**CRCP Project #840**
**American Samoa Department of Marine and Wildlife Resources**
**Data Management Initiative**

Deliverable Objective:

Compose a data management plan for each monitoring program of the American Samoa Department of Marine and Wildlife Resources (DMWR) that will thoroughly document the data collection, current data management practices, and outline any data management requirements. The document will serve as the first step toward realizing a standardized approach to coral reef monitoring data management within the jurisdiction and will mark substantial progress in both data standardization and data integration along with eliminating redundant data management tasks. This deliverable is part of CRCP project #840, whose objective is to provide data management support for DMWR's data collection programs.

Deliverable Summary:

A face-to-face meeting was conducted in October 2014 with DMWR personnel to discuss their respective monitoring programs and their data collections. The information collected during these meetings is documented in two data management plans that were developed using the NOAA Data Management Plan Template. A separate recommendations document was also composed to outline several recommendations to establish a solid information technology infrastructure and address several gaps in the data life cycle.
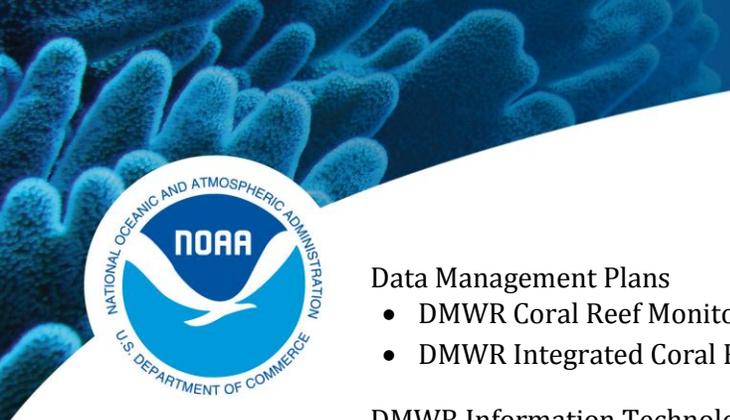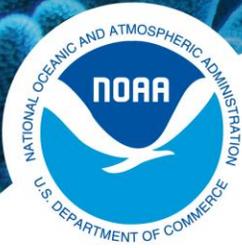
# Table of Contents

Data Management Plans

1. **General Description of Data to be Managed**
   - Name of the Data Collection
     - **American Samoa Department of Marine and Wildlife Resources Coral Reef Monitoring Program (ASCRMP)**
   - Keywords that characterize the data
     - **Marine Ecosystem**
     - **American Samoa**
     - **Tutuila**
     - **Reef Fishes**
     - **Coral Reef**
     - **Invertebrates**
     - **Benthic Cover**
   - Summary description of the data to be generated (abstract)
     - **Assessment of coral, macro-invertebrate, and fish populations and habitats in American Samoa for effective management and monitoring funded by the Coral Reef Conservation Program (CRCP) through the Coral Reef Advisory Group (CRAG);**
   - Anticipated temporal coverage of the data
     - **2005 – present**
     - **The survey methods were modified in 2013 to compliment the American Samoa Integrated Coral Reef Monitoring Program (ICRMP)**
   - Anticipated geographic coverage of the data
     - **American Samoa**
       - **Tutuila (12 permanent sites including 4 MPA sites)**
       - **Ofu & Olosega (opportunistically survey reef flats & pools)**
   - What data types will you be creating or capturing?
     - **Alpha-numeric data**
     - **Photos**
   - How will you capture or create the data?
     - **Diver visual observations**
     - **Underwater Camera**
     - **Photo analysis**

2. **Points of Contact**
   - Who is the overall point of contact for the data collection?
     - **Alice Lawrence, alicelawrence.mpa@gmail.com (fish)**
     - **Mareike Sudek, mareike.sudek@gmail.com (benthic)**
   - Who is responsible for verifying the quality of the data?
     - **Alice Lawrence, alicelawrence.mpa@gmail.com (fish)**
     - **Mareike Sudek, mareike.sudek@gmail.com (benthic)**
   - Who is responsible for answering questions about the data collection?
     - **Alice Lawrence, alicelawrence.mpa@gmail.com (fish)**
     - **Mareike Sudek, mareike.sudek@gmail.com (benthic)**
   - Who is responsible for data documentation and metadata activities?
     - **Alice Lawrence, alicelawrence.mpa@gmail.com (fish)**

- • **Mareike Sudek, [mareike.sudek@gmail.com](mailto:mareike.sudek@gmail.com) (benthic)**
- • Who is responsible for the data storage and data disaster recovery activities?
  - • **Alice Lawrence, [alicelawrence.mpa@gmail.com](mailto:alicelawrence.mpa@gmail.com) (fish)**
  - • **Mareike Sudek, [mareike.sudek@gmail.com](mailto:mareike.sudek@gmail.com) (benthic)**

