

## **Data Management Plan**

### **1. Types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project**

We will collect sediment samples and live species of *Nereis succinea* and *Clymenella torquata* from several sampling sites around Long Island, including Central Long Island Sound and Flax Pond. Frozen sediment samples will be archived; any remaining pore-water and fluid samples will be stored under refrigeration or frozen as outlined in the proposal description. We will collect observational, experimental and analytical data during the project which are obtained from sensors and instruments controlled through computer interfaces and linked to digital data storage. Experimental data to be collected include synchronized optode imagery, pore-water and overlying solute concentration and flux data, pore-water pressure data, and solid phase analyses data.

### **2. Standards to be used for data and metadata format and content**

Data management and sharing will follow the Division of Ocean Sciences Sample and Data Policy. We will follow established protocols, emphasizing careful lab book and spreadsheet management during the course of experiments and analyses. These efforts will include fully archived records of any data processing/calibration, which comprise essential metadata. Notebooks will be scanned on a regular bases and stored as Adobe PDF. All results will be archived in central spreadsheets (and backed up) and linked to raw analytical output and will be available for scrutiny on request. These data will primarily be in ASCII II and spreadsheet (xls) formats. Over the course of the project, all experimental data will be stored on at least 3 hard drives and one cloud based server; code developed for image analyses and modeling will also be backed up both locally and offsite. Data gathered from optical imaging sensors will be stored as raw b16 files, grey scale TIFF and false color PNGs. Data will be calibrated and assembled to self-documenting NetCDF (Network Common Data Format) files. This data format allows “stacking” 2-D data matrices (e.g., images) and combine the corresponding “data cubes” with supplementary data (e.g., time-stamps, metadata). This data format is thus ideal for imaging based data sets and has been accepted by the Biological and Chemical Oceanography Data Management Office (BCO-DMO) on a past project. Data will be uploaded to BCO-DMO data storage systems within 2 years of collection. PI Volkenborn will be responsible for data management and he has used NetCDF files and the data storage system of the BCO-DMO in the past.

### **3. Policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements**

We will file annual/final reports to the NSF describing our progress on experiments and analyses and results. We will openly share all our sampling and analytical procedures informally and via publication. The data will be broadly disseminated at seminars, workshops, international and national scientific meetings, and in peer-reviewed international scientific journals identified using Digital Object Identifiers (DOIs) and Open Researcher and Contributor IDs (ORCID). It is and will remain our policy to publish all our data (with full metadata and processing details available on request). For example, we are careful to publish absolute data rather than just ratios that mask the absolute values. As is typical in our field, the data are included in tables inserted within the primary published text, or appended. In some cases, as dictated by journal policies and the volume of data, material will be placed in supplementary online material available formally through journal links or by request to us (or via posts on a lab website). Data availability prior to publication will be treated on a case-by-case basis, particularly if publication is not timely (we don't anticipate this). Regardless, we expect to share unpublished data with the proviso that they only be used (published, used in presentations, etc.) subject to our approval. Because we have a history of rapid publication, we do not expect data to remain unpublished beyond a reasonable period.

#### **4. Policies and provisions for re-use, re-distribution, and the production of derivatives;**

Not applicable.

#### **5. Plans for archiving data, samples, and other research products, and for preservation of access to them**

We have a long history of archiving our samples well and sharing them for subsequent research within and outside our group. Fragile samples are kept frozen or refrigerated when appropriate. Dried powders are carefully labeled and archived in our storage facilities. Outside access is welcome and available on request with the proviso that plans, products, and publication be discussed up front. In all endeavors, our goal is to be generous and open. Our data will be published, where they become freely available and reproducible (with proper citation). Related metadata, primary analytical output, processing details, etc., will be archived by the graduate students, PI Wehrmann, and Co-PIs Volkenborn and Aller AND in a central lab database (on a lab computer AND burned on CD or DVD). Copies of all publications will be submitted to the Stony Brook University Academic Commons and stored for permanent open access by the public. In the case of specific journals, embargo periods of 1 year may apply. Additionally, we will archive relevant data via the Biological and Chemical Oceanography Data Management Office (BCO-DMO) and make them publicly available via this tool. BCO-DMO also ensures that the data are archived properly at the appropriate National Data Center.