

DATA MANAGEMENT PLAN

Primary Investigator: Houshuo Jiang, Donald Anderson, Michael Brosnahan

Institution: Woods Hole Oceanographic Institution

Project: The behavioral repertoire of dinoflagellates: high-speed, high-resolution imaging of ecologically important species-species interactions

NSF Division: OCE

Solicitation Info: PD 98-1650 - Biological Oceanography closing on 08/15/2015

Submission Date: 08/15/2015

Overview: Most data produced through this project will be in the form of HSMIS (High-Speed Microscale Imaging System) recordings, including: (1) time series of image files generated from high-magnification high-speed imaging of free-swimming dinoflagellates and species-species interactions in cell culture flasks, and (2) time series of flow velocity vector fields generated from high-magnification, high-speed, time-resolved microscale particle image velocimetry (μ PIV) measurements of the flow fields imposed by free-swimming dinoflagellates.

Additionally, the project leverages an ongoing field program supported through the NSF and NIEHS sponsored Woods Hole Center for Oceans and Human Health and other projects. Through these projects, the research team is deploying a custom built instrument platform at Salt Pond, a field site in the Nauset Marsh system of Cape Cod. The instrument platform produces continuous monitoring records of phytoplankton diversity and abundance as well as physical and chemical hydrographic conditions. All data collected at Salt Pond are copied to a local server at the field site and synchronized to two other servers in Woods Hole through a wireless Internet connection. Phytoplankton images are made available to the public in near real-time at <http://ifcb-data.whoi.edu/saltpond>. Additional measurements of nutrient concentrations and *Amoebophrya* spp. abundance will be made from water samples taken during weekly visits to Salt Pond. All sample processing will be completed at Co-PI Anderson's laboratory. Notes and results from these analyses will be recorded in laboratory notebooks and transferred to electronic files (e.g. Microsoft Excel, MATLAB) as appropriate and copied using our automated data backup systems.

Data description: We expect the total data volume from HSMIS recordings will be from several hundred GB to a few TB. The acquired data will be initially stored in hard drives of local PCs and in portable hard drives as backup storages. The initial data will be stored at our laboratories at the Woods Hole Oceanographic Institution and copied nightly through an EMC-networker based backup system at WHOI.

Data production rates from the Salt Pond observation raft are on the order of 1 GB/day. The raft is typically deployed in late March each year and recovered by the end of July, producing about 150 GB of images and over 10,000 CTD casts. Additional sampling results in counts and profiles of nutrient concentrations from weekly site visits.

Data analysis summary: From the original image data, we will retrieve various behavioral characteristics of free-swimming dinoflagellates and species-species interactions, including speed, body acceleration, turning rate, flagellar beat frequency, net to gross displacement ratio (NGDR), and reaction distance etc.

For the PIV flow field data, we will follow rather standard flow-field analysis procedures to calculate integrated quantities such as fluxes and spatial and temporal decay rates and to generate plots of velocity vectors and contours of various flow field quantities.

Field data will be analyzed according to our standard practices for monitoring as developed through our other projects. We anticipate the need to refine automated image classification software to optimize performance for some species (e.g. *Amphidinium carterae*, *Heterocapsa triquetra*, and *Protoperidinium hirobis*) that are not targeted through other ongoing projects but are known to occur in Salt Pond.

Includes field work? Yes

Expected data product #1

Data type: Observational, Experimental

Responsible investigator: Houshuo Jiang

Product description: All image and video data, and flow-field data.

Preservation plan: We will save all above described data into at least two backup copies using external hard drives over the course of the project. These hard drives will be regularly upgraded and stored in a secure location in our lab. Such a data archiving practice will be continued even after the project is over and for a long time (but subject to resource availability). We have been following the same data archiving practice for our current and past NSF funded projects. To make sure that our data will be accessible by the public from the Internet, we will use our project page <http://www.bco-dmo.org/project/637850> as the portal to our data. This webpage has been generated and currently maintained by the Chemical and Biological Oceanography Data Management Office (BCO-DMO). The webpage will include detailed descriptions of all the video data that we take and analyze for all published papers resulting from this project. The original high definition videos will be compressed into web friendly format and size for uploading to the webpage for public access. We will include contact information on the website for people who are interested in the original high definition video data to contact us. Upon receiving a data request, we will mail the requested high definition video data saved into DVD or Blu-ray disks to the requestor, for free. We realize it is currently impractical to make all the high definition video data (~5GB per event, and hundreds of GBs in total amount) accessible from the Internet directly. After the project is completed, we will build an additional project website through WHOI's webpage builder service. This webpage will mirror all the contents on the BCO-DMO page for our project and will be maintained for long-term access in case that the BCO-DMO page will no longer be available.

Timeline for data release: Immediate availability on request. Data will also be the primary basis of research manuscripts and other research products to be disseminated through national and international scientific meetings

Expected data product #2

Data type: Observational, Monitoring

Responsible investigator: Donald Anderson, Michael Brosnahan

Product description: All Imaging FlowCytobot images from Salt Pond deployments

Preservation plan: Triplicate storage on servers maintained by M. Brosnahan and WHOI CIS.

Timeline for data release: Immediate release (data hosted at <http://ifcb-data.whoi.edu/saltpond>)

Expected data product #3

Data type: Observational, Monitoring

Responsible investigator: Donald Anderson, Michael Brosnahan

Product description: All CTD profiles, also weekly profiles of cell abundance and nutrient concentrations at Salt Pond

Preservation plan: Triplicate storage on servers maintained by M. Brosnahan and WHOI CIS.

Timeline for data release: Immediate availability on request