## 3. Data Stewardship
- • What quality control procedures will be employed, or are employed?
  - • **Data is entered into an Excel spreadsheet**
  - • **The observer should confirm the data in the spreadsheet against the physical data sheet, although there is no formal QC process**
- • What is the overall lifecycle of the data from collection or acquisition to making it available to customer?
  - • **Diver visual observations are captured**
  - • **The observation data is transcribed to an Excel spreadsheet saved locally on the observers computer**
  - • **Photos are taken under-water**
  - • **Photos are downloaded to the observers computer**
  - • **Photos are analyzed to calculate benthic percent cover using Coral Point Count for Excel (CPCe)**
  - • **The observation data, photos, and CPCe outputs are backed up to an external hard drive periodically**
  - • **The physical data sheets are archived at the Department of Marine and Wildlife Resources office. DMWR will look into scanning the data sheets**
  - • **The observation data, photos, and CPCe outputs are not archived**

## 4. Data Documentation
- • Which metadata repository will be used to document this data collection?
  - • **Metadata records have not been created for the data collection. Part of the FY15 CRCP American Samoa Data Management Initiative proposal is to create a metadata record for this data collection that will be submitted to the NOAA Coral Reef Information System (CoRIS).**
- • In addition to discovery-level metadata, what additional metadata or other documentation is necessary to fully describe the data and ensure its long-term usefulness? How will that metadata be collected and updated? Is there a requirement to document this data collection in other metadata repositories?
  - • **There is no additional metadata**
  - • **Future updates to any existing metadata record will be supplied to CoRIS and serve to replace the existing metadata record**
- • What standards will be used to represent data and metadata elements in this data collection?
  - • **The metadata record will be formatted to comply with ISO-19115 in accordance with the NOAA-standard**

5. **Data Sharing**
    - Will the data be made available to the public? If so, what is the expected date of first availability? Is this a one-time data collection, or an ongoing series of measurements? Will there be a Principal Investigator hold or other delay between data collection and publication, and if so for how long? [Note: the Data Sharing for NOAA Grants Procedural Directive provides useful guidance for sharing data in a timely manner.]
        - **The data is not currently available to the public**
        - **In the future, the data should be available once the data has been entered and the QC process has been completed**
    - Will users be subject to any access conditions or restrictions, such as submission of non-disclosure statements, special authorization, or acceptance of a licensing agreement?
        - **Once the data becomes available, requestors will be subject to a (still to be developed) DMWR-specific data sharing agreement**
    - What data access protocols will be used to enable data sharing? The use of open standard, interoperable, non-proprietary web services is recommended (for example, OPeNDAP, or Open Geospatial Consortium (OGC) web services).
        - **Data will be downloaded as a comma-delimited file (csv) with corresponding XML metadata records**
        - **Photos will be provided in JPEG format**

6. **Initial Data Storage and Protection**
    - Where and how will the data be stored initially?
        - **DMWR Office: Physical data sheets housed at DMWR**
        - **Scientist's Computer: Observation data (Excel spreadsheets), Photos, CPCe output files (.cpc and .csv)**
    - How will the data be protected from accidental or malicious modification or deletion? Discuss data back-up, disaster recovery/contingency planning, and off-site storage relevant to the data collection.
        - **The observation data, photos, and CPCe outputs are backed up to an external hard drive periodically**
        - **DMWR will procure a server and work with an IT contractor to setup and administer the server as well as define a data backup schedule and a disaster recovery plan**
        - **The American Samoa Community College has the capabilities to serve as an off-site storage facility and has expressed partnering with DMWR in a disaster recovery plan**
    - If there will be limitations to data access, how will these data be protected from unauthorized access? How will access permissions be managed? What process is to be followed in the event of unauthorized access?
        - **Access to the server will be restricted to DMWR staff**
        - **The server will implement user-based file and folder permissions defined and managed by DMWR to prevent unauthorized access**

**7. Long-Term Archiving and Preservation**
- Will the data be archived and preserved with a NOAA Data Center (NODC, NCDC, NGDC)?
  - **There have been very preliminary conversations with NODC regarding data discoverability, accessibility, and archival. As of yet, none of the data has been archived with the data center.**

