DATA MANAGEMENT PLAN

Data management will be coordinated by the co-Principal Investigators and carried out by all project participants. The plan encompasses three areas: use policies, standards, and data and sample preservation and archival. The PIs will leverage existing systems wherever possible, including LSU university archival resources, Dryad, and OBIS. Following NSF guidelines, all data will be provided to the Biological and Chemical Oceanography Data Management Office (BCO–DMO; www.bcodmo.org/) as a permanent repository. All data generated during the course of this project will be made available for general use or derivative reuse upon publication (within 2 years of project completion). The PIs will contact their project office to discuss best practices within the Directorate.

I. Products of Research

Data products will consist of two classes of information: (1) larval behavior, survival, and growth data collected during laboratory experiments; and (2) Biophysical simulation output (hydrodynamic fields, particle trajectories, connectivity matrices); A third product (not a data product) will include code developed for biophysical modeling.

Larval behavior, survival, and growth data – Arrays describing larval survivorship and behavior in various climate experiments will be archived by Gravinese. Data acquired from these laboratory experiments, including the experimental systems environmental conditions (i.e., carbonate chemistry, salinity, and temperature) will be stored in laboratory notebooks and will be immediately backed-up on digital hard drives. Experiments will adhere to the Guide for Best Practices in Ocean Acidification (Riebesell et al. 2011; https://www.oceanbestpractices.net/handle/11329/339). Gravinese will also preserve digital copies of the biological data on a cloud storage repository (Dropbox). Biological data sets will be stored as excel (.csv) files.

Biophysical data – Raw modeling outputs are stored as netcdf files for data access and archival purposes, and include all model parameters as metadata. Hydrodynamic fields (including current velocities, and carbonate chemistry) are stored as large 4-dimensional arrays. Connectivity matrices – across years – will be stored as 3-dimensional arrays. Particle trajectories over space and time will be stored in large multidimensional matrices. The estimated data size is large, on the order of terabytes (TB).

Co-PIs Holstein, Xue, and Gravinese have a track record of collecting similar data as evidenced by their publication record. The data collected will be retained by the co-PIs indefinitely.

II. Data and Sample Storage and Preservation

All laboratory and model-generated biological data sets and metadata describing state variables and laboratory conditions will be uploaded to networked servers maintained at LSU, stored on multiple on-site backup drives, and stored on a secure cloud-based server (i.e., Dropbox, Google Drive). Original notebooks and information pertaining to the experimental design will be secured by the Co-PIs in their campus offices or laboratories. Scanned electronic copies of the laboratory notebooks will stored as described above for other electronic data. Data will be transferred and archived at the Biological and Chemical Oceanography Data Management Office (BCO-DMO; http://www.bco-dmo.org/) following processing and will comply with all NSF and OCE policies. No data transformations will be necessary to prepare data for preservation or data sharing and data will be structured in EML (see below) with associated 'read me' files to make readily usable.

III. Data Formats and Metadata

Our data will be stored and disseminated as clean matrices and will be stored in netcdf (.nc), comma separated files (.csv), Matlab (.m), and shape files (.shp), with associated metadata files. The biological and ecological data will be structured in Ecological Metadata Language (EML). We will also include a "readme" that will explain variables, structure of the files, etc. for each of the data sets.

IV. Data Dissemination & Policies for Data Sharing and Public Access

Data will be made publicly available in public repositories and via publications no later than two years after the completion of the project. Foreseeable users of the data are scientists, modelers, and managers investigating the response of marine metapopulations to climate changes, fishing pressures, coastal development, and local and regional conservation decisions. We will use the Dryad (http://datadryad.org/) platform to make publicly available the raw cleaned data. Dryad allows other scientists to download and work data without any restriction. We will provide any software code developed through a GitHub (www.github.com) "project" which will be maintained by the Co-PIs, and by reasonable request.

Access to raw, unprocessed, data products will be provided by reasonable request via contact with the Co-PIs. During Years 3 & 4 and after project completion, the results will be prepared for presentation and sharing in journals and at conferences.

Lesson Plans and Curriculum

Subsets of data will be made available to professional educators for use in their classrooms via an open access peer-reviewed publication. Gravinese will disseminate the lesson to professional educators via educational partnerships through the non-profit Youth Making Ripples Program and educational list-serves that include the National Marine Science Education Association free of charge.

Roles and Responsibilities

All personnel funded through this project will be responsible for maintaining and updating individual data sets and for writing detailed metadata and read me files that outline all components of how, when and where a particular data set was collected. A fraction of the salaries requested will be used for data management activities, in particular data entry, QA/QC, calibration, metadata management, and conversion to standard format.

This project will not involve the acquisition of data from either vertebrate or human subjects. The data acquired and preserved in the context of this proposal will be further governed by the PI's respective institutions policies pertaining to intellectual property, record retention, and data management, as referenced by each institutions employee handbook. We do not anticipate any significant intellectual-property issues involved with the acquisition of the data.