SCCOOS DATA MANAGEMENT PLAN
March 2019

**Data Manager**

Vicky Rowley is responsible for management of data operations including specific timelines of data implementation.

SCCOOS DMAC personnel maintain regular communication with the U.S. IOOS Program Office through a variety of mechanisms including in-person meetings, phone calls & webinars, email conversations, and GitHub repositories. The continuous communication ensures that the DMAC team is aware of all new practices and protocols, as promulgated by the IOOC and the IOOS Program Office, and also understands how and when to implement them.

The SCCOOS servers and data management software are maintained by the SCCOOS DMAC team so data management protocols can be applied as soon as practicable, limited only by resource restrictions (personnel time, budget, or server capabilities). The SCCOOS DMAC team plans to implement all new protocols as soon as possible and within one year of adoption by IOOS. Once the new protocol is received by the DMAC lead, an implementation plan, including a more detailed schedule, will be outlined and executed by the DMAC team.

**Background**

The Data Management and Communications effort provides scientists, decision makers, and the public access to products and data services that will facilitate a scientific basis for research and management of ocean observations. As an operational Data Assembly Center, SCCOOS aggregates and manages the integration and distribution of ocean data and products. Serving as a federal repository for physical and bio-chemical datasets, the Data Assembly Center provides data ingest, analysis, quality control, discovery, access services, visualization and archive. Observational information is made available in a variety of data formats to ensure that products are useful and easy to access, while preserving the necessary detail to support the scientific and educational communities. SCCOOS continues to foster advances in new visualizations and technologies, making their products more comprehensible and ensuring timely and sustained dissemination.

**Goals and Objectives**

SCCOOS Data Management and Communications goals are to develop and maintain the following capabilities:

- Deploy the information system components (including infrastructure and relevant personnel) for full life-cycle management of observations including product creation, public delivery, system documentation, and curation.
- Deliver accurate and timely ocean observations and model outputs to a range of consumers including government, academic, private sector users, and the general public utilizing standardization and specifications common across all providers.
- Establish a robust data exchange that is responsive to a wide breadth of customer requirements and user feedback, assuring the data handling flexibility to accommodate decision-support requirements.

SCCOOS will continue to provide timely access to high-quality integrated data and support regional user needs while complying with the national standards and protocols for sharing and archiving data following the "Guidance for Implementation of the Integrated Ocean Observing System (IOOS) Data Management and Communications Subsystem". SCCOOS will also continue to integrate a broad suite of observations in the form of raw data and products. Collaboration with partner RAs to advance stakeholder access to cross-regional data services as defined by commonalities in specified user requirements, is on-going through IOOS Data Management and Communications including workshops and webinars. SCCOOS is well poised to improve ingestion with a level of quality control based on Quality Assurance of Real Time Oceanographic Data (QARTOD) parameters, standardize the web services for all parameters, submit to National Centers for Environmental Information (NCEI) for archive, and enhance data visualization. Additional desired outcomes are described on pages 18-30 of the Strategic Operational Plan.

**Work Plan**

*Equipment:* SCCOOS Principal Investigators maintain equipment inventories, shipping logs and instrument history logs for equipment owned and/or operated by SCCOOS. All SCCOOS instruments are calibrated, validated, operated, and maintained in accordance with manufacturer's guidance and as recommended by the lead researchers responsible for the equipment.

*Computing Infrastructure:* SCCOOS stores observational data on servers located at Scripp's CoLo data center. SCCOOS operates two enterprise servers running Redhat operating system. Each server has Xeon Processors, 96 GB of RAM and shares a JBOD with 12 Terabyte of storage. Servers are backed-up nightly and copies stored offsite at the University of California San Diego Supercomputer Center. Annual disaster and recovery hard drives are sent to the University of California Santa Barbara for additional backup.

Currently, in the event of a catastrophic failure at our primary data center, our WordPress-based SCCOOS and HABMAP websites, including their associated MySQL databases, as well as the SCCOOS ERDDAP and FTP servers, could be quickly recovered with minimal data loss by bringing these servers up in the cloud and restoring the associated data from backup. Computer systems are being migrated to Docker, with associated code in SCCOOS' BitBucket source code repository

*Data Ingestion*: Observations are collected from a variety of platforms in a variety of ways, each of which has its own level of data processing maturity. SCCOOS accepts data from both automated and manual systems via SFTP, HTTP, serial over internet, Webforms and email. Observations are stored in flat files, databases and NetCDF files, all of which are available on the SCCOOS website.

*Data Management*: Automated Shore Stations, Burkolator and Harmful Algal Bloom data are managed through SCCOOS. Gliders, Coastal Data Information Program (CDIP) waves, and High Frequency Radar (HFR) data are managed by 3 national Data Assembly Centers (DACs).

*Quality Assurance:* SCCOOS is regularly in accordance with manufacturer guidance or industry best practice for calibrating, validating, operating, and maintaining equipment owned and/or operated by SCCOOS.

