Appendix F1.1. Data Stream Plan: Regional Partners, December 2023

1. Data and Information Types

A. Provide a contextual description of the data stream.

These datasets include regional data provided by local or state agencies, private companies supporting maritime activities in coastal waters, university projects, and research studies funded and conducted by local entities. Due to the nature of these datastreams, CeNCOOS is not able to provide archival services or quality control, see **Section 4.A** for more information. In this plan, we provide general information about how these regional partner data may be ingested into the CeNCOOS DMAC System.

These data streams are provided by the following regional partners:

Applied California Current Ecosystem Studies (ACCESS)
California Collaborative Fisheries Research Program (CCFRP)
California Cooperative Oceanic Fisheries Investigations (CalCOFI)
Multi-Agency Rocky Intertidal Network (MARINe)
Naval Postgraduate School
Partnership for Interdisciplinary Studies of Coasts and Oceans (PISCO)
Reef Check California (RCCA)

B. How many station locations are there for this data stream?

There following number of stations by partner are available through the CeNCOOS data portal:

Applied California Current Ecosystem Studies: cruise tracks
California Collaborative Fisheries Research Program: various sites
California Cooperative Oceanic Fisheries Investigation : 3
Multi-Agency Rocky Intertidal Network: various sites
Naval Postgraduate School: 4
Partnership for Interdisciplinary Studies of Coasts and Oceans (PISCO): 41
Reef Check California: various sites

C. What are the specific variables of the data.

Applied California Current Ecosystem Studies: Marine mammal and seabird abundance
California Collaborative Fisheries Research Program: Fish abundance
California Cooperative Oceanic Fisheries Investigations: Fish, egg, and larvae abundance.
Multi-Agency Rocky Intertidal Network: Fish abundance
Naval Postgraduate School:
wind_speed,toa_incoming_shortwave_flux,air_pressure,wind_from_direction,wind_speed_of_gust,air_temperature,dew_point_temperature,relative_humidity

Partnership for Interdisciplinary Studies of Coasts and Oceans (PISCO): sea_water_temperature
Reef Check California: Fish abundance

D. Provide information about the sampling platform or instrumentation.

The sampling platforms are variable among the regional partners and generally include:

Applied California Current Ecosystem Studies: human observation
California Collaborative Fisheries Research Program: human observation
California Cooperative Oceanic Fisheries Investigation: net tows
Multi-Agency Rocky Intertidal Network: human observations
Naval Postgraduate School- meteorological station
Partnership for Interdisciplinary Studies of Coasts and Oceans (PISCO)- sea water temperature stations
Reef Check California: human observation

2. Data Pathway

A. Is a data sharing agreement required?

Data are available publically.

B. In which format(s) was data received by CeNCOOS?

For meteorological or seawater temperature stations, the data were received from the originator's websites or by one-time file access.

C. How can the information be accessed?

The data are available through the CeNCOOS data portal, where it can be downloaded or explored through interactive visualizations. Specifically, data are available from two unique access points:
- File Downloads (CSV)
- ERDDAP

D. What file formats will be used for sharing data, if different from original?

The meteorological and seawater temperature data are shared as CSV and through ERDDAP via the CeNCOOS data portal. Data are also available for exploration in the CeNCOOS portals via interactive, graphical visualizations.

E. Describe how the data is ingested(e.g. the flow of data from source to CeNCOOS data portals) and any transformations or modifications made to share data in the CeNCOOS data portal.

Data are downloaded from the source to the CeNCOOS storage. Custom Java, Scala, and Python scripts are used to convert data formats suitable for internal and external interoperability services. Data are made available in the CeNCOOS portals through the access points and via graphic displays generated through internal JSON-format data requests from these services.

Graphic displays include a mapping service, customized interactive visualizations, and time-series plots of the unit values wherein each parameter is graphed independently. Back-end scripts handle the conversion of visualized data from CF standards to other, non-CF units that may be requested by the user. Data files may be downloaded by the user from the CeNCOOS data portal. A user request for a CSV file request pulls the data from the server cache. A user request for ERDDAP pulls data from the ERDDAP service using the same cache. For this data, no CF-standard names or units exist, therefore custom names of abundance_of_{scientific_name} were used.

Summary statistics generated within the interactive graphical displays may be requested by the user. Summary statistics may include minimum, maximum and mean values. Seasonal statistics, available on time series longer than 3 years, include mean, and 10th and 90th percentiles. Note: the number of points visually available to interactive users from the source data are limited when necessary using temporal binning, such as daily, weekly, monthly, seasonally and yearly.

F. What metadata or contextual information is provided with the data?

Data are shared in the CeNCOOS portals with descriptive project and file metadata describing the data and accompanying fields. Metadata are also available via ERDDAP: https://erddap.cencoos.org/erddap/index.html

G. Are there ethical restrictions to data sharing?

No

a. **If so, how will these be resolved?**

N/A

H. Who holds intellectual property rights (IPR) to the data?

The data provider(s).

I. Describe any effect of IPR on data access.

N/A

3. Data Source and Quality Control

A. Indicate the data source type (i.e. Federal, Non-Federal, University, State Agency, Local Municipality, Military Establishment (branch), private industry, NGO, non-Profit, Citizen Science, Private individual)

State, University, NGO, Private

a. If Federal data source, were changes applied to the data?

N/A

b. If Yes, describe any changes to the data that require documentation?

N/A

B. Indicate the data reporting type (e.g. real-time, historical).

Real-time:
Caltrans
Naval Postgraduate School

Historical:
Applied California Current Ecosystem Studies
California Collaborative Fisheries Research Program
California Cooperative Oceanic Fisheries Investigation
Multi-Agency Rocky Intertidal Network
Partnership for Interdisciplinary Studies of Coasts and Oceans (PISCO)
Reef Check California

C. If real-time, list the QARTOD procedures that are currently applied.

Two of the five required tests are currently applied: Syntax and Gross Range Tests. Refer to CeNCOOS Data Management System plan for details.

D. If real-time, list the QARTOD procedures that are planned for implementation.

Remaining required tests are planned for implementation within 12 months of certification.

E. What is the status of the reported data? (e.g. raw, some QC, incomplete, delayed mode processed but not QC'd)

QC by originator

F. Describe the data control procedures that were applied by the originator.

Contact the data provider for availability of QC information.

a. Provide a link to any documented procedures.

N/A

G. Describe the data control procedures that were applied by CeNCOOS.

Refer to Section 3.C of the CeNCOOS Data Assembly Center and Data Management Plan.

a. **Provide a link to any documented procedures.**
N/A

H. List the procedures taken for data that could not be QC'd as directed.
N/A

4. Stewardship and Preservation Policies

A. Who is responsible for long-term data archiving?
CeNCOOS is not responsible for archiving for long-term storage of these data streams. Weather data provided from the Navy Postgraduate School was provided without permission to long-term archive. PISCO is responsible for and does archive all of the data it collected as part of the program. See PISCO Data Access for more information. CCFRP and RCCA data are archived with DataOne. Long term data archives are primarily with DataONE, NCEI and OBIS or specialized data archives as necessary.

B. Which long-term data storage facility will be used for preservation?
N/A

C. Describe any transformation necessary for data preservation.
N/A

D. List the metadata or other documentation that will be archived with the data.
N/A