

DATA MANAGEMENT PLAN

Data management will be coordinated by the Principal Investigator. Data management will adhere and comply with data policies as described in the NSF Division of Ocean Sciences Sample and Data Policy document by following (i) timely submission of data to existing national repositories; (ii) data, results, and ancillary information made fully accessible to the public with minimum time delay; (iii) use of existing data standards and management systems as much as practicable; (iv) ensure data interoperability to the greatest extent possible; (v) employ best practices for data policy and data management. Observational data will be archived within three years of collection and made available to the scientific community, as required. Model files will be archived contemporaneous with publication based on model results. Below we describe the types of data, and more specific plans for archiving.

DM-1. Data Types

The work plan for the proposed efforts includes the collection of field sediment and water samples, field instrumental data from CTD, lander, profiler, field experiments, laboratory water and sediment analyses, field and laboratory notes, and existing data from a range of sources. We will also develop and run models of hydrodynamics and biogeochemistry. None of the data collected or generated during this project will present novel challenges for purposes of archiving and distribution.

DM-2. Physical Samples

Sediment cores will be collected in the field and then be either measured *in-situ* or subsampled into 1-cm slices. Water samples collected in the field will be either filtered or stored in water bottles, and then transferred back to the labs at LSU. No cores will be transferred back to LSU. Sediment and water samples will be labeled appropriately and archived at the location of the PI conducting the analyses in LSU *Energy, Coast and Environment Building*, either under fridge/freezer or at room temperature, as appropriate. We anticipate that all physical samples will be consumed for analyses during the course of the project.

DM-3. Lander, Profiler and CTD Data

A variety of field and lab digital data will be generated, including lander, profiler, CTD and other data. Field and laboratory notes will be scanned as PDFs either after field operations, or monthly in the labs, and archived with other digital data as described below. All data will be reviewed and transferred to redundant storage drives as soon as is practical after collection by the team responsible for collection, for initial quality control and data security. Laboratory data will be reviewed on a file-by-file basis by the researcher conducting the analysis, and transferred to redundant data storage as soon as is practical after analyses. These data will be stored locally at each participating laboratory on computers with backup capability, and will be collected quarterly for storage on a secure data FTP server at LSU Center for Computation and Technology (CCT) with backup function.

DM-4. HydroBioSed Model Input and Output

The open-source hydrodynamic / sediment transport model is available from the model developers (Moriarty et al. 2017), but the proposed effort will require modifications to account for carbon-based tracers. During model development, Harris lab routinely uses either SVN or GitHub repositories for version control and to facilitate sharing of code between researchers. The lab will develop model input files including initial conditions, bottom sediment characteristics, and forcing and water column conditions. Model output will be saved to represent ~hourly conditions including velocity, geochemical concentrations, sediment data, output from the turbulence closure, and changes in bed geochemistry. Model input and output include text files and NetCDF files (see below). These will be made publically available (see below) and both text and NetCDF files can be easily read using freely available viewers and toolboxes.

DM-5. Results and Data

DM-5.1. Data Access and Sharing

All data generated by this project will be made available to the scientific community. We will publish the data (and related interpretations) generated in this project in peer-reviewed scientific journals and associated supplemental data repositories, such as Interdisciplinary Earth Data Alliance. Theses and dissertations that result from this project will be made available by the libraries at LSU and College of William & Mary. Metadata will be available immediately after data are archived, and observational data files will be openly and freely available on the web within two years from date of collection. **Model files** will be stored on the College of William and Mary's "W&M Publish" archive, and will be uploaded contemporaneous with publications based on model results. W&M Publish is freely available, and provides a DOI number so that users can cite the data source. Maintained by library staff, this archive already holds a substantial collection of VIMS data. All **biogeochemical** and **routine hydrographic data** and meta-data will be submitted in a timely manner to the Biological and Chemical Oceanography Data Management Office (BCO-DMO) at Woods Hole, Massachusetts. The PI will work closely with BCO-DMO personnel to ensure all data generated and archived conform to the standards of the particular catalog. This also includes the reporting of all appropriate meta-data (data describing methods and protocols) in a timely manner. In addition to submitting data to BCO-DMO, Cai will also submit pCO₂, DIC, TA and pH data to the Carbon Dioxide Information Analysis Center (CDIAC) as have been done in the past by his lab. Cai has worked closely with CDIAC's data manager Mr. Alex Kozyr over the past decade to curate CO₂-related data from projects in the South Atlantic Bight, the northern Gulf of Mexico and the western Arctic Ocean. Upon completion of the cruise, the original underway data will be contributed by the vessel operator to the UNOLS central data repository managed by the Rolling Deck to Repository (R2R) project. Also, R2R will ensure that the original underway measurements will be archived permanently at NODC and/or NGDC as appropriate for the data type.

DM-5.2. Metadata Standards

Data collected in the field and laboratory on paper data sheets or instrument computers will be scanned or photocopied on the day that data are collected and entered into tabular spreadsheets. Textual/graphical documents will be made available in at least one of the following formats: plain text, PDF, Microsoft Word, or Microsoft Excel, owing to their widespread use in business and academia. Data generated from HydroBioSed model will be formatted in compact, portable, self-describing NetCDF files that conform to the standards used by the ROMS and oceanographic community. NetCDF files can be recognized and read with a wide array of open-source software.

DM-5.3. Data Storage, Archival and Preservation

Field and laboratory data like lander, profiler and CTD will be archived on a FTP server maintained by LSU Center for Computation and Technology. These data storages will be under the protection of LSU and will be modifiable only through a LSU login ID that utilizes strong password requirements. Access outside the campus network is done via a VPN and protected with LSU user IDs and passwords. Data encryption and host-based firewalls are available. At the end of the funding period, or after publication of peer-reviewed journal articles (whichever comes first), the subsidiaries of the Interdisciplinary Earth Data Alliance will be used for sensor-based (e.g. profiler) data preservation.

During development, the HydroBioSed numerical model results will be stored on VIMS file servers along with input files and the source code used. Once the model has matured to the point where it is ready for dissemination, the input files and results will be archived and publicly accessible through W&M Publish, contemporaneous with publication of results based on the model. Along with model source code, the specific input and output files from which peer-reviewed publications developed as part of this project will be saved as a "tagged release" and stored on W&M Publish.

REFERENCE:

Moriarty, J.M., C.K. Harris, et al., 2017. A model archive for a coupled hydrodynamic-sediment transport-biogeochemistry model for the Rhône River sub-aqueous delta, France. Virginia Institute of Marine Science, College of William and Mary. <https://doi.org/10.21220/V53P4Q>