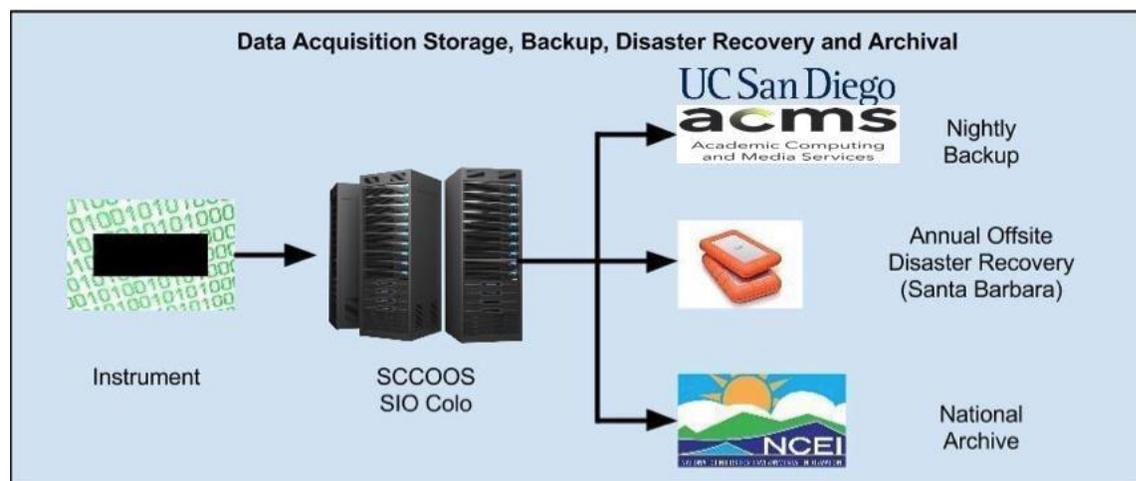***Quality Control****:* Publicly available data are quality controlled at minimum with the following QARTOD required tests: 1) timing, 2) syntax, 3) location, 4) range, 5) climatology. SCCOOS maintains an active role in the ongoing effort to develop and apply QC standards throughout the IOOS Regions. Highly leveraged programs such as CDIP wave data, HFR network, CalCOFI cruise data, and glider programs manage their own quality control which are compliant with the federal repository data management standards.

**Public Access and Dissemination**

All SCCOOS data are publicly available through the SCCOOS website. The HFR, glider, and CDIP wave observations are all transmitted to the National Data Buoy Center for inclusion on the World Meteorological Service Global Telecommunication Service. SCCOOS actively participates in on-going efforts to standardize data distribution through the use of web services such as the Sensor Observation Service and the Open-Source Project for a Network Data Access Protocol (OPeNDAP)/THREDDS. SCCOOS participates with IOOS partners in a highly distributed system of interoperable components. HFR, gliders, Automated Shore Station and CDIP wave buoys are registered in the IOOS Service Registry. Registration of datasets and services with IOOS provides basic monitoring of service availability and response time.

Collaboration with international programs such as OceanSITES and the Joint Commission of Oceans and Meteorology (JCOMM) are on-going instructive for CF compliancy and data handling issues.



