

## **Data Management Plan**

The proposed work involves the collection of geochemical tracer data, in addition to the already planned measurements, within two NSF-funded research cruises (see below). All other data collected in these cruises is already covered under the approved Data Management Plans and is only briefly described below. In general, for both these funded projects and for the work proposed here, project investigators will conform to NSF's policy on dissemination of research results described in the NSF guidelines (see References). Within two years of acquisition, processed data will be submitted with documentation and metadata without restrictions to the relevant data center. This period is necessary for the processing of raw data and also allows the postdoc working with the data sufficient time to perform analyses and complete his research. Our data policy encourages open collaborative sharing of data, both between participants and with the general earth systems science community. Evidence of data submission for prior projects is provided in the NSF Prior statement.

The two funded cruises that this project relies on are:

- 2015 – Mooring recovery and hydrographic survey of Sermilik Fjord (PLR-1332911, Andres and Straneo, WHOI Co-PIs).
- 2016 – Mooring recovery, redeployment and CTD survey as part of OSNAP (OCE-1258823, Lozier, Straneo, Bower, Johns co-PIs).

CTD, lowered (Sermilik) and shipboard ADCP (OSNAP), and moored measurements were already planned as part of these projects and the processing and archiving of these data is described in the project-specific data management plans. Briefly, this involves processing and quality control with respect to national and international standards, where available, and submission to national Data Centers – primarily NODC, the Joint Archive for Lowered ADCP data (U. Hawaii) and ACADIS (Advanced Cooperative Arctic Data and Information Service (ACADIS, <http://www.aoncadis.org>) for the Greenland data.

### **1. Data Collected**

Approximately 200 samples, distributed over two cruises, for noble gases, isotopes (hydrogen, helium and oxygen) will be collected within this project. This will be in conjunction with the already planned collection of synoptic CTD and velocity measurements.

### **2. Data Analysis**

Sample analysis is done on a fully automated system that digitizes and records not only the mass spectrometric data but also key environmental properties (pressures, temperatures, etc.) critical to QA/QC of the results. The data are reduced and processed using custom VB.Net software and subsequently using MATLAB implementing sound statistical protocols (Glover et al., 2011). The reduced data are stored in an RDBMS (PostgreSQL) where it is linked to the pertinent (hydrographic and geographic) metadata. The raw data is automatically replicated on at least two different computer systems, one of which is backed up daily to a remote site. The RDBMS, which is situated on a separate dedicated computer, is also backed up daily to a remote site.

### **3. Data Archival**

Within the timeframe imposed by NSF (see above), all final data (noble gas and tritium concentrations and isotope ratios, and including a copy of the hydrographic data) and metadata will be submitted to both NODC and BCO-DMO ([www.bco-dmo.org](http://www.bco-dmo.org)). These will include links to the cruise and data reports for the two projects, which will be archived as a technical report at the WHOI/MBL Archive that is available to the public.

**References:**

Glover, D.M., Jenkins, W.J., Doney, S.C., 2011. Modeling methods for marine science. Cambridge University Press, Cambridge, U.K.

Division of Ocean Sciences Data and Sample Policy III. General Data Policy

[http://www.nsf.gov/pubs/2004/nsf04004/nsf04004\\_1b.htm](http://www.nsf.gov/pubs/2004/nsf04004/nsf04004_1b.htm)

Division of Ocean Sciences Sample and Data Policy

<http://www.nsf.gov/pubs/2011/nsf11060/nsf11060.pdf>