1. **General Description of Data to be Managed**
   - Name of the Data Collection
     - **American Samoa Department of Marine and Wildlife Resources Integrated Coral Reef Monitoring Program (ICRMP)**
   - Keywords that characterize the data
     - **Marine Ecosystem**
     - **American Samoa**
     - **Tutuila**
     - **Reef Fishes**
     - **Coral Reef**
     - **Invertebrates**
     - **Benthic Cover**
   - Summary description of the data to be generated (abstract)
     - **Assessment of coral, macro-invertebrate, and fish populations and habitats in American Samoa for effective management and monitoring funded by the US Fish and Wildlife Service**
   - Anticipated temporal coverage of the data
     - **Merged two existing monitoring programs in 2015**
       1. **Key Reef Species Monitoring Program (1996; 2002; 2005 – 2007; 2008 – 2012) that surveyed Tutuila (24 sites), Manua (~20 sites), and Swains (5 sites) for fish and benthic cover.**
       2. **Community-based Fisheries Management Program (2004 – 2008; 2010) that surveyed 11 villages around Tutuila for fish and benthic cover.**
     - **The survey methods of the merged program were modified to compliment the American Samoa Department of Marine and Wildlife Resources Coral Reef Monitoring Program (ASCRMP).**
   - Anticipated geographic coverage of the data
     - **American Samoa**
       - **Tutuila (18 permanent sites)**
   - What data types will you be creating or capturing?
     - **Alpha-numeric data**
     - **Photos**
   - How will you capture or create the data?
     - **Diver visual observations**
     - **Underwater Camera**
     - **Photo analysis**

2. **Points of Contact**
   - Who is the overall point of contact for the data collection?
     - **Alice Lawrence, alicelawrence.mpa@gmail.com (fish)**
     - **Mareike Sudek, mareike.sudek@gmail.com (benthic)**
   - Who is responsible for verifying the quality of the data?
     - **Alice Lawrence, alicelawrence.mpa@gmail.com (fish)**
     - **Mareike Sudek, mareike.sudek@gmail.com (benthic)**
   - Who is responsible for answering questions about the data collection?
     - **Alice Lawrence, alicelawrence.mpa@gmail.com (fish)**

- **Mareike Sudek, [mareike.sudek@gmail.com](mailto:mareike.sudek@gmail.com) (benthic)**
  - Who is responsible for data documentation and metadata activities?
    - **Alice Lawrence, [alicelawrence.mpa@gmail.com](mailto:alicelawrence.mpa@gmail.com) (fish)**
    - **Mareike Sudek, [mareike.sudek@gmail.com](mailto:mareike.sudek@gmail.com) (benthic)**
  - Who is responsible for the data storage and data disaster recovery activities?
    - **Alice Lawrence, [alicelawrence.mpa@gmail.com](mailto:alicelawrence.mpa@gmail.com) (fish)**
    - **Mareike Sudek, [mareike.sudek@gmail.com](mailto:mareike.sudek@gmail.com) (benthic)**

3. **Data Stewardship**
   - What quality control procedures will be employed, or are employed?
     - **Data is entered into an Excel spreadsheet**
     - **The observer should confirm the data in the spreadsheet against the physical data sheet, although there is no formal QC process**
   - What is the overall lifecycle of the data from collection or acquisition to making it available to customer?
     - **Diver visual observations are captured**
     - **The observation data is transcribed to an Excel spreadsheet saved locally on the observers computer**
     - **Photos are taken under-water**
     - **Photos are downloaded to the observers computer**
     - **Photos are analyzed to calculate benthic percent cover using Coral Point Count for Excel (CPCe)**
     - **The observation data, photos, and CPCe outputs are backed up to an external hard drive at periodically**
     - **The physical data sheets are archived at the Department of Marine and Wildlife Resources office. DMWR will look into scanning the data sheets.**
     - **The observation data, photos, and CPCe outputs are not archived**

4. **Data Documentation**
   - Which metadata repository will be used to document this data collection?
     - **Metadata records have not been created for the data collection. Part of the FY15 CRCP American Samoa Data Management Initiative proposal is to create a metadata record for this data collection that will be submitted to the NOAA Coral Reef Information System (CoRIS).**
   - In addition to discovery-level metadata, what additional metadata or other documentation is necessary to fully describe the data and ensure its long-term usefulness? How will that metadata be collected and updated? Is there a requirement to document this data collection in other metadata repositories?
     - **There is no additional metadata**
     - **Future updates to any existing metadata record will be supplied to CoRIS and serve to replace the existing metadata record**
   - What standards will be used to represent data and metadata elements in this data collection?
     - **The metadata record will be formatted to comply with ISO-19115 in accordance with the NOAA-standard**

5. **Data Sharing**
   - Will the data be made available to the public? If so, what is the expected date of first availability? Is this a one-time data collection, or an ongoing series of measurements? Will there be a Principal Investigator hold or other delay between data collection and publication, and if so for how long? [Note: the Data Sharing for NOAA Grants Procedural Directive provides useful guidance for sharing data in a timely manner.]
     - **The data is not currently available to the public**
     - **In the future, the data should be available once the data has been entered and the QC process has been completed**
   - Will users be subject to any access conditions or restrictions, such as submission of non-disclosure statements, special authorization, or acceptance of a licensing agreement?
     - **Once the data becomes available, requestors will be subject to a (still to be developed) DMWR-specific data sharing agreement**
   - What data access protocols will be used to enable data sharing? The use of open standard, interoperable, non-proprietary web services is recommended (for example, OPeNDAP, or Open Geospatial Consortium (OGC) web services).
     - **Data will be downloaded as a comma-delimited file (csv) with corresponding XML metadata records**
     - **Photos will be provided in JPEG format**

