

Data Management Plan

Project Title: Collaborative Research: Calibration and application of vascular plant and aqueous microbial biomarkers to examine transformations of dissolved organic matter

Data Manager: Dr. P.J. Hernes (PI), University of California, Davis

All project data generated from field sampling will be submitted to the National Oceanographic Data Center (NODC). NODC is the United States facility established to acquire, process, store, and disseminate oceanographic data from the United States and other countries. The NODC operates as a component of the National Environmental Satellite, Data, and Information Service (NESDIS) of the National Oceanic and Atmospheric Administration (NOAA) of the U.S. Department of Commerce. Data from this project will include field measurements for transects conducted in the San Francisco Bay Estuary (SFBE) in year 2 (water temperature, specific conductivity, dissolved oxygen, turbidity) and discrete bulk and chemical constituent measurements from the SFBE transect and incubation studies (DOC, UV-Visible and fluorescence excitation-emission matrices, lignin phenols, fatty acids, amino sugars, and D/L amino acids). A separate database (Access or equivalent) will be constructed and made available with unknown compounds from the CuO oxidation technique for all incubations and transects, and will include mass spectra and retention time tags and queries for data processing. All data will be accompanied by detailed metadata files documenting relevant information about sample collection procedures, transport / handling procedures, analytical procedures, uncertainties, temporal and spatial domains, etc. Data will be described in accordance with NODC standards. The investigators will work closely with NODC curators to ensure accurate and complete documentation in accordance with the NODC designated level of service. The project will upload data files and collection-level metadata annually to NODC as soon as they have been quality controlled and processed. NODC makes the data publicly available through their system immediately upon receipt and subject to any approved embargo period. To facilitate tracking of reuse and fair credit to data providers, NODC will provide a recommended formal citation for the data set, including a persistent identifier.

Before submitting to NODC, these data will be collated by PI Hernes and organized in Microsoft Excel spreadsheets. While the individual laboratories that generate the various data streams will be responsible for maintaining records of data quality (standard curves, measures of analytical error, etc.), the collated data will also be screened for anomalies and the analytical error will be reported. Where possible, re-analyses will be completed to double check anomalous values. Outliers included in the final, submitted data set will be flagged to alert subsequent data users. Any subsequent changes to data after submission will be noted in the appropriate meta data files. Data files sizes will be on the order of tens to hundreds of megabytes.

Project investigators will work closely with NODC curators to provide all information necessary for data preservation in accordance with the NODC designated level of service.