Data Management Plan

The proposed research will primarily generate data from the analysis of samples or observations collected during the seasonal project cruises, which will be undertaken in project years 1 and 2 in conjunction with cruises of the NSF-funded Bermuda Atlantic Time-series Study (BATS). For the trace metal sampling, processing and analysis, we will use GEOTRACES program protocols and QA/QC procedures to ensure accurate and precise measurements, including the analysis of certified reference materials, where possible, and use of appropriate procedural blanks. In the Project Description we indicate the project member responsible for each specific measurement.

In advance of field work in project year 1, PIs and project participants (US and UK) will attend a 2-day planning meeting at Old Dominion University, which will be supplemented by regular teleconferences, to produce detailed implementation plans for survey observations, station locations, water-column sampling and water budgets, and to discuss modeling activities of the UK group. At sea, all sampling events will be recorded in a digital event log. Following each cruise, underway data will be contributed by the vessel operator to the UNOLS central data repository managed by the Rolling-Deck-to-Repository (R2R) project, which will ensure that the underway measurements are archived at the National Oceanographic Data Center. In addition, all primary CTD data (standard and trace-metal CTDs) and underway data (ADCP, thermosalinograph, meteorology, echosounder) will be immediately transferred to the BATS resident servers at BIOS and all primary data will be subsequently archived.

The trace metal data generated from the water column sampling (trace metal CTD and pump-collected particulate samples, as well as individual cell metal contents) will be integrated with supporting hydrographic (trace-metal CTD) data to create a merged seawater data set, which will be submitted by the project PIs to the Biological and Chemical Oceanography Data Management Office (BCO-DMO) for management and archiving within two years from acquisition/analysis, to be made available online from the BCO-DMO data system following standard NSF requirements; these data will also be submitted to the GEOTRACES data center at NODC. Sedwick, Buck and Twining have contributed similar data sets to BCO-DMO and the GEOTRACES data center from previous NSF-funded field projects.

The various hydrographic, chemical, biological and ship's data generated from the standard BATS program operations will be managed and archived following the established BATS protocols and will be available for our proposed project research (see letter of commitment from BATS PI's Johnson and Bates). The CTD data processing will typically be completed within one week of the cruise following analyses of salinity and dissolved oxygen. The data processing steps and QA/QC procedures are outlined in the BATS methods manual and will also be applied to the trace-metal CTD data. Following sample analysis of the BATS biogeochemical variables (2-4 months after cruise), data will then be plotted against each other and then individual data profiles will be plotted against QC windows (range of reported data). Final data quality checks will be performed using an interactive plotting routine of the probability distribution function of each parameter for each cruise at which point data will be assigned WOCE quality flags.

Every six months new QA/QCed BATS data (including the trace-metal CTD data from this new project) will be added to the database for download via the BATS FTP site. Additionally, BCO-

DMO at WHOI have a link to BATS in their user-interface and will update it as new data are deposited. As per previous conversations with BCO-DMO staff and their policy, when the archive timeline is triggered they will submit our supplied data to the appropriate archival site such as the National Ocean Data Center. Similar to BCO-DMO, staff at the CLIVAR and Carbon Hydrographic Data Office (CCHDO) routinely pull data from the BIOS BATS FTP site and subsequently convert the data to formats (WHP-Exchange and WHP-NetCDF) suitable for hosting on their web portal and OceanSITES.

The current BATS webserver is a cloud based FTP server using the Google cloud 'App Engine' which has excellent uptime (Google SLA states > 99.95%) and access speeds for community access. New file types other than traditional ASCII files will be added in the near future which will include Excel, CSV and Matlab. For the Matlab formats we will provide two systems whereby users can either download standard Matlab (*.mat) files for each cruise or one *.mat file for the CTD and bottle data for all cruises. For the latter, extraction scripts will also be provided such that the users have enhanced flexibility for data extraction on their host computer following the simple download.

Meteorological data from the Bermuda Weather Service will be downloaded to the BATS data server on regular intervals, while Mercator weekly forecast data will be accessed just prior to each cruise and Mercator hindcast data will be downloaded each month. All of the physical 1-D modeling will be performed on the BATS Linux workstation and results made available through the BIOS/BATS FTP portal.

All primary and processed data (CTD, trace-metal CTD, underway thermosalinograph and meteorology, ADCP and discrete biogeochemical samples) hosted at BIOS will reside on both the primary BATS Linux workstation (Dell PowerEdge R730, new as of Jan 2016) and secondary Linux workstation. Both servers support a RAID 1 data storage system which in turn is backed up at BIOS by the IT department using a Crashplan near real-time replication service. These systems are adequate and provide scalability for BATS and new data collected from this project through 2020.

The NEMO-PISCES model that will be used in this work is freely available under the CeCILL free software license. We will use version 3.6 of the NEMO model, including dynamic ligands and various modified versions of the PISCES biogeochemical model, which will be made available by PI Tagliabue to interested users.