**Data Management Plan**

The PIs are committed to the goal of making the highest quality data, metadata, models, model output, and summaries of research available to the scientific and management communities. The PIs have a history of sound and productive data management in projects supported by DoD, EPA, NOAA, NOAA-Sea Grant, USDA-NRI, and NSF to support this claim. Providing accessibility to data, metadata, models, model output, and research summaries will make the products from this project useful to researchers, students, coastal zone managers, and educational users as well as to the public.

*Types of data:* This project will generate the following databases:

*DataFlow and associated water quality data:* Data collected during DataFlow cruises will be entered into MS Excel workbooks. Locations and issues related to sampling will be noted in a comments section for each workbook. A protocol book for all analyses will be kept for the project. All data will undergo QA/QC and be available to the public via a project web portal within one year of collection.

*Experimental data:* Data collected during light/dark incubations, benthic metabolism experiments, sinking rate experiments, and DOC production and lability experiments will similarly be entered into MS Excel workbooks, with notes, protocol books, QA/QC, and public availability as described above.

*Sequencing data:* The sequence data obtained from this project will be deposited at the National Center for Biotechnology Information (NCBI). Additional sequence data will be available on the project website at VIMS.

*Microscopic and qPCR data:* Samples for microscopic and qPCR algal analyses will be labeled and tracked in sample logs. Each sample will have a unique 4-digit number and corresponding log sheet. Microscopic and qPCR data generated for each sample will be entered into MS excel spreadsheets, where the unique 4-digit label can be linked to other relevant metadata. Cell abundance data, generated from qPCR, will be calibrated against laboratory culture material from clonal isolate cultures of the blooming phytoplankton species, which are maintained at VIMS. Cell counts for each of the cultures will be determined by visual counts and DNA will be extracted from a known number of cells to use as positive control material and for generating standard curves by serially diluting the DNA to achieve a range of cell number equivalents that are reliably measured by the specific assay. A standard curve containing 5 points or more along with negative and positive controls are included with each qPCR assay run. Quantification is determined by using the automatic Ct and baseline settings. Dissociation curve analysis is performed for all SYBR® Green assays and the melting temperature of all positive samples are compared to the control for verification of identity.

*Ecosystem model:* The model developed from this project will be maintained on PI and student computers subject to automated daily backup to a local drive as well as a remote site. A user-friendly, online version of the model will be served on the VIMS Coastal Systems Ecology and Modeling Program website and routinely updated:

www.vims.edu/research/departments/bio/programs/semp/models/index.php

The user interface will describe the model and how to conduct scenarios, and will be complemented by a short, ~ 3-page downloadable user guide served on the same website. Model code will be made available upon request following peer-reviewed publication.

*Data and metadata standards*

All data will be stored in both raw and edited formats. Field notebooks will be stored and organized during each sampling cruise. Data entry into Microsoft Excel/Access software will be checked using both manual and automated techniques. Field data will be geographically referenced using hand-held GPS units. GPS data will be differentially corrected and stored as NAD83 UTM coordinates. We will follow the best management practices for metadata and data outlined by the Biological and Chemical Oceanography Data Management Office (BCO-DMO), which are available on line (http://www.bco-dmo.org/files/bcodmo/BCO-DMO_Guidelines.pdf). Measurements will be collated on spreadsheets and where possible converted to data formats compatible with Ocean Data View (odv.awi.de). The data will be available electronically to all researchers involved in the project, through a shared "Drop Box" among PIs.

The PIs will monitor the overall status of the project's information dissemination activities. Technicians, students, and PIs will maintain the highest standards for data quality assurance/quality control (QA/QC), maintenance, and access. The latest techniques and federal standards will be incorporated into the data collection and management program, from automated data entry and QA/QC programs to the use of software that provides for Federal Geographic Data Committee (FGDC) compliant metadata. The status of data management will be reviewed bi-monthly to ensure that communication and oversight are constant and accurate.

*Data archiving*

For this proposal Anderson will be responsible for coordinating submission of the geochemical, microbial, and rate data with BCO-DMO (*Biological & Chemical Oceanography Data Management Office*) at Woods Hole. If the project is funded, as a first step we will submit a data inventory form to the data archive. Metadata for each type of experiment and measurement will be collated into a single file. Once per year, when annual reports are submitted, the data specific to this project (including new data and any updates/corrections of previously submitted data) will be provided to the BCO-DMO. Upon completion of the project, the *final* data will be archived with BCO-DMO.

All data will be retained for at least three years beyond the award period, as required by NSF guidelines. 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 in lab repositories of the data generator. All the data are backed up in near real time using TimeMachine at VIMS, which has an institutional long-term storage.

*Data sharing*

All data produced as part of this project are intended for publication in peer reviewed journals and/or will be made available to the public without reservations, restrictions, or limitations. Presentations and peer-reviewed publications resulting from the proposed work will be documented on the respective web sites of the PIs. Archived primary data will become available to interested parties at BCO-DMO upon completion of the project as specified by NSF.