NOAA Data Sharing Template
I.   Type of data and information created

What data will you collect or create in the research?
Contextual statement describing what data are collected and relevant URL (IOOS Certification, f 1. ii)
Manual shore stations consist of a suite of sensors that are attached to piers at several locations along the California coast. These manual stations measure temperature and salinity daily in the nearshore coastal ocean. These data can provide local and regional information on mixing and upwelling, land run-off, and algal blooms.

The SCCOOS Manual Shore Stations consist of:

1.)   Bodega Bay
2.)   Charleston/Coos Bay
3.)   Crescent City
4.)   Pacific Grove
5.)   Granite Canyon
6.)   Morro Bay
7.)   Neah Bay
8.)   Newport Beach
9.)   Port San Luis/Avila
10.)   Pt. Dume
11.)   San Clemente
12.)   Santa Barbara
13.)   Santa Catalina Island
14.)   Santa Cruz
15.)   Santa Monica
16.)   SE Farallon Island
17.)   SIO Pier
18.)   Trinidad Beach
19.)   Trinidad Bay

What data types will you be creating or capturing?
Manual Shore Stations collect temperature and salinity data at the surface and at the seafloor.

How will you capture or create the data?
Describe how the data are ingested (IOOS Certification, f 2.)
Data are ingested through manual sampling stations. After measurements are taken, various locations report their measurements.

Describe how data are managed (IOOS Certification, f 2.)
The data are managed at SIO/SCCOOS server. Once ingested, SCCOOS processes these data. The data are stored on disk in ASCII, NetCDF, and SQL formats. Back-up occurs offsite at the UCSD Supercomputer Center.

Describe the data quality control procedures that have been applied to the data. (IOOS Certification, f 3.)
QARTOD data quality control of In-situ temperature and salinity data will be implemented.
*(*http://www.ioos.noaa.gov/qartod/temperature_salinity/welcome.html*)*

If you will be using existing data, state that fact and include where you got it.
Manual shore stations supplement automated shore station measurements. One of the longest manual temperature records is at the Scripps Pier.

What is the relationship between the data you are collecting and the existing data?
N/A

II.   **Expected schedule for data sharing**

Adheres to the NOAA Data Sharing Procedural Directive. The System is an operational system; therefore the RICE should strive to provide as much data as possible, in real-time or near real-time, to support the operation of the System. (IOOS Certification, f. 4.)

How long will the original data collector/creator/principal investigator retain the right to use the data before opening it up to wider use?
Raw data are available in near-real time. Within 8 minutes of the collection, data are archived online and made publicly available on the SCCOOS website. In the near future, processed data will also be available utilizing quality control methods from the Quality Assurance of Real Time Ocean Data (QARTOD) manual, In-Situ Temperature and Salinity Data.

How long do you expect to keep the data private before making it available? Explain if different data products will become available on different schedules (Ex: raw data vs processed data, observations vs models, etc.)
Data are available in a raw format every 8 minutes. In the future, data will be processed and provided within a short period of time. Processing procedures are being developed but, processing will be expected to take a few minutes once the code is established.

Explain details of any embargo periods for political/commercial/patent reasons? When will you make the data available?
N/A

III.   Standards for format and content

Which file formats will you use for your data, and why?
How can the information be accessed? (IOOS Certification, f 1. ii)
SCCOOS shares data in a variety of file formats.
1. ASCII - Text that easily read and parsed by people and programs via the web. Available for download in Comma or Tab Separated Values.
2. OPenDAP - A data transport architecture and protocol widely used by earth scientists.
3. SOS - Sensor Observation Service is a mandatory service which is an integral part of IOOS.
4. NetCDF - A self-describing, machine-independent data format that support the creation, access, and sharing of array-oriented scientific data.

Raw data are written to a daily data log file. Once data is written into a raw data file, a separate process then parses the raw data into a mysql database. From the database, data are immediately available on the SCCOOS webpage. Additionally the raw data are save into Tab-Separated Values and NetCDF files. TSV files are available as hourly or monthly files. All TSV data are also zipped by station for convenient download. NetCDF file are made and the data is made available through the SCCOOS THREDDS server via OPenDAP and SOS protocols.

SCCOOS is in the process of updating its data flow. In the near future, once the raw data are acquired into a raw log file a raw NetCDF file will be made. From the original raw NetCDF file, further processing will occur applying quality control and flags and a final NetCDF will be made. The final NetCDF file will be made available on THREDDS and an ERDDAP server.

What file formats will be used for data sharing?
All of the Above.

What metadata/ documentation will be submitted alongside the data or created on deposit/ transformation in order to make the data reusable?
The metadata are available in multiple formats, ASCII FGDC, XML FGDC, and Supplemental metadata can be found here: http://www.sccoos.org/meta/browse .

ISO metadata is also available, on the following link: http://sccoos-obs0.ucsd.edu/thredds/SASS/catalog.html.

What contextual details (metadata) are needed to make the data you capture or collect meaningful?
Metadata details contain basic information regarding these data. In the future, processing steps will be described in detail providing better quality data.

How will you create or capture these details?
These details will be captured through quality control tests and techniques adopted from QARTOD manuals.

What form will the metadata describing/documenting your data take?

