

DATA MANAGEMENT PLAN

Our project will provide for the preservation, documentation, and sharing of data collections, curriculum materials and other related research and education products using the guidance provided in NSF 17-037 OCE Data and Sample Policy and we will conform to the overall NSF philosophy of data management and dissemination, embodied in the NSF Award and Administration Guide (AAG) Chapter VI.D.4. The project will generate new data and incorporate existing data streams from ongoing projects. All significant findings from the proposed work will be promptly prepared and submitted for publication with authorship that accurately reflects the contributions of those involved.

1. Description of data types to be collected in the course of the project

Glider datasets will include digital profiles output by the vehicles' engineering and science sensors (as a function of time as well as sorted into pressure bins) and metadata including serial numbers, make, model, and calibration information. Water samples collected by shipboard methods will be analyzed for nutrient and oxygen concentrations, chlorophyll-a concentrations, O₂/Ar, and particle size / community composition from flow cytometry.

2. Data and metadata formats and standards

Glider datasets are stored in netCDF files which implement formats and standards set by the U.S. IOOS National Glider Data Assembly Center (NGDAC) (<https://github.com/ioos/ioosngdac/wiki>) including Climate and Forecast (CF) compliant conventions for metadata. For chemical and biological samples, the project will follow best management practices for metadata and data outlined by BCO-DMO, which are available at (http://www.bco-dmo.org/files/bcodmo/BCO-DMO_Guidelines.pdf). All data will be collected and archived using most commonly accepted data formats, including Excel spreadsheets, and where possible converted to data formats compatible with Ocean Data View (odv.awi.de).

3. Data storage and access during the project

Throughout active deployments, the gliders will transmit data streams to both BIOS and Rutgers University which will be processed through an automated system for immediate availability. Basic information and quick plots are posted on the Rutgers RU-COOL website in near real time (<http://marine.rutgers.edu/cool/auvs>). On a daily basis, BIOS will send NetCDF files to the NGDAC which distributes them via web servers and the Global Telecommunications System (GTS), and archives them at the National Centers for Environmental Information (NCEI). NetCDF and Matlab files will also be posted and updated daily on the MAGIC program website for public access.

4. Mechanisms and policies for access, sharing, re-use, re-distribution

All data produced as part of this grant are intended for publication in peer reviewed journals and will be made available to the public without reservations, restrictions, or limitations. Publications will cite requisite information, references, and accession numbers to allow public access to data. Glider data will be publically disseminated through the U.S. IOOS NGDAC (ERDDAP and THREDDS servers), GTS, and the program website. Biological and chemical data from field samples will be made available through BCO-DMO.

5. Plans for archiving data, samples, and other research products, and for preservation of access to them

For short- and long-term storage of raw data, images, and data files, as well as their derivative products and downstream analysis and work, the data will be backed up locally in institutional long-term storage. Ship-based hydrographic data acquired by the BATS program will be archived at BCO-DMO and NCEI as part of their routine. Grundle will be responsible for submitting all ancillary biological and chemical data collected by this project to BCO-DMO as well. Curry will assume responsibility for glider data processing, quality control and submission to the US IOOS NGDAC and NCEI.