6. **Initial Data Storage and Protection**
   - Where and how will the data be stored initially?
     - **DMWR Office: Physical data sheets housed at DMWR**
     - **Scientist's Computer: Observation data (Excel spreadsheets), Photos, CPCe output files (.cpc and .csv)**
   - How will the data be protected from accidental or malicious modification or deletion? Discuss data back-up, disaster recovery/contingency planning, and off-site storage relevant to the data collection.
     - **The observation data, photos, and CPCe outputs are backed up to an external hard drive periodically**
     - **DMWR will procure a server and work with an IT contractor to setup and administer the server as well as define a data backup schedule and a disaster recovery plan**
     - **The American Samoa Community College has the capabilities to serve as an off-site storage facility and has expressed partnering with DMWR in a disaster recovery plan**
   - If there will be limitations to data access, how will these data be protected from unauthorized access? How will access permissions be managed? What process is to be followed in the event of unauthorized access?
     - **Access to the server will be restricted to DMWR staff**
     - **The server will implement user-based file and folder permissions defined and managed by DMWR to prevent unauthorized access**

**7.    Long-Term Archiving and Preservation**

- Will the data be archived and preserved with a NOAA Data Center (NODC, NCDC, NGDC)?
    - **There have been very preliminary conversations with NODC regarding data discoverability, accessibility, and archival. As of yet, none of the data has been archived with the data center.**

**IT Infrastructure Recommendations**
- Procure a server with existing CRAG funds based on the following requirements:
  - Purchase multiple hard drives to allow for enabling software or hardware RAID
  - Memory levels at or exceeding the recommended operating system standards
  - Final hard drive capacity should be at least 1TB (2+TB is preferable)
  - Server hardware configuration should allow for future hard drive expansion
  - Redundant power supplies
- Migrate users to an industry standard server
  - Ensure copies of Windows operating system and user software are properly licensed
- Procure a managed switch and join the diverse networks within DMWR, separating working groups onto differing subnets if desired for operational purposes
- Universally apply an industry standard anti-virus / anti-spyware / anti-phishing solution
- Isolate the network through an industry standard hardware firewall and insure Network Addressing Translation (NAT) is appropriately applied to the internet connection
- Fund continued maintenance and support of investment in IT infrastructure

**Data Management Recommendations**
- Scan physical data sheets into pdf files to eliminate the reliance on data sheets and provide redundancy in the event of a disaster
- Transfer DMWR files to the procured server to provide common file access to staff
- Define a backup schedule of data files and pertinent metadata documents on the procured server to protect against accidental or malicious data modifications or deletions
- Develop a disaster recovery plan, including but not limited to an off-site backup solution
  - Two Options for Off-Site Backup:
    - Hardware-based (USB drive swaps, tape backups, etc.)
    - On-line: may take advantage of the existing infrastructure at the American Samoa Community College (ASCC) that ensures the mitigation of most disaster threats. ASCC has offered 250GB of server space (at no cost) and with FTP access. DMWR would need to develop an automated process that transfers data to ASCC on a regularly scheduled basis and delegate the necessary resources to verify the data transfer.
- Establish a department-wide relational database on the procured server to house all of the data ever collected by the DMWR monitoring programs (including but not limited to site metadata, species metadata, observation data, and photo analysis data):
  - American Samoa Coral Reef Monitoring Program (ASCRMP)
  - Integrated Coral Reef Monitoring Program (ICRMP)
    - Key Reef Species Monitoring Program
    - Community-based MPA Monitoring Program
  - No Take Monitoring Program
  - NOTE: Any relevant analytical queries (ex: calculation for fish biomass or benthic percent cover) should also reside in the relational database to establish a "single source of the truth" for the entire department

- Develop a user-friendly, web-based data entry application to enter observation data and related metadata directly into the relational database with stringent hard and soft field-level validation criteria to ensure data quality
- Explore channels for data sharing and collaboration in American Samoa:
  - National Park Service
  - American Samoa Sanctuaries
  - American Samoa Department of Commerce: Marine Data Portal
  - NOAA Coral Reef Ecosystem Division (CRED)
- Compose data documentation (metadata) of survey methods and data collections that will reside with the NOAA Coral Reef Information System (CoRIS) to ease data sharing, comprehension, and usability
- Archive observation data, photos, and CPCe output files with the National Oceanographic Data Center (NODC) to address discoverability, accessibility, and posterity of DMWR data