Data Management Plan

The Data Management Plan encompasses: types of data, standards, shared use policies, preservation and archival. We will work within ITGC guidelines and leverage existing systems wherever possible.

I. Types of data

A. Continuous data streams

i. **Ship's underway data** includes continuous **meteorological** data (air temp., wind, etc.), ***bathymetry*** (multi-beam), **ADCP**, and **sea surface hydrography** (SST, SSS, pCO_2); usually binned for every 1 minute. We propose 1 expedition, about 34 science days, total.

B. Semi-continuous data streams

i. **CTD**. Hydrographic profile data (at a specific time and location) includes pressure (depth), temperature, salinity, turbidity, etc., binned for every 1 m of depth below the surface. We expect 2–3 casts per day, with 30 science days over the expedition for ~100 casts. We also plan continuous "yo-yo" casts (up and down repeatedly) over several tidal cycles, which would generate an additional 50–100 casts over 24–48 hr. periods.

ii. **Gliders**. Sensor data will be collected continuously during deployment (several days to weeks). Sensors include location, depth, hydrographic and biogeochemical data.

iii. **Seal tags.** Sensor data will be collected continuously during deployment (several months to a year). Sensors include location, depth, hydrographic and biogeochemical data.

C. Discrete sample data

i. **Niskin bottles (part of CTD/Rosette).** Water and particulate samples associated with CTD cast. We will use conventional as well as trace metal clean bottles. Every Niskin fired will be assigned a unique sample identifier that will link every water sample collected to others as well as to the CTD data collected at the moment of firing (bottle file).

D. Outputs from numerical models

i. Results from the ocean ice-shelf model. Results from the Regional Ocean Modeling System (ROMS, *myroms.org*) include daily-averages for most physical and biogeochemical variables. The results are divided into monthly files each representing 5–25 GB of data.

II. Standards and formats to be used for metadata and data:

Metadata related to shipboard collection of samples will be managed by the cruise Chief or co-Chief Scientists. Use of standards or reference material will be described, and estimates of precision and accuracy included as part of the metadata. As much as possible, data will be archived in ASCII format, which is the most flexible and readable over the long term.

Trace chemical data will conform to the metadata standards established by the GEOTRACES International Data Assembly Center (*http://www.bodc.ac.uk/geotraces/data/policy/),* in consultation with BCO-DMO. In the case of ICP-MS trace metal data, calculated dissolved and particulate trace element concentrations will be reported to BCO-DMO along with detailed information on the number and type of blanks (e.g., analytical instrument blanks, digestion blanks, filter blanks, and process blanks that integrate the entire sample collection, digestion and analysis process). For seawater analyses, we will report data for NIST-traceable, certified and GEOTRACES reference seawater. We will also report concentrations of commercially-available standard particulate reference materials (PACS-2: Marine Sediments for Trace Metals, NRC Canada; BCR-414: Trace elements in plankton, Community Bureau of Reference, Institute for Reference Materials and Measurements, Belgium) digested and analyzed in the same batches as the particulate samples from the project expeditions.

Carbonate system analysis will be referenced to Certified Reference Materials (CRM) purchased from Scripps Institute of Oceanography (Dickson et al. 2007). Gas standards, if we need to use our own

underway pCO_2 system, will come from NOAA Earth System Research Laboratory Global Monitoring Division, Carbon Cycle Greenhouse Gases group (*https://www.esrl.noaa.gov/gmd/ccl/services.html*).

Results from the ROMS are stored in NetCDF format. This format can be accessed with software freely available on the Internet. Each file is self-documented and includes metadata (dimensions, units, etc.) for every variable following the Climate and Forecast (CF) convention.

III. Data access and sharing policies:

All field data collected will be available immediately to **all ARTEMIS team members** as well as **other participants in the larger ITGC project**, following initial quality control and quality assurance processing by individual investigators. **Metadata** will be submitted to the **Antarctic Master Directory**, via the USAP Data Coordination Center (*http://www.usap-data.org/*) as a Directory Interchange Format (DIF) entry. We agree to the NSF and NERC policy that requires investigators to make samples and data available to other researchers at no more than incremental cost and within a reasonable time. We are committed to making all data publicly available through peer-reviewed publications and public databases with as few restrictions as possible. Immediately following the expeditions, metadata (descriptions and locations of samples collected) will be shared with cruise leaders for inclusion in the cruise report. As samples and raw data are analyzed and processed, these products will be made available to all project collaborators via secure servers at University of Georgia, US, and University of East Anglia, UK. These servers are backed up regularly and can be accessed by collaborators needing the data. All data and derived data products collected will be posted with its metadata to **public access scientific databases** as required by NSF/NERC within two years of collection, or the end of the award, whichever comes first. Most of our data will be oceanographic, so databases such as the Biological and Chemical Oceanography Data Management Office (**BCO-DMO;** *http://www.bco-dmo.org*) are a good choice. NSIDC (*https://nsidc.org/data/submit/intro*) may also be an option. Ship data is sent directly to Rolling Deck to Repository (*http://www.rvdata.us;* R2R). We will work with the ITGC Project Management group or Program Officers to determine the best sites. St-Laurent will be responsible for handling all model data requests relative to the ice shelf-ocean model (ROMS), within and outside the research group. A single model realization represents 1-5 TB of information and thus exchanges with outside groups may be limited to subsets of the simulation (e.g. 3-monthly averages).

IV. Policies and provisions for re-use, re-distribution:

No proprietary data or products will result from the proposed research.

V. Plans for archiving and preserving data:

Oceanographic data will be archived long term as part of the data archives describe above (e.g., BCO-DMO). These resources have standardized file formats including relevant metadata (e.g. sample location, depth, age model), and the abstract of the original publication. Data archived with these electronic depositories are public domain, and readily available to the public and international scientific community. Individual PIs will be responsible for archiving their data, but Yager and Heywood will oversee the process.

Results from the ice shelf-ocean model (ROMS) that are used in publications will be archived at ODU's Center for Coastal Physical Oceanography (see "Facilities") for a period of five years following the end of the project. After this period, only the software configuration and input files used to generate the results will be available. Where the subglacial hydrology code is used in conjunction with the ROMS, those files will be archived with the input files.

Updates on the status of metadata and data archive will be provided in annual reports. Compliance with the Data Management Plan will be documented in the final project report. URL's for archived metadata and data will be included under "Products-Websites."