

Data Management Plan

This project will produce various data products: image files; abundance and size data extracted from images files; isotope and SEM data from sample assays; outputs from physical oceanographic models; and K-12 lesson plans. If the project is funded, each type of data will be managed as detailed below by the PIs and Co-PIs and made available within two years of collection, as required by OCE's data policy.

Image files. Zooscan. PI Décima will oversee the archiving of images of *Pyrosoma atlanticum* and broad taxonomic zooplankton groups obtained using the Zooscan scanner from archived samples. They will be classified using Ecotaxa (<https://ecotaxa.obs-vlfr.fr/>) algorithms. All images will be uploaded and stored on Ecotaxa for general scientific use. The platform is free of charge and will become soon an international infrastructure hosted at IFREMER. PI Décima will also work with Scripps IT to set the data/images to back up to the current SIO-PIC servers. Raw data extracted from the images (zooplankton morphometrics) will be submitted to the Biological and Chemical Oceanography Data Management Office repository (BCO-DMO, <http://www.bco-dmo.org>). BCO-DMO ensures that a permanent archive of the data is maintained.

SEM. Unprocessed SEM images will be stored on the SIO-PIC local server, because Ecotaxa does not host these types of images, and available upon request. However, PI Décima will discuss options with BCO-DMO once the project is underway to identify other options for long-term storage on publicly accessible repositories. The raw data from the SEM images will be submitted to BCO-DMO.

Data. Zooplankton abundance and size data, stable isotope measurements, and SEM-based prey quantification will be overseen by PI Décima and submitted BCO-DMO. Metadata for type of assay (bulk, CSIA-AA, SEM-based prey composition) will be collated into a single file and supplied to BCO-DMO.

Regional Ocean Modeling System (ROMS) output. All ROMS model output is in NetCDF-4 file format that conforms to Climate Forecasting (CF) conventions (<http://cfconventions.org>) for metadata, and Common Data Model (CDM) (<http://www.unidata.ucar.edu/software/thredds/current/netcdf-java/CDM>) conventions for specifying coordinate systems and scientific feature types to facilitate interoperability with numerous client software tools used in the oceanographic community.

All model output will be supervised by co-PI Edwards, stored on a UCSC server for at least three years beyond the completion of the project. Presently, ROMS model output has been made available on the UCSC THREDDS data server (available through <https://oceanmodeling.ucsc.edu/>). An example of such a configuration can be found at <http://oceanmodeling.ucsc.edu/reanalccs13/#txtOutput> which links to https://oceanmodeling.ucsc.edu/thredds/catalog/wc12.0_ccsra13_01/catalog.html from which model output can be downloaded using the widely used Open-source Project for a Network Data Access Protocol (OPeNDAP). Model output calculated in this project will be made available through a similar THREDDS data server catalog.

Population Model Output: All MATLAB m-files created for the population model will be made available on a GitHub site and supervised by co-PI Prairie. Additionally, all model output will be submitted to the BCO-DMO site for this project.

K-12 “Phenomena-based” lessons: The NGSS-aligned curriculum resources will be shared through multiple networks. At the regional scale, the San Diego County Office of Education and San Diego Science Project will archive the curriculum in the ‘#ProjectPhenomena Database’ that is accessible online (<https://www.sdcoe.net/ngss/phenomena-and-the-ngss>) and promote its use quarterly. At the statewide scale, teacher leaders will share the lessons at the annual meeting of the California Association of Science Educators during Year 3. Additionally, the curriculum will be disseminated through the California Science Project and California Environmental Literacy Initiative.