

Data Management Plan

Co-PI Li will be responsible for managing the model outputs and data which he will use. He will assure that the project data are verified, stored, and backed up according to standard research practices.

Types of Data: We will implement the necessary data management systems to support the full data life cycle, with a focus on assisting project investigators in creating and sharing quality-controlled data and in implementing tools that assist project researchers and other data users in discovering, accessing, visualizing, and interpreting data required for their respective analyses. Data generated by Li's group are mostly the outputs from the numerical ocean models of the West Florida Shelf. Other data for this project will include 1) biological and chemical data collected from field sampling including an array of nutrient and carbon analyses and biological community characterizations (Chl *a*, pigments, microscopic analyses), and 2) experimental data focused on nutrient rate processes and 3) physical data characterizing the water column, including temperature, salinity, density, pH and DO. *Digital photomicrographs* will also be produced from laboratory and field microscopy.

Data and Metadata Standards: One goal of this project is to ensure that the data that we derive become community resources for understanding coastal resiliency in estuaries. As such, we will make full use of existing and emerging standards for sharing environmental data. We will also provide standard HTTP and FTP access to download datasets that are more easily used as files rather than services, but will ensure that published data are accompanied by appropriate metadata descriptions either delivered by the data service or accompanying the data download. Metadata descriptions for all datasets will conform to appropriate metadata specifications.

As samples are collected, processed and data are generated they will be loaded into spreadsheets; tabular data and metadata will be kept primarily on Excel or similar spread sheets. Microscope images will be stored as Tif files. Spreadsheets will be backed up in near real time using TimeMachine or a similar program. *Lab and field notebooks* containing the raw data generated by this project, detailed descriptions of procedures and methodological approaches, deviations from protocols, specific equipment, and chemical reagents utilized for this project will be cataloged by all project participants.

Policies for Data and Research Products: Most of our data for the modeling component of this project will be the outputs and derived datasets from model simulations. All model source code developed by this project will be open source and will be distributed under an open source license. We will use an open source code repository for our coding and development, which will enable us to coordinate our activities and to potentially engage developers and contributors from outside of the project team who wish to contribute to our model development activities.

Model outputs, derived datasets and research products will be made publicly accessible within two years of their generation. All published data and research products will include appropriate attribute and citation information (including mention of NSF sponsorship). Research results will be published in peer-reviewed journals.

Data from field collection and laboratory experiments will be submitted to the Biological and Chemical Data Management Office (BCO-DMO) where they will be publically available at the end of the project or after publication. All data generated as a result of research performed shall be the intellectual property of the PI overseeing the research. As such, the PI associated

Laboratory will retain an implied copyright for these data. Peer-reviewed papers will be written by PIs which present these experimental data, along with details of their collection and data interpretation. Additional southwest Florida monitoring (e.g. temperature, salinity, pH, Do, Chl a) data used by the PIs that is collected prior to, during or after completion of this project by federal and/or state supported monitoring programs in the region shall be included in data submissions to BCO-DMO after consultation with the appropriate state/federal agency. The majority of this data falls under state of Florida 'Sunshine' Laws and is publically available upon request.

Plans for Archiving Data: Our project will create a body of model outputs. Development of the required repositories, protocols, services, and methods for enabling shared access to the full range of anticipated data types, is part of the intellectual contribution of this project. We will build on existing data-sharing infrastructure at UMCES. All model outputs and derived datasets will be archived for a minimum of three years after the conclusion of the project.

Our project will create a body of scientific data, and geodatabases that will be valuable to research and management communities beyond our project team and into the future. Development of the required repositories, protocols, services, and methods for enabling shared access to the full range of anticipated data types, particularly to enable their efficient integration with the proposed models, is part of the intellectual contribution of this project. Primary datasets and other research products will be published, cataloged, and archived on existing enterprise production servers. All work of graduate students, including dissertation data, will be stored on these servers.

For physical sample archiving, project PIs have large freezers which are used to archive physical samples after initial analysis until the complete dataset is assembled. Any images will be submitted to Microbe Library <http://www.microbelibrary.org/>, a site created and hosted by ASM to generate a database of microbial images for use in undergraduate education and will also be submitted to Micro*Scope <http://starcentral.mbl.edu/microscope/>. All data will be archived as soon as it becomes available on storage devices located in each of the PIs labs and on their institutional servers, which are backed up both locally and remotely to ensure data security and prevent loss. In addition to on-line repositories, project results will be presented at scientific meetings and in peer reviewed publications.

Short-term storage and data management: During the project duration, data will be stored at PI labs or computers and backed up weekly or more frequently as specified by each institution.

Long-term deposition and preservation strategy. At the project end, or upon data publication, data will be transferred to BCO-DMO for public access and long-term storage.

Length of archival: Data will be kept in individual institution storage for at least five years, or until it has been successfully uploaded to and made publically available through NCEI or a nationally or internationally funded database specific to that data.

Documentation and Sharing of Data and Samples: All aspects of the research will be documented in both physical (Lab Notebooks) and digital formats. Data access and sharing will be communicated through all publications with the relevant data and data access. Logs will be kept by PI for all samples collected and analyses performed. The PI will collaborate closely on the final grant Report and on data management in their respective laboratories. Archived and frozen (-80 °C) sub-samples backup will be available to interested parties for their independent analysis following publication of results.