m intervals on both sides of the 15-m transect, generating 30 photographs per survey site.

- IMAGE PROCESSING: Photoquadrat images are organized by site, color-corrected (if necessary), and quality controlled to remove non-photoquadrat/poor quality images. An optical validation script is used to guarantee that all photoquadrat sites have the respective correct metadata.   Post image analysis quality control steps include spot checks of the machine-generated output data for completeness and adequate analysis of optical data.

- AUTOMATED ANNOTATION OF BENTHIC IMAGES: Benthic habitat images were quantitatively analyzed using the web-based, machine-learning, image annotation tool, CoralNet (https://coralnet.ucsd.edu; Beijbom et al. 2015). Ten points were randomly overlaid on each image and the machine-learning algorithm "robot" identified the organism or type of substrate beneath, with 300 annotations (points) generated per site. Benthic elements falling under each point were identified to functional group (Tier 1: hard coral, soft coral, sessile invertebrate, macroalgae, crustose coralline algae, and turf algae) for coral, algae, invertebrates, and other taxa following Lozada-Misa et al. (2017).   The description of each functional group is included in the benthic image analysis classification scheme. (Citation: Lozada-Misa P., B. D. Schumacher, and B. Vargas-Ángel. 2017. Analysis of benthic survey images via CoralNet: a summary of standard operating procedures and guidelines. Pacific Islands Fish. Sci. Cent., Natl. Mar. Fish. Serv., NOAA, Honolulu, HI 96818-5007. Pacific Islands Fish. Sci. Cent. Admin. Rep. H-17-02, 175 p. doi:10.7289/V5/AR-PIFSC-H-17-02.)

5.1.1. If data at different stages of the workflow, or products derived from these data, are subject to a separate data management plan, provide reference to other plan:

5.2. Quality control procedures employed (describe or provide URL of description):
Machine-generated fully-automated benthic estimates of site-level coral cover (Tier 1) are highly comparable to those generated by human analysts (Pearson's r > 0.97, and with bias of 1% or less; Williams et al. 2019).

Pre- and post-automated benthic image analysis quality control procedures include spot

checks for:

1- Image quality integrity: focus, white balance, distance from bottom (too high or too low)

2- image inventory completeness (all images acquired for each island are included in the analysis)

3- Duplicate sites

4- Complete and coherent upload of imagery to CoralNet

5- Integrity and completeness of robot-generated image annotations

6. Data Documentation
The EDMC Data Documentation Procedural Directive requires that NOAA data be well documented, specifies the use of ISO 19115 and related standards for documentation of new data, and provides links to resources and tools for metadata creation and validation.

6.1. Does metadata comply with EDMC Data Documentation directive?
No

6.1.1. If metadata are non-existent or non-compliant, please explain:
Missing/invalid information:
- 1.7. Data collection method(s)

6.2. Name of organization or facility providing metadata hosting:
NMFS Office of Science and Technology

6.2.1. If service is needed for metadata hosting, please indicate:

6.3. URL of metadata folder or data catalog, if known:
https://www.fisheries.noaa.gov/inport/item/57610

6.4. Process for producing and maintaining metadata
(describe or provide URL of description):
Metadata produced and maintained in accordance with the NOAA Data Documentation Procedural Directive: https://nosc.noaa.gov/EDMC/DAARWG/docs/EDMC_PD-Data_Documentation_v1.pdf

7. Data Access
NAO 212-15 states that access to environmental data may only be restricted when distribution is explicitly limited by law, regulation, policy (such as those applicable to personally identifiable information or protected critical infrastructure information or proprietary trade information) or by security requirements. The EDMC Data Access Procedural Directive contains specific guidance, recommends the use of open-standard, interoperable, non-proprietary web services, provides information about resources and tools to enable data access, and includes a Waiver to be submitted to justify any approach other than full, unrestricted public access.

7.1. Do these data comply with the Data Access directive?
Yes

7.1.1. If the data are not to be made available to the public at all, or with limitations, has a Waiver (Appendix A of Data Access directive) been filed?

7.1.2. If there are limitations to public data access, describe how data are protected from unauthorized access or disclosure:

7.2. Name of organization of facility providing data access:
NOAA National Centers for Environmental Information (NCEI)

7.2.1. If data hosting service is needed, please indicate:

7.2.2. URL of data access service, if known:
https://accession.nodc.noaa.gov/0157633
https://accession.nodc.noaa.gov/0204646

7.3. Data access methods or services offered:
Data can be accessed online via the NOAA National Centers for Environmental Information (NCEI) Ocean Archive.

7.4. Approximate delay between data collection and dissemination:
Unknown

7.4.1. If delay is longer than latency of automated processing, indicate under what authority data access is delayed:

8. Data Preservation and Protection
The NOAA Procedure for Scientific Records Appraisal and Archive Approval describes how to identify, appraise and decide what scientific records are to be preserved in a NOAA archive.

8.1. Actual or planned long-term data archive location:
(Specify NCEI-MD, NCEI-CO, NCEI-NC, NCEI-MS, World Data Center (WDC) facility, Other, To Be Determined, Unable to Archive, or No Archiving Intended)
NCEI_MD

8.1.1. If World Data Center or Other, specify:

8.1.2. If To Be Determined, Unable to Archive or No Archiving Intended, explain:

8.2. Data storage facility prior to being sent to an archive facility (if any):
Pacific Islands Fisheries Science Center - Honolulu, HI

8.3. Approximate delay between data collection and submission to an archive facility:
Unknown

8.4. How will the data be protected from accidental or malicious modification or deletion prior to receipt by the archive?
Discuss data back-up, disaster recovery/contingency planning, and off-site data storage relevant to the data collection

The image analysis data is captured in two different locations: exported files from CoralNet that reside on a file server, which are ingested into a table in the PIFSC Oracle database. Both the file server and PIFSC Oracle database are maintained and regularly backed up by PIFSC ITS.

9. Additional Line Office or Staff Office Questions
Line and Staff Offices may extend this template by inserting additional questions in this section.