Metadata is available in FGDC and ISO formats.

Which metadata standards will you use and why have you chosen them? (e.g. accepted domain-local standards, widespread usage)
FGDC and ISO 19115 metadata are both accepted standards and mandated by the US Federal Government.

IV.   Polices for stewardship and preservation

What is the long-term strategy for maintaining, curating and archiving the data?
Points of contact- Individuals responsible for the data management and coordination across the region (CV's attached); (IOOS Certification f 1. i)
Julie Thomas - Employee 38 years, Principal Investigator/Program Manager
858-534-3034
jothomas@ucsd.edu

Darren Wright - Employee 10 years, Programmer/Analyst
858-534-3032
darren@ucsd.edu

Jen McWhorter - Employee 1 year, Administrative Analyst
858-534-3032
jen@cdip.ucsd.edu

Identify the procedures used to evaluate the capability of the individual (s) identified in subsection 997.23(f)(1) to conduct the assigned duties responsibly. (IOOS Certification, f 1. iii)
The University of California has a process in place for personnel evaluation. These evaluations are on file with UC San Diego Human Resource. All personnel listed have received excellent evaluations.

Which archive/repository/database have you identified as a place to deposit data?
Documents of the RICE's data archiving process or describes how the RICE intends to archive data at the national archive center (e.g., NODC, NGDC, NCDC) in a manner that follows guidelines outlined by that center. Documentation shall be in the form of a Submission Agreement, Submission Information Form (SIF) or other, similar data producer-archive agreement (IOOS Certification, f 6.).
SCCOOS manual shore station data are maintained, curated and archived at the Scripps Institution of Oceanography (SIO) and the University of California, San Diego. The SCCOOS long term strategy for maintaining, curating and archiving data involves sending data and metadata to the National Centers for Environmental Information (NCEI).

Data physically exist on three servers on two campuses. Recent and historic data are in several formats (raw, MySqlDB, TSV, ZIP and NetCDF) on two server at the Scripps Institution of Oceanography. A nightly backup of these data are stored on the University of California, San Diego Supercomputer Center.

What procedures does your intended long-term data storage facility have in place for preservation and backup?
Local redundant HDD storage on the SCCOOS Servers, the UCSD Supercomputer center, once the archive process is in place, NCEI.

How long will/should data be kept beyond the life of the project?
Data are indefinitely stored.

What data will be preserved for the long-term?
All data are publicly available and preserved.

What transformations will be necessary to prepare data for preservation / data sharing?
Raw data are decoded and formatted, analyzed and will eventually be quality controlled.

What metadata/ documentation will be submitted alongside the data or created on deposit/ transformation in order to make the data reusable?
FGDC standard metadata are available per deposit and transformation. NetCDF files have complete metadata and when implemented, will also have quality control flags.

What related information will be deposited?

V.   Procedures for providing access

What are your plans for providing access to your data? (on your website, available via ftp download, via e-mail, or another way)
Describe how data are distributed including a description of the flow of data through the RICE data assembly center from the source to the public dissemination/access mechanism. (IOOS Certification, f. 2.)
SCCOOS Access to Data (http://sccoos.org/data/overview/)
1. THREDDS data are organized into Archived and Realtime folders:
   a. OPENDAP - provides URL that can be used in Python/Matlab to automatically grab NetCDF file of data from server. Also provides option to download user-specified variables/timeperiods as ASCII or Binary file.
   b. HTTPServer - option to download the whole NetCDF file.
   c. NCML (NetCDF Markup Language) - XML document used to define a CDM dataset, and to allow user to add/delete/change metadata and variables, or combine data from multiple CDM files.
   d. ISO - XML metadata record for each station.
   e. UDDC (Unidata Data Discovery Convention) - tool to determine how well file metadata conforms to list of recommended metadata attributes.
   f. SOS - web service interface which allows querying observations, sensor metadata, and representations of observed features. Defines means to register/remove sensors and insert new sensor observations.
2. SCCOOS Request Data http://sccoos.org/query/ Returns SCCOOS data for web downloads.
3. SCCOOS website
4. http://shorestation.ucsd.edu/active/index_active.html

Will any permission restrictions need to be placed on the data?
Current funding is provided by the National Oceanic and Atmospheric Administration (NOAA). When used for web displays and online resources, please provide a link to the SCCOOS homepage. For instance, in standard html:

Data courtesy of SCCOOS

For offline references, please choose the appropriate form from the recommended acknowledgements below.

  • Short form (figure captions, etc.)

"... data from SCCOOS"

  • Longer form (in text)

"...data were furnished by the Southern California Coastal Ocean Observing System."

  • Full form (acknowledgements at conclusion of papers, etc.)

"...data were furnished by the Southern California Coastal Ocean Observing System (SCCOOS), a regional partner of the United States Integrated Ocean Observing System (IOOS®)."

With whom will you share the data, and under what conditions?
Data are publicly available.

Will a data sharing agreement be required?
In general, a data sharing agreement will not be required. However, data should be properly acknowledged.

Are there ethical and privacy issues? If so, how will these be resolved?
N/A

Who will hold the intellectual property rights to the data and how might this affect data access?
The funding agency & the University of California, San Diego through a contractual agreement.

VI. **Previous published data**