**Picture 1.** Data acquisition, storage, disaster recovery and archival

**Data Sharing**

SCCOOS adheres to the NOAA Data Sharing Procedural Directive. All real-time and near real- time data managed by SCCOOS are freely available through open services, without delay or restriction. Avenues for accessing the data include the SCCOOS website: sccoos.org. SCCOOS does not restrict access to any data it collects or serves.
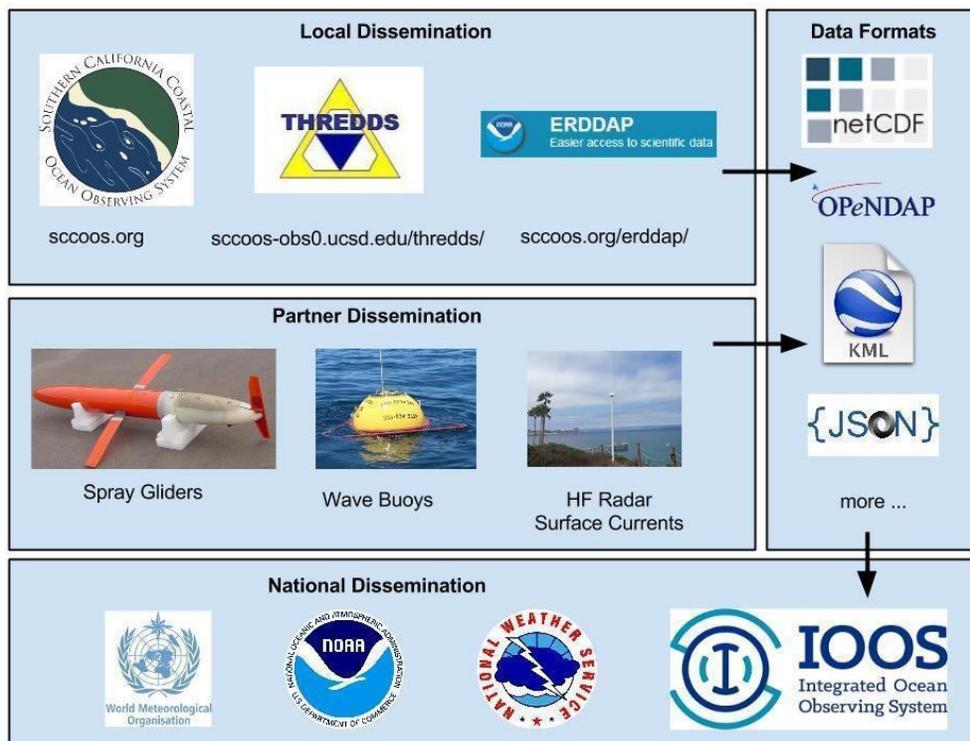
| | WMO GTS | THREDDS | ERDDAP | SOS | IOOS Catalog (DAC) | Web Page | Primary Funder | Operated | Maintained (QC, Archive) |
|---|---|---|---|---|---|---|---|---|---|
| Automated Shore Stations | | x | x | x | x | x | SCCOOS | SCCOOS | SCCOOS |
| Burkolator | | x | x | x | | x | SCCOOS | SCCOOS | SCCOOS |
| Gliders | x | x | x | | x | x | NOAA/SCCOOS | NOAA/SCCOOS | NOAA/SCCOOS |
| Harmful Algae Blooms | | | | | | x | SCCOOS | SCCOOS | SCCOOS |
| High Frequency Radar | x | x | | x | x | x | SCCOOS | SCCOOS | SCCOOS |
| | | | | | | | | | |
| **State & Federally Leveraged** | | | | | | | | | |
| Automated Information System | | | | | | x | USCG | USCG | USCG |
| CalCOFI | | | x | | | x | NOAA/State of CA/NSF | NOAA/State of CA/NSF | NOAA/State of CA/NSF |
| CDIP Wave Data | x | x | | x | x | x | USACE | CDIP | CDIP |
| Satellite Data | | | | | | x | NASA/JPL | NASA/JPL | NASA/JPL |

**Table 1:** Table indicates which communication; web services and access options are available for SCCOOS distributed data.

**Picture 2:** Real-time and near real-time data managed by SCCOOS are freely available through open services, without delay or restriction.

**Metrics**

The Google Analytics application demonstrates the high activity on the SCCOOS website. In 2017, there were 185,046 page views and 40,017 users.

**Data Portals**

In addition to the IOOS base funded tasks, SCCOOS continues to host project specific data portals. A few examples follow are:

- The West Coast ocean acidification data portal called, IPACOA http://sccoos.org/data/oa/.
- The San Diego Coastkeeper's Coastal Champion award winning data portal for University of California Environmental Health and Safety. The display focuses on the La Jolla Areas of Biological Significance site. http://sccoos.org/data/asbs/
- The Tijuana Plume Tracker webpage that displays the probability of plume interaction with the coastal zone. http://sccoos.org/data/tracking/IB/
- The integrated waves and HFR site for the Pt Mugu Navair. http://www.sccoos.org/projects/navair/
- The integrated harbors webpage waves, HFR, nautical charts and winds for the Ports of Los Angeles, Long Beach and San Diego. http://sccoos.org/data/harbors/

**Archive**

There are only two existing datasets that are not archived at NCEI at this time:
1.) HABs, will be archived by June 2019 at NCEI,
2.) Burkolator Ocean Acidification, SCCOOS is in discussion with Matt Biddle at NCEI and preparing to submit a request or archive by April 2018.

Presently, Automated Shore Station data, CDIP's wave data, HFR data and glider data are all archived at NCEI.

Certain datasets that are registered with the IOOS Service Registry contain FGDC and ISO 19115-2 metadata are available through ERDDAP and THREDDS.

**Additional Information**
**Individual Sensor Plans for Data Streams Managed by SCCOOS**

- Automated Shore Station
- Burkolator- Ocean Acidification
- Harmful Algae Blooms (HABs)

**Leveraged Programs Sensor Plans**

- High Frequency Radar (HFR) (IOOS DAC manages ingestion, quality control, and archive)
- Gliders (PI manages ingestion, quality control; IOOS DAC manages dissemination and archive)

**Leveraged Programs Sensor Plans (These programs have their own DAC which manages their own ingestion, quality control, dissemination, and archive.)**

- California Cooperative Oceanic Fisheries Investigation (CalCOFI)
- Coastal Data Information Program (CDIP)