

Data Management Plan

(1) Data and products from the observational component of the project:

Moffett's group will comply with NSF data policy which requires the collection and submission of all data resulting from our cruises. This includes hydrographic data, nutrients, dissolved and particulate iron and other metals, and iodine speciation. Moffett will also report the appropriate meta-data (data describing methods and protocols). Data will be submitted to the Biological and Chemical Oceanography Data Management Office at the Woods Hole Oceanographic Institution following their standards for data and metadata format and content. We will also report all of our hydrographic data to R2R (Rolling Deck to Repository) immediately upon completion of each cruise. We will follow GO-SHIP guidelines on how to produce good quality hydrographic data before submission. We will also take steps to ensure that our data are GEOTRACES Compliant. This includes submission of data and metadata to the GEOTRACES Standards and Intercalibration Committee and documentation for intercalibration activities. Moffett participates in ongoing activities to intercompare data for Fe and other metals under the auspices of GEOTRACES, including reporting of consensus standard values obtained by his lab to the Standards and Intercalibration Committee. There has not been an intercomparison of iodine speciation yet. It is anticipated that we will have conducted an intercomparison of iodine speciation methods with the Cutter (ODU) and Hardisty (MSU) labs during the award period so that our iodine data will be GEOTRACES compliant as well. From an educational perspective, the students involved in the project will be expected to become proficient in depositing data in these repositories and retrieving data submitted by others. In keeping with NSF policy, all data from this project will be submitted within 2 years of completion of the 2021 cruise and password protection will be removed as soon as the data are published.

(2) Data and products from the modeling component of the project:

Output from model simulations, which will include both physical (T, S, velocities) and biogeochemical (tracer, rates, fluxes) variables. The raw output will be post-processed and analyzed for presentation at meetings and for publication in scientific journals. Standard model fields (i.e. physical and biogeochemical tracers, current velocities, and the main biogeochemical rates) at monthly resolution will be stored at UCLA, and provided to, and made available through, BCO-DMO and NOAA NCEI in standard NetCDF format. In addition, the archival output fields will be stored at UCLA with daily-average values, suitable for use by others for independent analyses and for embedding local small-scale nests. These raw model outputs, which are generally too large and resource intensive to serve to an interactive portal, will be made available to the public upon request.

New model code and files for the ROMS-BEC physical-biogeochemical model, including: code for new biogeochemical components (iodine, benthic-release tracers) and revised Fe-cycle formulation; new grid configurations, initialization and forcing fields. The code used for the project's numerical simulations will be documented and made available via a public GitHub page as soon as results are published, as is our practice, and no later than two years after the end of the project. Peer-reviewed articles based on the model will link to a public GitHub page, which will host the version of the code used to produce the results.

Data standards and formats

Model output will be archived as self-contained NetCDF files, whose meta-data is readable via the NetCDF libraries with a wide range of software analysis tools. Model code will be released as FORTRAN (ROMS-BEC) and MATLAB code (analysis tools), and will include inline

documentation and “how-to” tutorial files.

Policies for data access and sharing

There are no confidential or private data expected in this study. The data syntheses, model code and processed model output will be stored on publicly accessible repositories (GitHub, BCO-DMO and NOAA NCEI) and linked by a project web page hosted at UCLA as soon as is feasible following processing or publication, and no later than two years after the end of the project. Due to the large size, raw model output will be archived on the PI’s storage server and made available upon request. All data, standard model output at monthly resolution, and raw model output will be released following the procedures indicated above, at time of publication of a peer-review article based on the data or model output, and no later than two years after the end of the project, whichever is soonest.

Data archive plan

All data and model output will be archived on public repositories as outlined above, and backed-up on PI Bianchi’s RAID server in the Department of Atmospheric and Oceanic Sciences at UCLA, where it will be preserved for as long as possible, but at least for two years following the end of the project period.